Redundancy as a Criterion for Multimodal User-Interfaces

Antti Pirhonen

JYVÄSKYLÄ STUDIES IN EDUCATION, PSYCHOLOGY AND SOCIAL RESEARCH 147

Antti Pirhonen

Redundancy as a Criterion for Multimodal User-Interfaces

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 1998

Editors
Paula Määttä
Department of Special Education, University of Jyväskylä
Kaarina Nieminen
Publishing Unit, University Library of Jyväskylä

Cover
Hannu Havusto, Fujitsu Computers Oy

ISBN 951-39-0382-6 (nid.), 978-951-39-4357-8 (PDF)
ISSN 0075-4625

Jyväskylä University Printing House, Jyväskylä
and ER-Paino Ky, Lievestuore 1998

To all who would like to make things
better – even if it takes more time

ABSTRACT

Pirhonen, Antti
Redundancy as a criterion for multimodal user-interfaces
Jyväskylä: University of Jyväskylä, 1998. 141 p.
(Jyväskylä Studies in Education, Psychology and Social Research
ISSN 0356-1070; 147)
ISBN 951-39-0382-6 (nid.), 978-951-39-4357-8 (PDF)
Finnish summary
Diss.

The study concerns information presentation in multimodal user-interfaces. A conceptual environment for analyzing information presentation is constructed on the basis of theoretical reasoning. The applicability of the model is empirically tested.

The central concept of the proposed conceptual environment is redundancy. Since the word has been used in diverse environments and has existing connotations, a functional definition for it is composed on the basis of earlier use and current needs. The result of this phase of the work is a conceptual environment in which redundancy is one extreme, referring to identical meanings among a set of user-interface elements.

Redundancy as a phenomenon in multimodal user-interfaces is analyzed in light of various theories and models concerning human cognition. The main concern in this part of the work is the human ability to process several messages simultaneously. On the basis of existing research, a suitable theoretical framework is sought for the efforts to avoid combinations of output elements in which messages interfere with the processing of each other. Each of the analyzed theories proved applicable in some respect even though only one of them was originally meant to be applied in a multimedia environment.

In the study, the proposed model is also tested in a concise experiment in which 14 multimedia design students analyzed 4 given user-interfaces on the basis of the proposed framework. The subjects worked in pairs, discussing the user-interface before them and filling in a form in which they were to classify single elements according to their role and then to classify the relationships that they found important. The data consisted of video of screen-events and synchronized audio information containing the discussions. The aim of this phase was to find out the usability of the proposed model. A phenomenographical analysis of the data showed that the proposed conceptual environment is easy to adopt and that its use in the analysis reveals features that would otherwise be unnoticed.

The study supports the use of redundancy as a central concept in user-interface design, providing that redundant combinations are not seen as an end in themselves but that the concept of redundancy is used in the analysis and creation of multimodal virtual objects consisting of multiple user-interface elements.

Keywords: Redundancy, Multimodal, User-interface, Cognitive models, Information presentation

Preface

On July 7 1972 at the University of Jyväskylä there was a public examination of a doctoral dissertation entitled "The effects of directive teaching materials on the affective learning of pupils". The defendant, Jouko Kari, presented work based on an interesting experiment whereby affective learning was investigated with sophisticated audio-visual material. In the terminology of today, Kari had created a complicated multimodal learning environment.

Today, Professor Kari is the supervisor of my work, in which multimodal information presentation is again the domain. It might be appropriate to consider what has changed in this domain since the early 70's. Kari had multiple modals available in the form of an audio tape recorder, a slide projector, a slide synchronizer, and a screen. Today, all that is in a compact package. Multimedia workstations support interactive applications and are spread all over the community, from homes to offices, unlike the heavy and complicated audio-visual equipment of the past. However, when one considers information presentation, the basic setting remains. The modalities available are the same. Today, there is much more quantity, whereas the quality is sometimes even worse. For example, the resolution of a typical picture in a multimedia application is much lower than that of an ordinary film. It seems that the requirements of quantity often override the requirements of quality.

A similar phenomenon, the continuously increasing requirement of producing *more* instead of *better*, can be seen in the academic world, too. The research work of today seldom allows one the chance of concentrating on one's topic, free from external pressures. In that sense, I have been privileged. I am grateful for that and would like to thank the persons who have made it possible for me to disengage myself from my regular employment as a teacher and be able to dedicate myself practically full-time for four years to research work and to prepare this thesis.

First I must express my gratitude to my supervisor Professor Kari who has followed and supported my studies almost entirely throughout the 90's. One of his most significant actions concerning my work was to introduce me to Jorma Niinivaara, the then Director of Fujitsu Computers. This led to a two-year cooperation with this gigantic computer manufacturer. I had an opportunity to participate in stimulating projects and to expose my theoretical reasoning to real application development. Jorma Niinivaara and his successor Heikki Roikonen arranged scholarships for me, which enabled me to concentrate on my research work without constant financial worries. For making it possible for me to keep my employment during this research period I would like to thank school principals Kari Fagerholm and Kari Asikainen for their flexibility.

Doing interdisciplinary research is frequently quite lonely work. When I present my work in order to get advice, people tend to first pay attention to the most unfamiliar part – to a computer scientist I appear a strange humanist, and to an educational scientist I am a technocrat. I am thus all the more grateful for the contacts I have made where my interdisciplinary approach has been treated with interest instead of suspicion. The support that I have received from my project group UCRET (User's Cognitive Resources Evoking Technology) has been especially important. Also, my cooperators in Fujitsu Computers have shown an open mind instead of dividing people into engineers and humanists.

I am fortunate to have been able to have a workroom in the building of Technology Education in the Department of Teacher Education, whose personnel have offered most inspiring and relaxing coffee breaks. The spirit of building E never fails.

Jyväskylä, December 15, 1998

Antti Pirhonen

CONTENTS

1 INTRODUCTION

In the past, the focus of application development was to implement programs in a way that optimizes the use of computational power of a computer. The form of interaction between a computer and a user was a secondary problem, because the form of information in the interaction process consisted mostly of simple characters. Additionally, the users were few and could be trained to use an interface that seemed cryptic for an inexperienced person.

The rapid growth in the computational power and the spreading use of computers since the 1970's has changed the focus of interest. The key question is no longer how to make the computer (or microprocessor) calculate as effectively as possible; the computations needed in usual applications, like word processing, are easy tasks for contemporary microprocessors. The focus has shifted from the computer to the whole that consists of a user, a computer, and a task (Card, Moran & Newell, 1983, pp. 403-404). The computer is usually only a tool that is supposed to help in performing a task. The co-operation between a computer and a human being should be seamless. The roles of a machine and a user when performing a task should be appropriate. The interaction between a human being and a computer should also support the performing of a task.

The quality of a user-interface is obviously seen as a key factor concerning fluent human-computer interaction. Because of the importance of this fluency, the quality of the user-interface also defines, to a large extent, the quality of the whole product (e.g., Brown & Cunningham, 1989, p. 1; Barker & King, 1993). The requirements for a high-quality user-interface have usually been formulated as a quality that is called *usability*.[1] The definitions of usability

1 The criteria for usability were discussed (Bennett, 1972) before the concept was named. After naming (Bennett, 1979), the concept has been defined and the definition has been

usually consist of a list of criteria. Different lists are quite analogous with the list that Shneiderman (1987) launched as a list of criteria for a high-quality user-interface:

- time to learn,
- speed of performance,
- rate of errors by users,
- subjective satisfaction, and
- retention over time.

In his book, Shneiderman proposes the *direct manipulation* principle as a way to success in user-interface design according to these criteria. In direct manipulation all the vital functions and information are easily accessed and manipulated via graphical virtual objects, and the consequences of a user action are instantly illustrated; a user manipulates physical objects via virtual counterparts (p. 201). Other concepts that characterize the same underlying attitude are *ease-of-use* (analyzed by Miller, 1971) and *transparency*. All these principles aim at making the user unaware of the existence of the user-interface itself. The goal is an illusion, in which the user has a feeling of complete and immediate control over the application.[1] In this kind of approach the user-interface is seen as a separate technical solution for delivering information between a user and a computer. When a computer program is used to perform a task, the functions of a computer are seen to consist of two kinds of functions. First, there are some kinds of core functions that are directly related to the content of a task. Second, there are user-interface functions that are only instrumental in order to use the core functions. (Some implementations of the approach are reviewed by Blattner & Glinert, 1996). The aim is to create software that would support spending as big a proportion of the human resources as possible on using the core functions. In other words, according to this view, the design of the software should support the user's concentration on the content of the task instead of the technical performance of it.

Laurel (1990, 1991) criticized the traditional method of separating the user-interface from the rest of the application and proposed her theatre metaphor for application development. Vertelney and Booker (1990) would even include hardware design in application development. On the other hand, the attempts to emphasize the role of cognitive models at the cost of domain specific expertise (Barnard, May, & Blandford, 1992; May, Barnard, & Blandford, 1992) can be seen as a quite contradictory effort. Construction of domain-free cognitive models is a basis of the efforts to create expert systems for the needs of design.

---

further developed (Bennett, 1984; Galitz, 1993; Shackel, 1984). There are slight differences among the definitions from different sources.

1 The directness of direct manipulation is really only an illusion. In fact, for example, executing a program via direct manipulation interface is technically a much more mediated and complicated procedure than via an "old-fashioned" command-line interface.

The problem of taking the user-interface as a separate part of an application is that in real life such a separation is not possible. When a carpenter is used to his regular hammer, the hammer has more than only instrumental value for him. Likewise, even if a user-interface is often seen as a tool to perform something important, it is an essential part of an application and the task-performing process. It cannot be changed or re-designed without affecting the use of the whole application. Therefore, the design of a user-interface cannot be separated from the content; nor can the criteria for the quality of a user-interface be content-free. It is not clear that the recent tendencies, like visual bias in information presentation, are always desirable in user-interfaces. Especially in education, properties of instructional media, content, context, values and many other complex properties are so firmly connected to each other that the idea of a standardized user-interface for all needs is indefensible (see Peled, Peled, & Alexander, 1992). For example, some educational studies have shown that the advantages of a graphical user-interface (GUI) and direct manipulation are questionable in problem-solving tasks (Cope & Simmons, 1994; Holst, Churchill, & Gilmore, 1997). Also, the power of animation has been questioned when used in instruction for a problem-solving task (Dyck, 1995). Although a GUI fulfills the criteria of usability, the results of its effectiveness in different tasks are contradictory (Brown & Schneider, 1992; Hazari & Reaves, 1994; Petre & Green, 1990). Even the communicative value of one basic element of a GUI, an icon, is probably overestimated (Griffin & Gibbs, 1992). The open structure of a GUI is a very important property concerning educational applications. The structure of an application affects, among other things, the role of a teacher (see Jackson, Edwards, & Berger, 1993). In other words, in a class, human-computer interaction is not the only interaction.

An attempt to separate the user-interface from the rest of the application is an attempt to separate the form and the content of an application. The discrimination of different aspects of an application becomes evident in the kind of application development projects in which domains of expertise and therefore domains of responsibilities concerning the project have strict boundaries. However, an application development is usually such a complicated project that it cannot be divided into independent sub-projects each of which would only require one kind of expertise. For example, if the tasks of a programmer and a specialist in ergonomics are separated, the latter would be forced to provide the programmer an atomic level description about the desired functions of an application. Since the same requirement would concern each pair of experts in the group in turn, the result would be an endless list of demands, many of which would probably be contradictory with each other. Skills in working as a team would, of course, be in a central role. But a truly interdisciplinary cooperation is hardly possible if the expertise of each domain is strictly divided. Not only should the project group be interdisciplinary in nature, but also each single member of the group should have – in addition to his or her own discipline – some expertise in other disciplines involved. For example, Card, Moran and Newell (1983) suggest the system designers become "the main agents to apply psychology" (p. 12).

Wickens and Carswell (1997) expressed the inappropriateness of separation of expertise in content and cognition by stressing the importance of domain knowledge of the psychologist. With this kind of policy, in which a project group consists of broadly educated participants, two important benefits are gained. First, communication is easier when each member of a group understands at least to some extent each others' disciplines. Second, many atomic level decisions can be made alone without consulting the rest of the group at every single turn. As a result, the quality of application is probably higher as the participants are better able to do creative work when they are competent and empowered to make decisions independently without an endless sequence of meetings.

In the current work, the level of abstraction is aimed at staying on a level that makes a truly interdisciplinary approach possible. This work does not anchor in the widespread conventions and standards when discussing a user-interface. Therefore concepts like usability and the related approach are questioned. User-interface is not handled as a distinguishable part of an application. Here, it is neither essential nor even desirable to be able to make a distinction between content and an interface. But in order to limit the domain of this work, *only the part of interaction that concerns information presentation is dealt with. It is a question of a message from a designer to a user, which implies both the meaning and the form*.

1.1 The current approach and research tasks

The usability approach usually leads to prototyping. The weak point of prototyping is that prototypes are based on conventions. When designing usable user-interfaces, usability testing of prototypes is only able to provide information about details. Prototypes seldom question the whole basis of the old models. Thus, the starting point for analysis of human-computer interaction is the interaction that has been implemented in the existing systems. However, when striving towards virtual worlds (Fisher, 1990), the models of interaction styles should not be taken from our interaction with already implemented computerized environments but from our interaction with our natural environments.[1]

In the present work, the focus is on the information flow toward a user. Referring to the idea presented above, the challenge is to outline principles that transfer essential qualities of receiving information from our natural environments to the design of artificial environments. The cornerstone of the approach is the simple fact that most messages we receive from our natural

[1] This ecological approach, originated from the ecological approach to visual perception (Gibson, 1979), stresses the significance of environment to individual perceptions.

environments are actually multimodal sets of messages with to a high extent common meaning.[1]

The objective of the present work is to introduce a new kind of approach to the design of multimodal user-interfaces. Central to this approach are the relationships among presentation elements; *redundancy* is presented as a unifying concept. Expressed in another form, the tasks of this work are

1. to find a functional definition of redundancy for the needs of design of multimodal user-interfaces, and
2. to find an applicable theoretical basis for the analysis of the cognitive consequences of presenting several elements simultaneously.

The first task means an attempt to create a conceptual tool for the analysis of combinations of presentation elements. In other words, the objective is to support the creation of multimodal messages. The second task is performed in order to develop the definition of redundancy in accordance with human cognitive ability.

1.2 Disciplines applied

1.2.1 Cognitive sciences and related approaches

Awareness of the need for understanding the human being in human-computer interaction is a result of the growth in the variety of users and the complexity of applications. In order to make computers useful for as wide a variety of users as possible and to broaden the domain of applications, the engineers needed to know what kind of machine a human being is able to use. The problem was and is how to support the desired behavior in human-computer interaction. Experts on cognitive psychology have been the most popular consultants in this problem. The designers of a user-interface ask psychologists how information should be presented to a user so that it will be understood.

In the described approach the human being is basically seen as a compulsory constraint in the development of technology. Resources are invested in user-interface design because of external pressure – a high quality user-interface is instrumental to gain primary targets like better user-satisfaction, which – in turn – has commercial value. The features of human mental processes are taken as a static, given framework that has to be accepted. Technology has to be developed in terms of human mental qualities.

When exploring mental processes, it is the objectives of the activity that define the perspective and the quality of possible results. To some extent, it can be argued that different points of view complement each other and an interdisciplinary approach gives a broader view of the entity. Frequently, this

---

[1] For simplicity, the phenomenon is later referred to with the expression "multimodal message," although it can be interpreted as a set of separate messages.

kind of view is an over-simplification of reality; in practice, different approaches to the research inevitably lead to examination of quite different entities. The unifying concept, say, mental processes, is then in too abstract a level to support the construction of an over-all view.

Palmer and Kimchi (1986) outlined different kinds of approaches to human mental events. They defined three categories, which can be seen as a continuum:

- *Physical approach.* When the focus of interest is on observable physical phenomena, we are closest to the natural sciences. In a physical approach, reductionism is at an extreme level, but this is hardly seen as a problem as the primary interest is in physical details. However, if the primary object of interest is changes in voltage of a single neuron or the composition of a compound found in a certain part of the brain, it is quite an abstraction to argue that the subject is mental processes. Rather, the nervous system is an object of interest in itself.
- *Functional approach.* In a functional approach, the physical mechanism that performs a process is not the essence. The process itself is what matters. When adopting a functional approach, the researcher is interested in how the mind works rather than which particular physical objects cause the phenomenon.
- *Phenomenological approach.* At the other end of this classification, the interest is in conscious experiences. The phenomenological approach differs from physical and functional approaches in both ontological and epistemological aspects. Appropriate descriptions of mental processes are argued to be impossible to make without subjective observations and interpretations. The object of interest, conscious experience, is at a quite different level than physical entities.

Hardly any theory represents purely one of these approaches. That is why this classification should be seen as a continuum. But when defining the classification as a continuum, it is reasonable to ask what are the properties that define the level of a certain theory in this dimension. The most obvious property is the degree to which a theory is based on measurable facts. At one extreme, the conclusions are logically inferred from numerical data, while at the other, conclusions are subjective interpretations of subjective observations. The other property is the belief in the structure and nature of reality. The extreme physical view is finally constrained by known physical laws, whereas phenomenological discourse is not restricted to concepts referring to physical entities. In other words, the classification is based on differences in epistemology and ontology.

Because physical and phenomenological approaches are the theoretical ends of the continuum, practically all real cases are located somewhere between these extremes. Before deciding on an approach, a researcher has to consider carefully the appropriateness of the expected sort of data, as well as other consequences of the choice. The basis for the choice is, inevitably, the belief in the nature and structure of reality.

Materialism often leads by logical reasoning to a physical approach. If physical laws are seen to explain atomic mental events that constitute all more complex behavior, it seems reasonable to spend resources on investigation of these elementary entities. On the other hand, much qualitative research is done on a materialistic base. In the latter case, physical laws are acknowledged as the basic framework of reality, but when the object of research is extremely complex, like mental events, a more abstract approach is chosen to get at least some kind of view about the object. A logical inference is that, according to this view, all approaches other than the physical one are only temporary concessions while waiting for physics to finally explain the human mind at an atomic level.

To give a similar description about non-materialistic view is certainly impossible because of the wide variety of possible approaches that can be described as non-materialistic. In such an attempt, we are bound to discuss the nature of science and the criteria for scientific validity of different approaches.

One important divide between materialistic and non-materialistic views is the understanding of the nature of conscious experience. For example, John Taylor, professor of mathematics and director of the Centre for Neural Networks at King's College (London), declared (1995) that even conscious experiences will be explained in physical terms. He believed that in ongoing projects some kind of conscious experiences can be created. Philosopher David J. Chalmers, whose ideas have been central in a recent interdisciplinary forum called consciousness studies, rejects materialism with philosophical arguments (1996, pp. 123-124). Chalmers has found conscious experience such a unique and important phenomenon that it deserves to be handled as fundamental (1995). This fundamental is, according to him, "over and above the physical facts" (1996, p. 124). The ontology of the current physical view simply does not cover the phenomenon.

Chalmers (1995) found an analogy between the current problem of explaining consciousness and one that took place when a phenomenon called electromagnetism was discovered. Electromagnetism could not be explained in terms of known physical laws. Therefore, the view of reality had to be broadened by expanding the ontology of physics. New concepts and laws were needed in order to analyze the phenomenon. Similarly, Chalmers (1996) argued that conscious experience is a phenomenon that can neither be denied nor explained in terms of current science. That is why he found it necessary to broaden the view by creating a unified theory of physics and psychophysics.[1] (pp. 276-277)

The debate about consciousness indicates the need to find a theoretical basis that would be valid in more aspects than the ones that handle publicly observable entities. It is not a question of rejecting the physical approach but of broadening the view.

Broadening the view of the physical approach would broaden the epistemological as well as the ontological basis of reasoning. Basically, it is a question of balance between two principles of science that are frequently difficult to combine; science should seek for truth and objectivity. In extreme objectivity the problems are so reduced that they have little to do with reality. The criterion for formulating the problems is then whether they can be reliably

[1] The concept of psychophysics is defined by Chalmers himself.

examined with quantitative methods. At the other extreme, problems are real, but the access to them is beyond logical reasoning.

Finally, it cannot be said that this controversy between the strictly physics-based view and a broader view is a controversy between objectivity and subjectivity. As knowledge is nowadays defined as a subjective phenomenon, the role of subjectivity in the construction of scientific knowledge does not need to be neglected. Instead, once identified and acknowledged, the role of subjectivity can be analyzed. The chain from laboratory observations to knowledge inevitably contains subjective elements, as do clearly qualitative studies. Above all, a researcher has to *believe* that the results of an experimental setting are applicable in real life.

As a summary it can be said that in order to find an approach that can produce information for the needs of knowledge construction, a balance must be found between physical and phenomenological approaches. It should not be too subjective in order to provide information that can be accessed publicly and can be used in science. On the other hand, it should not rely blindly on the possibilities of physics to finally explain the universe with the help of quantitative methods. If this balance is found between the two extremes, physical and phenomenological, it is very likely to be characterized as functional.

Of course, there is nothing new in the contradiction between physical and phenomenological approaches. Especially in humanities, there has been a vivid debate for decades about different approaches. The discrepancy between the approaches has usually concerned methodological questions, thus directly referring to differences in epistemology. But argumentation about ontological questions would probably lead the discussion more directly to see possible differences in the content of analysis instead of the form. In other words, if a severe contradiction between qualitative and quantitative methodologies emerges, it is probably a question of differences in ontology and thus, in comprehension about what is the object of research. If there is mutual understanding about the nature of the object of research, methodological choices have a firm basis.

1.2.2 The information processing approach as a functional approach

While physical approach is implicitly tied to materialism, the phenomenological approach leaves the question about the related ontology unanswered – instead it deals with epistemological questions. The functional approach is even more unclear in this sense. Functional descriptions can be made on any ontological and on any epistemological basis.

The information processing approach (from now on referred to with the established abbreviation *IP-approach*) is a typical approach used to study the human mind in a functional manner. Although an approach may be functional, that does not necessarily restrict its possible underlying epistemologies. However, the bulk of IP-oriented research seems to try to rely on data from carefully designed quantitative experiments. Palmer and Kimchi (1986) stress its close relationship to the physical approach. Massaro and Cowan (1993) try to get rid of the fuzzy nature of the IP-approach by applying unambiguous strategies, falsification (see Popper, 1959, p. 33) and strong inference (see Platt, 1964). Van der Heijden and Stebbins (1990) seem to wish to do about the same thing by distinguishing the approach from cognitive psychology to manifest the emphasis on overt behavior instead of cognition.

Palmer and Kimchi (1986) presented a meta-theoretical view of the IP-approach, in order to outline the domain. Even if their arguments are not universally accepted (van der Heijden & Stebbins, 1990), they provide a lucid basis for understanding what the IP-approach actually is about. Palmer and Kimchi stated five assumptions, which are, according to them, acknowledged by almost all IP-psychologists:

1. Information description. All mental events can be described as three-stage information processing events – input, operation, and output.
2. Recursive decomposition. All mental events can be recursively decomposed to sub-events until the level of primitive events is reached.
3. Flow continuity. All operations must be able to be performed on the basis of input information, i.e. outputs of preceding events.
4. Flow dynamics. To perform an operation, all input information and sufficient time is needed.
5. Physical embodiment. In informational events information is embodied in states of the system and operations in changes in the states. The states of the system are called representations, and changes in the states are called processes.

Even if these five qualities are only basic assumptions, not necessarily objects of research, they reveal the nature of the approach. The IP-approach is based on the interest in mental processes as a sum of identifiable sub-processes. Thus, explanations are usually reductive. The epistemology of the approach is adopted from the natural sciences – most data of experiments are quantitative in nature, and therefore results are based on statistical analysis.

The five assumptions (Palmer & Kimchi, 1986) did not contain any assumptions about generalization, but the methods used reveal one weakness of the approach: in practice, IP-psychologists seem to be interested in human beings as a species. In many attempts to outline the architecture of human mental processes, individual differences are forgotten. Besides differences between individuals, differences between people from different cultures show how complicated the process in which information processing procedures are developed is (e.g., Griffin & Gibbs, 1992).

The temptation to find an analogy between mental processes and a computer is often too high to be resisted. Adopted vocabulary (memory[1], storage, processing, etc.) reflects the orientation (Neisser, 1967, p. 6; van der Heijden & Stebbins, 1990). It is even argued that the computer as a metaphor of the human mind has a central role in cognitive science (Gardner, 1985, pp. 6,

1 The conceptions of memory seem to be especially firmly connected to computer metaphor (see the review of Craik and Lockhart, 1972).

40), not to speak of the long tradition of attempts to simulate cognitive processes with a computer (e.g., Collins & Loftus, 1975). But by handling human mental processes as mechanical, programmed functions, only a narrow view of mental processes can be achieved. According to Gardner (1985, p. 6), issues concerning e.g. affections, as well as historical, cultural and contextual aspects are consciously given a minor role in order to simplify the research domain. With such a notion, Gardner did not give too ambitious an image about the work done in cognitive science. On the basis of Gardner's statement, it can be understood that in cognitive science only problems that are easy to examine are handled. Of course, it can not be argued either that there is no reality beyond the physical facts. However, Gardner evades the challenge articulated by Chalmers (1995). Chalmers divided the problems into easy ones and to hard ones. Easy problems may be solved with traditional methods, whereas hard problems may not (he refers above all to problems concerning conscious experience). Chalmers lists typical strategies to avoid handling hard problems. Gardner's strategy, in Chalmer's concepts, seems to be to "explain something else." Since the most interesting problems seem too complicated, simpler ones are handled instead.

The cost of limiting the domain in cognitive sciences to "easy problems" could be discussed endlessly. The approach of cognitive science, and thus also of cognitive psychology, can be seen in the methodological choices. The empirical research in the domain is based on carefully designed experiments and quantitative analysis. In order to broaden the view, a wider variety of methodological approaches should be applied. However, disparaging the work done by IP-researchers would be a grave mistake. Whether we agree with them about everything in their approach or not, many of their findings cannot be denied.

The IP-approach is used in the current work as a framework or paradigm to the extent it is relevant. The IP-approach, although its appropriateness will be questioned, provides some useful concepts and guidelines in order to analyze events that take place in intentional information presentation. In the current work, IP-theories are utilized with respect, though critically, thus leaving the door open to a broader view.

1.2.3 Communication studies

Cognitive sciences promote the current research interests by providing information about mental consequences of information presentation to the user, whereas the contribution of the point of view of communication is more difficult to characterize. The difficulty is due to the essential differences between schools of communication. Fiske (1990) divides them into the process school and the semiotics school. To the process school he attributes all doctrines that are derived from information theory (Shannon & Weaver, 1949). In the semiotics school, the interpretation of signs is the objective of interest.

1.2.3.1 Process models

The metaphor of process models of communication is to copy or transfer a certain information unit from a sender to a recipient. The illustrations of the models usually contain at least a sender, a recipient (or, when focusing on the content, instead of a sender and a recipient, a sent and received meaning or message), and an arrow from the former to the latter. Different variations contain various intermediate stages on the way from a sender to a recipient.

Process models can be criticized for their extremely mechanical view of communication. In them, sending a message is like programming a computer; the receiver is influenced by delivering him a message that is encoded with a coding system that he is supposed to be able to use in decoding. Communication skills consist of skills of encoding and decoding and of using a communication channel. Knowledge about the receiver is important in order to encode in an appropriate manner, i.e., in order to express a meaning in a way that it is likely to be interpreted in a desired way. This kind of mechanical view of communication is quite incompatible with contemporary paradigms of the human being as an active and intentional creature. However, its common basis with IP-models is salient. The computational approach (Gardner, 1985, pp. 6, 40), or information theory (Shannon & Weaver, 1949), provide a paradigm that unifies the whole communication process with its all mental and technical stages. Within this paradigm, communication can be described as an information flow from one physical place to another. No sharp conceptual distinction is made between the operations performed by a human and the ones performed with the help of technology.

One advantage of process models is that they describe the roles of the persons involved in the communication process. This kind of conceptualization supports the analysis of complex mediated communication processes like application development. Vacherand-Revel and Bessière (1992) illustrated the complicated relationship between designers, authors and learners involved in CAL (computer assisted learning, figure 1). Their chart shows how information flow from a designer to a user, despite the interactivity between a user and a computer, is one-way in nature. It also shows how the user-interface is a tool for a designer to send messages to a user. Therefore, in computerized environments it is possible to speak not only about human-computer interaction but also about one-way communication between the designer and the user.

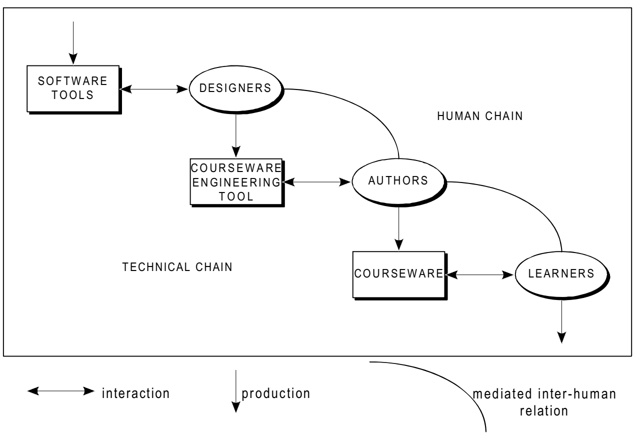

FIGURE 1 The actors, tools, and learners involved in CAL. (Vacherand-Revel & Bessière, 1992)

1.2.3.2 Models based on semiotics

The communication models that are based on semiotics do not handle communication as a process of sending and receiving messages. Instead, they analyze the interpretation of signs. In the process models, the receiving of a message is a more or less passive process in nature, and only the sending of a message requires activity. In semiotics, the receiving of a message is an active interpretation process of signs. Sending a message, in process models, is to deliver information in a form that is designed to match the mental structures of a receiver. Mental structures are an object of interest in order to be influenced as effectively as possible. Instead, in the semiotics school, a message is not a unit of information that is *pushed* into the mental structure but a sign that is first actively picked up and then interpreted by the receiver.

The comparing of process and semiotics models is obviously problematic because of their completely different approaches. Even the use of concepts like "sender," "receiver" and "message" are slightly inappropriate concerning the semiotics models. However, in the current work, the different approaches are used in parallel. The conceptual basis is in process models because the models are mainly applied in the process of "sending a message" from a designer to a user. But the active role of the user ("receiver") is emphasized as in the semiotics models.

1.2.4 Educational sciences

Education has already been referred to as a domain that has high requirements concerning user-interfaces (p. 15). Educational sciences also provide methodological support for the current study (p. 27). However, the most significant point concerning educational sciences here is that it is an approach that integrates the other applied disciplines, cognitive psychology and communication studies, for the needs of the current work.

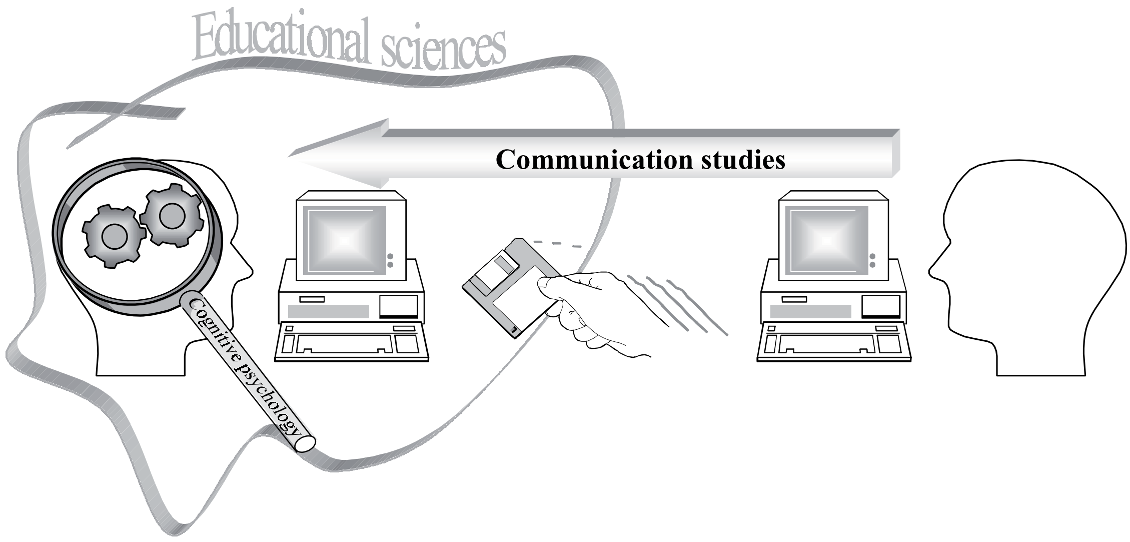

FIGURE 2 The roles of the different disciplines in the current work.

Figure 2 illustrates the roles of the applied disciplines. Cognitive psychology is used to provide descriptions of users' cognitive functions and structures that have to be taken into consideration in creating the form of presented *information*. Communication studies, in turn, provide a framework for the analysis of the dynamics of creating and interpreting messages. In them, especially in the semiotic school, the focus of interest is on *meaning* instead of information. Information has only instrumental value in the creation of meanings.

The role of educational sciences in the current work is not as simple to describe. Using a simple metaphor, it can be said that cognitive psychology and communication studies are like the two eyes of a human being. Each provides a slightly different perspective on the object of the gaze. Educational science is used here as the brain that combines those two different images into a single three-dimensional impression.

The perspective of the current work is, in the terms of educational science, "*user as an active learner*." The phrase can be interpreted from the perspective of cognitive psychology as "*user as an effective information processor*" and from the perspective of communication studies as "*user as an active interpreter of messages*." The educational sciences provide a paradigm that covers not only concepts but also values that basically give meaning to the activities of studying user-interfaces. The perspective of education is truly *user-centered*, as an approach in which the properties of the user are prioritized is often called in system development literature. In education, the user is not a set of properties that is supposed to work as a part of a technological system. In education, all aspects are submitted to the intentions of a human being. So this work does utilize cognitive psychology and communication studies in order to achieve a better comprehension about the human being as an active, intentional, and creative creature.

1.3 Structure and methodology of this work

The first stage in performing the research tasks (p. 16) is to form a preliminary definition of redundancy (chapter 2). That definition serves as a basis for further analysis. Second, the preliminary definition is exposed to the theories of attention (chapter 3) and two other IP-oriented theories (chapter 4). Chapters 5 and 6 are about the third stage. That is an experiment in which the applicability of the preliminary definition of redundancy is tested in a user-interface analysis task with multimedia design students.

The order of chapters does not illustrate the true progress of the work. The basic ideas and conclusions have matured gradually and different stages have chronologically overlapped each other. The conceptions concerning redundancy and the conceptions concerning the contribution of the IP-orientation to design have been developed in a reciprocal relationship with each other. Therefore, even the preliminary definition of redundancy is influenced by the analysis of the experiment, for example. For the same reason, discussion about the concept of redundancy is distributed among several chapters. Different aspects are discussed in different chapters according to each particular chapter’s contribution to the definition of redundancy.

It is important to note that even if there are early chapters that contain theoretical reasoning and then a description of an experiment, this work does not follow the standard structure of an empirical study. In a traditional setting, the introduction of theoretical background aims at the formation of hypotheses, which are then empirically tested. In this work, both theoretical reasoning and empirical testing contribute to the development of the model.

In the present work, redundancy is approached from two main distinct points of view: concerning attention (and other IP-oriented theories) and concerning the conceptions of redundancy (the preliminary definition and its empirical testing). However, the statement that these two approaches provide different points of view does not mean that they are only different perspectives on the same phenomenon. They also differ with respect to the quality of information they are able to convey; not only are the perspectives different but also the way the phenomenon is looked at. The difference in the quality of information between the two approaches is a result of the difference in the nature of the data available. Therefore, the different points of view are described here in separate chapters.

1.3.1 The analysis of IP-oriented theories

The theories and models of attention have been developed for several decades. The changing paradigms in research and the needs of society have influenced both the quality and the quantity of interest paid to the topic. But on some level, the same few themes concerning the human way of prioritizing the constant flow of information from the environment through the senses have been more or less present in research for a century (Lovie, 1983). However, since the prevailing paradigm defines the perspective, concepts, and acceptable methods (Kuhn, 1970), the interpretation of old research reports is not a trivial process. At least, if the results of research from former paradigms is intended to be taken as verified facts, the reports should be seen as representatives of their paradigm and interpreted from the point of view of the current paradigm. For example, in 1972 Ingling reported on the basis of her experiment that it is easier to detect certain characters (targets) among distracters if the targets and the distracters were of different categories, e.g., letters and numbers. The setting of the experiment was reported in detail, probably in order to validate the test. Now the detailed information about the experiment serves as a way to evaluate the appropriateness of the experiment from a contemporary point of view. The description of the experiment makes it possible to see the actual test situation as a whole. Thus the physical context can be taken into consideration and the results can be evaluated according to the current view that emphasizes the context.

The analysis of the theories concerning attention thus implies two-dimensional interpretation. First, as described above, the theories should be seen as products of their era and paradigm. In practice, this means that the theories should be interpreted with the concepts and from the point of view of the contemporary paradigm. Second, in the present work, the theories are interpreted as a means to analyze user-interfaces. Even if these two dimensions of interpretation are not explicitly separated in the analysis of the theories, they should be seen as underlying principles according to which the theories are handled here.

The two other IP-oriented theories, dual coding theory (chapter 4.2) and the ICS-model (chapter 4.3) are easier to analyze than the theories of attention, since both of them have updated versions from the 90's. The ICS-model is even created for the needs of analyzing human-computer interaction.

1.3.2 The method of the experiment

The original idea of arranging the experiment was to get a preliminary comprehension about the relevance of the results of the theoretical analysis in practice. I.e., the plan was to define the concept of redundancy on the basis of theoretical reasoning, then to teach the subjects the definition and to test whether they could detect the related phenomenon when analyzing user-interfaces. Finally, however, when the data was collected it proved to deserve a different kind of analysis and a more central role than what was intended. Although in the original idea the researcher first explicitly defined the concept and then tried to transfer the definition to the subjects, in the actual experiment the subjects simply did not resign themselves to be passive receivers of information. They only used the pieces of provided information as elements when they actively constructed their own mental representations about the phenomenon. The result of that activity was too interesting to be neglected even though it was not possible to classify the data according to the original idea.

The purpose of the experiment thus changed from testing a theory to creation of a theory. The final setting resembled the one that is described by Glaser and Strauss (1968). This method, grounded theory, is based on an idea according to which data and theory are in a reciprocal relationship with each other; the theory is developed gradually in parallel with data collection and data analysis (Strauss and Corbin, 1990, p. 23). However, grounded theory is a method with a strict sequence of procedures (Tesch, 1992, p. 24). Since the sequence is not followed in the present work, the method used cannot be labeled as grounded theory despite similarities in principles. Therefore, when trying to name the method used, the label must be searched for among methods that do not provide as strict limits for the procedures as grounded theory.

The data provided information about different types of conceptions about redundancy and about changes in those conceptions. Marton (1978) used the same approach when he developed a method for the needs of learning research. The method, which was later named phenomenography, is a way to analyze and classify conceptions about a phenomenon. When applying Marton's method, the object of interest is our experiences of reality rather than reality itself. Marton designed his method to be "directed towards experiential description" (1978, p. 6). His aim is not to construct abstract cognitive models. The phenomenographical approach is contextual. The task is to analyze and categorize different conceptions of some real-life phenomenon. Marton argues that the differences in conceptions among individuals often resemble the sequence of different conceptions in the history of science. Marton also recognizes a similar evolution in the change of an individual's conceptions (Marton, 1988).

Categorization of conceptions is an essential part of phenomenography. According to Marton, there is a relatively limited number of qualitatively different ways to conceptualize certain phenomenon. By analyzing the categories, the essential structural differences in conceptualization are searched for. The resulting structural framework provides a way to see the relationships between different kinds of conceptualizations and thus a way to understand different kinds of conceptions.

Marton's background is in educational sciences, and phenomenography is usually related to education. In the empirical part of the current work, there are features that make it suitable to be handled within the prevailing learning paradigm. Learning in the sense it is used in contemporary cognitive theories (Resnick, 1989, p. 2) is such a broad concept that it can be applied – at least to some extent – to almost all analysis concerning human mental processes. Learning is nowadays understood as an active interpretation of information on the basis of former knowledge, just the thing the subjects of the current experiment were supposed to do. Additionally, their former knowledge was based primarily on the given instruction; even if the phenomenon (redundant messages) is an essential part of our everyday life, its conceptual analysis seemed unfamiliar for the subjects. Even the word used to refer to the central concept, redundancy[1], seemed new for all the subjects. Therefore, the setting can be easily seen as a learning process that is based on an intentional teaching act.

Phenomenography is an appropriate tool for performing the current research tasks. Above all, the data proved suitable for analysis by Marton's method.

1 In the experiment, the Finnish word "redundanssi" was used when referring to its English counterpart ("redundancy"). Usually, redundancy is translated to Finnish with other, original Finnish words, but their meanings are limited to certain contexts and they are not appropriate when analyzing the concept referred to by the English word "redundancy." When translating in that way, the change in the content of the concept is quite analogous to that of substituting in English "repetition" for "redundancy."

2 REDUNDANCY

2.1 Introduction

In verbal communication, word choice is one of the most important actions. The expectations of the consequences of a choice reflect our view of the communication act. If communication is seen as a mechanical coding and transmission of meanings – as in models that are based on information theory – our attitude toward our ability to transmit meanings is probably quite optimistic. But if we take the receiver of a message as an active interpreter of signs – as in the semiotics school – our assumptions concerning the evoked meanings are bound to be more careful. Mental representations that a word activates or creates are to a large extent subjective, dependent on complicated associations and the results of creative processing.

In the present work, *redundancy* (as well as the related adjective *redundant*) is used as a central concept in a conceptual model for analyzing multimodal messages. Because it is not a novel word, its appropriateness concerning the current use has to be evaluated in the light of its former use. *When we communicate with words, we choose words that we think are most likely to be interpreted in the way we wish them to be interpreted*. That is why we have to take into account all possible levels of meaning, especially if the concept in question has no clear physical referent. When assessing the appropriateness of the concept of redundancy, we have to compare the current referent with other possible referents or interpretants (Peirce, 1958, pp. 296-297) or significations (Saussure, 1983, p. 67). Because we do not have direct access to these abstract, subjective objects, the concept of redundancy is now interpreted on the basis of definitions found in the literature. The references cited her cover different disciplines and approaches, e.g. computer science, psychology, and everyday language use.

2.2 Definitions of redundancy

When exploring the use of the word redundancy, one of the most common views arises from information theory. Shannon and Weaver (1949, p. 13) define it as an unnecessary part of a message, or a portion that could be left out without loss of information. This view dominates in technological environments, in which the effectiveness of data transmission and storage is the prime concern. Expressions like *unnecessary*, which are used for example by Shannon and Weaver (1949, p. 13), Longman Dictionary, (1984, p. 1243), and the psychologist Staniland (1966, p. 15), who deals with psychological problems derived from communication engineering, and *nonessential* (Longman Dictionary, 1984, p. 1243) are usually used to describe the concept in this sense. The same tone can be seen in expressions like *superfluous*, *tautological*, and perhaps *verbose* (Longman Dictionary, 1984, p. 1243). This view is most obvious in computing, thus referring to, for example, optimization. But it is also used in linguistic studies where the possibility of abbreviating the linguistic code has been investigated (Weik, 1977, p. 293; Edwards, 1992; Paul, 1992). For example, it is argued that written English is about 50% redundant (Shannon & Weaver, 1949, p. 13; Web Dictionary, 1997), which means that the number of letters could be reduced by half without losing the possibility of distinguishing words from each other.

A similar view – redundancy as the concept that a message would have been complete with less data – is also expressed in somewhat more neutral terms: "…can be eliminated without loss of essential information" (Maynard, 1975, p. 158); "The amount by which the decision content of a message exceeds the entropy… Redundant – Equipment, facilities, or data over and above the minimum required for a specific purpose or effect" (Weik, 1977, p. 293).

These two ways of describing the concept have a very important difference. In the first one, redundancy is clearly a phenomenon that is undesirable, something we have to get rid of. In the second one, neutrality is achieved by more extensive explanation, thus avoiding the transference of emotional tones related to single, strong expressions (e.g. *unnecessary*) to the concept of redundancy. However, strong, negative expressions seem to dominate the use, while there are no clearly positive ones, but only neutral definitions of the concept to counterbalance the tone.

Another, even more clearly negative use of redundancy deals with superfluous workers in industry (Staniland, 1966, p. 15; Longman Dictionary, 1984, p. 1243). Although this use of the word is common in everyday language and has no direct relation to uses in information technology (despite the fact that information technology is one of the main reasons why workers become redundant), the extremely negative tone is surely a burden for the word if it is intended to be used in order to create more positive connotations.

In domains like computer science and artificial intelligence, the concept of redundancy is also used in a quite different way: to signify a method for increasing the security of a system. Expressions like "Having backup components or overlapping information that can be used to complete a task in the event of a minor failure…important architectural principle…" (Mercadal, 1990, p. 241), "…to assure reliability and to maintain a communication channel" (Web dictionary, 1997), "serve to facilitate a check…" (Weik, 1977, p. 293) are bound to give an impression of something essential and important. This – undoubtedly positive view – could be found in dictionaries of computing, although not in the Longman dictionary, which, however, explained the concept quite extensively.

Thus, it can be argued that in computing, redundancy has quite separate meanings, both positive and negative in nature. In everyday language, negative tones are likely to dominate.

2.3 Communication studies and redundancy

Information theory is often referred to as the mathematical theory of communication. But in communication studies, the word redundancy has a completely different tone from that of Shannon and Weaver's theory. In these studies, the concept is used to signify an important principle of integrating two or more messages, possibly of different modalities.[1] Most of them concern television (e.g., Reese, 1983; Drew & Grimes, 1985; Hanson, 1993; Basil, 1994), but in the same manner, the concept is used in some studies concerning interactive environments (e.g., Edwards, 1992; Miller, 1993) or just signal detection (Green & Anderson, 1956). Additionally, there are studies that handle the same phenomenon but do not use, for some reason, the word redundancy. For example, Braden (1992) uses the phrase "visual-verbal symbiosis" (launched by himself in 1983) when referring to the redundancy of visual and verbal elements in audio-visual material. Another expression that refers to redundancy is used by Drew and Grimes (1985): close coordination between audio and video. Bahrick and Gharrity (1976) described the relationship between redundant pictorial elements as "independent," referring to a similar effect of both stimuli.

The way the word redundancy is used in communication studies is brilliantly illustrated by Edwards (1992). As seen in figure 3, Edwards's view of communication is clearly derived from information theory: there is a sender, a communication channel, and a recipient (here, both sender and receiver can be machines as well as humans). Communication is an act in which the sender wishes to change the state of the recipient by sending a message through the communication channel (first row of the picture). In the second row, messages are sent and the desired change in the state of the recipient takes place. The

1 A closely related approach is the use of the concept of redundancy when speaking about text illustration and learning. In the research review of Levie and Lentz (1982) and in the studies it referred to (such as Levin & Lesgold, 1978, and Haring & Fry, 1979) the word redundancy was used to distinguish conditions in which illustration was relevant to the text to those in which it was not.

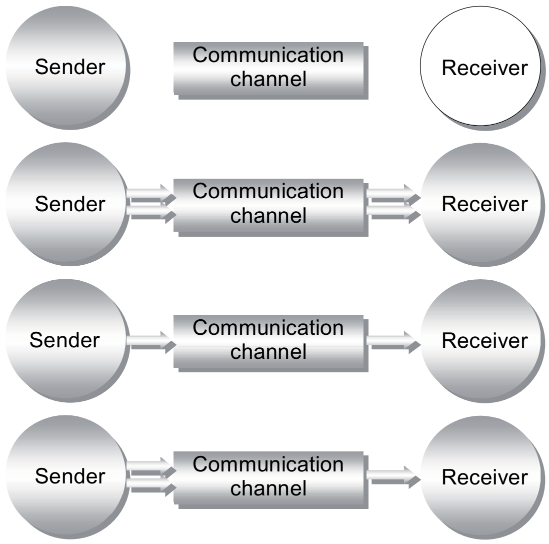

FIGURE 3 The amount of sent and received information in communication act. (Edwards, 1992)

third row illustrates a situation where the amount of information is reduced to as little as possible but still has the same effect in the recipient. Because it was possible to reduce the amount of information, some of it must have been redundant. In the last case, the amount of information is reduced in the communication channel. But still enough information passes through the channel in order to obtain the desired effect. Thus, in the last case, redundancy compensates for the deficiency of the communication channel.

Edwards's analysis is aimed at illustrating the possibilities of creating user-interfaces for disabled people. For him, deficiencies in the communication channel imply physical deficiencies of humans. But Edwards's illustration also sums up different views of redundancy and provides a metaphorical basis for discussing redundancy. First, as mentioned, the model clearly refers to the mathematical theory of communication. Second, it illustrates how redundant information can be reduced, still causing the same effect in the recipient. Third, it illustrates how redundancy, in certain cases, reduces the demands put on the communication channel.

Hsia (1971, 1977) also deals with redundancy as a feature of human communication. Hsia's view is derived from information theory, too. He legitimizes the application of information theory in human mental processes by arguing that "…it [the nervous system] is no more than a collection of cells with wired connections functioning analogously to a modern telephone system…" (1971). As a result of his approach, the need for redundancy is described in the same manner as it is described in computing: to ensure the transmission and storage of information. Hsia speaks about over- and under-redundancy and optimal redundancy between these two. It is argued that too much redundancy causes uncertainty, while too little of it makes the task difficult. Thus, Hsia's view of redundancy is quite neutral: redundancy itself is neither good or bad, it is only a tool of communication and should be used in an appropriate way.

However, at least one view is still missing. Providing essential information for both blind and deaf people would hardly on its own make redundancy a desired phenomenon on television. The studies in the domain concern perfectly hearing and seeing people. Why should they be provided the same information in both audible and visual formats? Neither the reports of Reese (1983), nor Drew and Grimes (1985) contain analysis about the phenomenon itself. They only use the concept as given, and report the statistical connection between redundancy and recall or understanding.

2.4 A redundancy model

In order to be able to locate redundancy as a concept in wider mental structures, the approach of Basil (1994) is worthy of consideration. He quoted Ekman and Friesen (1969), who listed different possible relationships between the verbal and nonverbal behavior of an individual. The authors specified the following relationships: "The non-verbal act can repeat, illustrate, accent, or contradict the words; it can anticipate, coincide with, substitute for or follow the verbal behavior; and it can be unrelated to the verbal behavior" (p. 53). Basil then applied this list when formulating a list of possible relationships between messages of different modalities: redundancy, substitution, complementing, contradicting, and emphasizing.

The list of Ekman and Friesen, and the interpretation of Basil, are a good basis for forming a list that would form a continuum or dimension, in which redundancy is only one stage. By defining such a dimension, we could get a tool to illustrate different kinds of cases that cannot be described with one single word.

When defining a dimension, possibly definitions of the ends of the dimension would be enough. At least it is easiest to begin with them. When we have two messages and consider possible relationships between their contents, finding the extreme cases causes no problems: the messages can be identical or totally unrelated. From now on, these extremes are called *distinct* and *redundant* cases.

The relationships could be illustrated simply with circles that are thought to border a set of facts from a space of all facts. Figure 4 is a reduced, abstract illustration of the proposed dimension. The first and the last row represent the extreme ends of the dimension, and the second and the third rows different intermediate cases.

The essence of the drawing is to illustrate the change in the quantity of information in common between two messages. While in *distinct* conditions the messages have nothing to do with each other, in *redundant* conditions they concern exactly the same entities. In practice, most cases can be located somewhere between these extremes. That is why the intermediate area has to be elaborated.

In the second row, the circles are tangents to each other and thus, have only one point in common. But they still form a figure together that is no longer two circles but a single whole. This figure represents a case in which the messages are about the same object, but contain quite different information about it. For example, the first message could be “Peter has long, dark hair” and the other one “Peter uses an earring.” The aspect in common between these messages is that they both handle Peter. But they tell quite different things about the boy. These messages *complement* each other, thus forming an overall

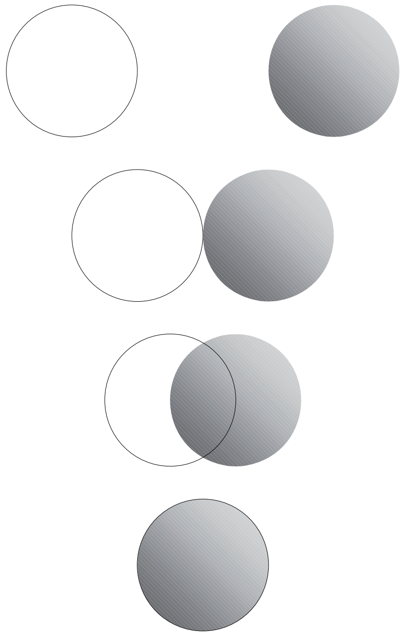

FIGURE 4 Possible relationships between two elements. (Pirhonen, 1998)

view about a youngster. The relation can be called *complementing*, as Basil (1994) did. In Ekman and Friesen's vocabulary, the other message *augments* the first one. But if the first message was "Peter has long, dark hair", and the other "Peter is a president of a multinational company", the other message is likely to change our mental representation of Peter, formed on the basis of the first message. It can be argued that the other message complements the first one, but above all, it substitutes the meaning of the first message with a new one: instead of a youngster, we form a mental image of a powerful, bohemian leader. The relationship is here called *substituting* (Basil, 1994). But in figure 4, the same case illustrates both complementing and substituting relationships. The feature that separates these two types of relationships is that in complementing conditions, the first message anticipates (word used by Ekman & Friesen, 1969) the second one. In other words, when considering whether a relationship is complementing or substituting, we have to consider how the relationship fits our conventions.

In the chart (figure 5), the horizontal axis illustrates the amount of information in common in two messages. The vertical axis refers to how the combination of two elements fits our conventions. The area surrounded by dotted lines is assumed to be a closed area, within which practically all relation definitions can be located. When moving from the origin rightwards, the possibilities for variance in the vertical axis decrease. This is due to the logical fact that when the relative quantity of information in common increases (i.e., the relation becomes more and more redundant), the relative quantity of information that is contained in one element only decreases. When moving

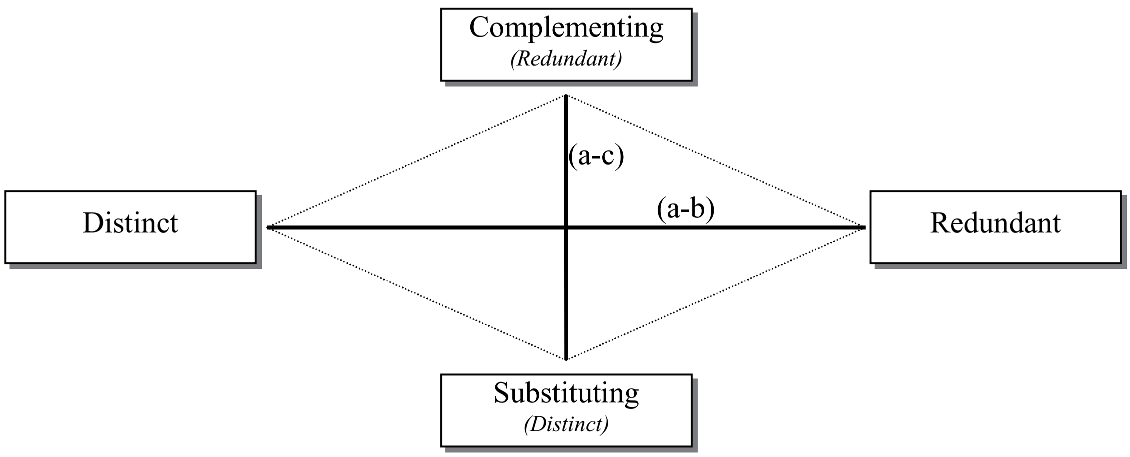

FIGURE 5 Different relationship types. (Pirhonen, 1998)

leftwards, the possibilities for variance decrease again. But the reason is that when the messages have less and less to do with each other, we are likely to have less and less expectations about the other message on the basis of the first one. Thus, the criterion for defining the level of a relation on the vertical axis gradually loses its sense.

When discussing the meaning of the vertical dimension, another interpretation appears. The criterion of the level on the vertical axis was said to be defined on the basis of how the combination of elements fits our conventions. Thus, when considering the relationship of two elements we inevitably have a third element, too: the existing mental representation concerning the item. In figure 6, items "a" and "b" refer to a single output element, "c" to a mental representation about the object.[1] This kind of abstraction makes it possible to simplify the model even more: the ends of the vertical axis can be interpreted as *redundant* and *distinct* with mental representation, thus substituting the concepts *complementing* and *substituting*, respectively (italics in figure 5). However, even if a model might seem more compact with analogous dimensions, the original concepts (*complementing* and *substituting*) would probably be more descriptive.

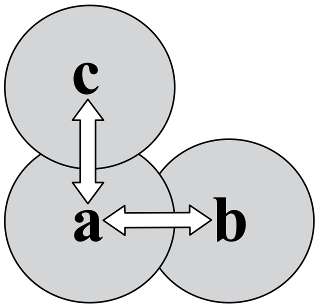

FIGURE 6
Relationships that define the dimensions of the model.

To be able to cover all the expressions of Ekman, Friesen (1969) and Basil (1994), the chart needs at least two more explanations. First, there is a usual case of relationship that Ekman and Friesen (1969) refer to when they use the verb *accent* and Basil with the word *emphasizing*. If illustrated analogous to figure 4, the relationship would look like the one represented in figure 7. The other message (small circle) emphasizes or accents the first one. But how could this kind of case be interpreted in order to be located somewhere in the dimensions defined earlier? From the point of view of the larger circle, the smaller is completely redundant, containing only the information that could be found in the larger circle as well. But from the point of view of the small circle, the larger contains much augmentation that could be further analyzed on the basis of its conventionality. Because the decision of the definition of this relationship is related to point of view, it is here mentioned separately. That way coherence with earlier studies is also easier to maintain. In this model, the name of the relationship (*emphasizing*) is taken from Basil.

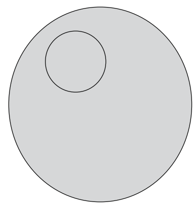

FIGURE 7
Emphasizing relationship.

---

[1] In the model, relationship a-b refers to the horizontal, a-c to the vertical axis. This illustration is made from the point of view of item "a", so we are not interested in the relationship b-c. The purpose is to analyze the relationships of item "a" to the other items. This way the two-dimensional model can be applied to analyze combinations of several elements. Otherwise, if all the relationships were illustrated in a single graph simultaneously, there should be n(n+1)/2 dimensions, where n is the number of output elements analyzed.

Another case that needs explanation is expressed by Ekman and Friesen (1969) with the word *contradict*. Basil (1994) uses the derived attribute *contradicting*. These expressions refer to a case where two messages are in conflict with each other. If the first message is "Mary has green eyes" and the second "Mary has blue eyes", the messages are contradictory. Evidently, only one can be true if we compare the messages with the entity they refer to. But because we earlier defined the model (figure 5) to handle facts, the false message is out of this range. It is possible that two persons might form messages that contradict each other, even though both persons think their message is true. But in such a case, the persons have different realities. For example, a color-blind person lives in a different kind of world than a person who is able to see all colors, which are usually defined as colors that are in the range of visible light. A hypothetical person with an ability to see colors whose wave-length is longer than that of red or shorter than that of violet has, respectively, quite a different reality than most of us. But because the environment in which this model is created primarily concerns combinations of messages that are coded by one person or a group of people that cooperate with each other, a case where two messages contradict each other is either a mistake or a really exceptional case. In other words, it is hard to find a reason for an intentional design of contradictory elements in a multimedia product. That is why this case, even though mentioned by both sources used, is left outside the model.

2.5 Creating redundant messages

The proposed model is not an end in itself, but rather a way to see the conceptual environment where redundancy is located. As mentioned earlier, redundancy is only a theoretical extreme of the dimension between it and the distinct case. Thus, it is argued that two messages are never perfectly redundant. To make this point clearer, an example of redundancy presented by Edwards (1992) has to be elaborated. Edwards's example is from everyday life, traffic lights. Traffic lights inform us whether we are allowed to go or have to stop, in two different codes: First, there are colors. Second, we are shown simultaneously whether we can go or not in another way: the position of a light provides the same information.[1] Thus, color-blind people are able to follow the

---

[1] Traffic lights as an example is brilliant also because they may be able to send even more than two simultaneous messages that inform whether you are allowed to go or not. In addition to messages that are based on color coding and positional information, the lights for pedestrians usually have either iconic information (two figures, illustrating walking and standing) or verbal information (e.g. texts "walk" and "don't walk"). Audio messages are also common. For example, in Finland most traffic lights for pedestrians inform you that you are not allowed to go by a broken beep, while a continuous beep denotes permission to go.

The combination of visual and audio messages in traffic lights serves as an especially clear example of the significance of redundant information in user-interfaces. The audio information is added to traffic lights, presumably, in order to help visually handicapped people. But the message is inevitably sent to all pedestrians who approach the traffic messages of traffic lights perfectly. At first sight, this seems a really pure case of redundancy. The possible messages, which could be verbalized as "stop" and "go", are so simple that redundancy seems perfect. But even in this extremely simple communication act questions about the real meanings of color and positional information could be made: Why it is just red that means stop? What is the message of the fact that prohibition is indicated with red and permission with green? Or what is the message of the order of the lights: prohibition first and above, permission only after and below that. Some anthropologist could consider whether these tell us something about our culture and society. Although this kind of reasoning can be claimed to be excessively detailed, it serves as an extreme case to show how each message can be interpreted deeper and deeper, and finally, something is different in the interpretations of the two messages. At least, this should be taken seriously to prevent naïve simplifications such as dividing different combinations of messages into two classes according to whether they are related or not, as well as to avoid calling some relationship of messages as "completely redundant" (as Haring and Fry did, 1979). For example, in the experiments concerning text illustration reviewed by Levie and Lentz (1982) this kind of simplification – or information reduction – has been made in order to carry out quantitative analysis. The question is whether some essential information has been lost in the reduction process, i.e., whether something important is missed when neglecting the part of a message that is not redundant with the rest when striving toward redundant combination. Levin and Lesgold (1978) handled the item by beginning with simplification ("completely overlapping"), then softening the expression by explaining that "they are redundant in that they convey the basic propositional content of the text."

---

lights. Finnish people, for example, are used to getting audio messages at traffic lights. When they visit the U.S., they are bound to have difficulties in recognizing when it is their turn to cross the street. They might be used to looking around when waiting for their turn, not to staring at the traffic lights. When they hear the sound change, they presumably confirm their perception by gazing at the lights before stepping in the roadway. The lack of audio messages forces them to act in a new way.

3 HUMAN ABILITY AND REDUNDANT MESSAGES: THE CONTRIBUTION OF ATTENTION STUDIES

Attention is not a new subject of interest in cognitive psychology. In 1891, William James discussed the same phenomena that were still under extensive research a couple of decades ago. Through different paradigms and different eras in the philosophy of science the interest in the phenomenon itself has stayed. Attention is a central quality of human interaction with the environment. Therefore, interest in attention studies arises especially when discussing human beings in a new kind of environment.

3.1 Definitions

Attention is a familiar word in everyday language. The scientific definitions of the concept have not given the concept some new meaning. They merely seem to be attempts to manifest the essence of the existing concept and the related word. Even one of the most productive researchers of the domain, Daniel Kahneman, refers to the dictionary (Kahneman, 1973, p. 3) in order to make sure that he has considered all aspects. James (1891) defined attention in this way: "It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects of trains of thought." (pp. 403-404). James's concept of attention is also visible in his list about the effects of attention: according to James, attention makes us perceive, conceive, distinguish and remember, and it shortens reaction time as well (pp. 424-425). Titchener (1908) used the phrase *clearness in consciousness* (chapter V) when he discussed attention. The phrase illustrates how he stressed the role of affections. Affections and attention were described to be in a reciprocal relationship with each other (chapter VIII): affection usually (in voluntary attention) precedes attention, and it is likely that there are no affections without attention. In Sperling's (1984) approach, the functions of different memory systems play a central role. He defines attention as a function of choosing information from a very short visual memory (Visual Information Storage, originally defined by Sperling) to store in short term memory when the capacity of visual storage exceeds the capacity of short-term memory. Swets (1984) gives a more general definition of attention as "behavioral processes that evidence an increment or decrement in the effectiveness with which an organism handles current information from a given source in its environment." Sack and Rice (1974), in turn, provide a functional definition by describing three characteristics of or qualities of attention:

- selectivity (the ability to establish the focus of attention),
- resistance to distraction (the ability to maintain the focus of attention), and
- shifting (the ability to switch the focus of attention from one target to another).

The verbal manifestations of the conceptions of attention among researchers resemble each other markedly. The agreement is not based on an all-inclusive, explicit verbal definition. The concept is probably too complicated and multidimensional to be expressed verbally. But the agreement must be quite strong since attention as a subject has retained its position through decades without major discussion about the definition of the concept. Attention can be seen as a loose conceptual framework when discussing related matters. Rather than an independent phenomenon, attention is a label of one branch of research in cognitive psychology.

3.2 Attention as an object of research

According to the most significant authors in the field, interest in research on attention has varied markedly (e.g., Broadbent, 1958, p. 109; Kahneman, 1973, p. 1). The most severe decline is argued to have taken place because of Watsonian behaviourism at the beginning of this century, continuing to the 50's. Behaviourism is seen to have swept all cognitive psychology from science.[1] The

---

1 There seems to be a general agreement that in the behaviouristic era (1910-1950) there was hardly any research concerning attention. Lovie (1983) marvels at this impression: according to his survey, the proportion of papers concerning attention (and related phenomena) among published psychological reviews has remained quite steady, about 0.1 to 0.4 percent of all papers, between 1910 and 1960.

Lovie discusses the possible reasons for such a strong belief. First, he supposes that the tone of the behaviourists was harsh, thus attracting much notice. On the other hand, the representatives of cognitive psychology might have overstated the dominance of behaviourism in order to be able to interpret the rise of cognitive psychology as a true Kuhnian revolution against behaviourism.

The author also argues that there is a clear continuity in the domain. However, since the point of view and the vocabulary have varied according to the prevailing paradigm, it is not easy to observe the connections between studies of different eras.

rise of cognitive psychology against behaviourism[1] has been seen as the start of modern research on attention.

Besides changes in scientific paradigms, pressures from outside the academic world have effected an interest in attention. Human functions concerning attention satisfy the needs of interaction with our natural environments. But when new environments are created, information about the human way of interacting with the environment is needed. The less the new environment resembles our natural environment, the more detailed information about human qualities is required to make interaction work. Especially, this has been the case when designing technological environments. It is probably not a coincident that James's work took place at the time of the industrial revolution. In the literature, the most common example of a practical need for expertise in human-machine interaction is, however, the growth of air-traffic in the 50's. The vast quantity of information handled by airfield control made the technology so complicated that the ability of a human being to control all the functions simultaneously had to be investigated (Kahneman & Treisman, 1984).

The rapid development of microprocessors and related hardware during the last two decades has resulted in the high performance of applications. That computational power has to be used for something, at least for commercial reasons. The needs of most usual uses of microcomputers, like word processing, have already been satisfied a long time ago. Thus, if there is no need to enhance the performance of a program, the other possibility for using the growing capacity is to increase the number of applications that run in parallel or the number of functions of an application. This tendency has lead to user-interfaces in which a vast quantity of information is accessible simultaneously – in theory. We need expertise about human ability to control a large quantity of different kinds of information in order to design user-interfaces in which information is accessible in practice, too.

3.3 Different approaches to attention

Because research in the domain of attention has usually been a response to practical needs to make human-machine interaction work, the point of view of the research has been one of human limitations. In other words, the experts of technology have asked for information about the counterpart of technology to be able to design a working whole. The human being is, in this view, one part of a system, and the system would not work if different parts were not fully coordinated with the whole. Human factors have been seen as a static framework, an item that cannot be changed but about which information is needed to be able to create technology that is in coherence with it.

---

[1] Van der Hejden and Stebbins (1990) question the contradiction between the information processing approach and behaviourism, partly by distinguishing the IP-approach from cognitive psychology; i.e., they seem to propose a shift in the meta-theoretical framework where the IP-approach is usually placed.

In order to find the real limits of human cognitive performance, the research concerning attention is one of the most obvious sources. Attention has been seen as a tool with which man can cope with more information than he can really process. Thus, attention is required in extreme cases where human limits are approached.

Different models and theories of attention have traditionally been divided into categories according to the approach (figure 8).

3.3.1 Structural vs. capacity models

The first major division between different approaches to attention is the division between structural and capacity models. Structural theories explain the limitations in the ability to process multiple information in parallel by assuming a bottleneck in which all information is processed in a serial system. (The dichotomy between the early and late selection theories relates to the question of the stage in which this bottleneck takes place: in perception or in response). Capacity theories assume that there exists a common pool of processing resources available, and when the demands of one or more tasks exceeds the amount of resources available, the performance deteriorates. (Wickens, 1984)

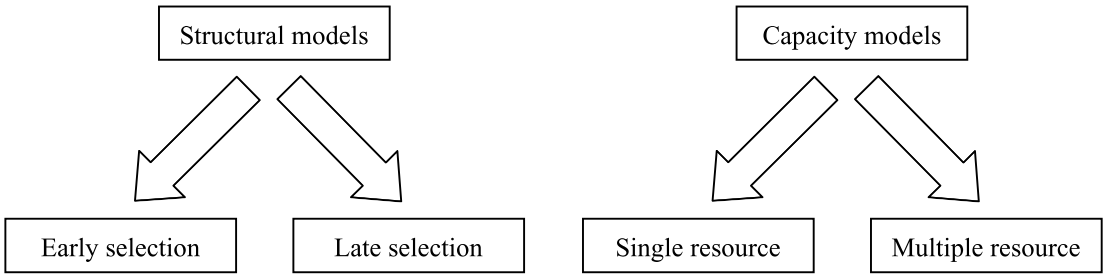

FIGURE 8 A rough classification of different approaches to attention.

3.3.2 Structural models

Structural models can be divided into early and late selection models according to the hypothetical location of the bottleneck in information processing. The early selection model emphasizes perceptual overload and the human need to select information to be able to concentrate on the essential. The late selection model is concerned with the difficulties in selecting an appropriate response. The assumed role and therefore the definition of attention differ according to the model. (Kahneman and Treisman, 1984)

3.3.2.1 The filtering paradigm (early selection)

In the 50's, the main problem concerning attention was considered to be perceptual limitations. The results of experiments were interpreted in the first theory of attention by Broadbent (1958). The model that is presented in the theory is usually known as the early-selection model. According to this model, information (stimuli) is stored only briefly at the pre-attentive level in a very short-term store. The capacity of that store is not defined. From short-term storage, information is transferred to higher levels via a limited capacity channel. Because only a limited amount of information can be transferred at a time, some kind of selection must occur. Some selection criteria are automatic, such as physical intensity and the time from the previous occurrence of information of the same class.

In Broadbent's theory, multiple tasks can be performed simultaneously, but the actual processing is always sequential, never parallel. When performing multiple tasks, it is a question of rapid switching between the tasks. While waiting for a turn to enter the limited capacity channel, information may be stored for a short period (a maximum of a few seconds) in short-term storage. According to Broadbent, the total rate of information flow determines performance, independently of the number of tasks (pp. 34-35, 298-299).

The most important findings and interpretations are presented in Broadbent's second chapter. The methods used in experiments to which Broadbent refers are usually known as dichotic listening and shadowing. In dichotic listening, the subject is presented different passages to each ear. Usually, a subject is told to concentrate on one passage and reject the other. In a shadowing task, a subject repeats continuously what he or she hears. A typical experiment was conducted and reported by Moray (1959). In these kinds of methods, the processing of the information presented to the unattended ear is usually investigated.

The explanatory power of Broadbent's theory is questioned (e.g. Kahneman, 1973, pp. 112-113) because it was based mainly on experiments about processing of audio information. This way, Broadbent avoided handling for example the complex processes concerning visual scan. But the cost of that simplification was that the results could be applied only to a very limited set of cases.

Filter theory was soon modified when evidence was found for processing of the unattended stimulus (Kahneman and Treisman, 1984). According to the updated theory, the filter only reduces the amount of information received in the rejected channel (Moray, 1959; Treisman, 1960). Anne Treisman (1960) concluded that the results of her experiment supported – to some extent – Broadbent's filter theory. But she found a slight processing of the rejected passage in a modified dichotic listening shadowing task experiment in which the channels were changed with each other in the middle of the passage. The subjects usually followed the same passage without even noticing the change of the channel. Most of them, however, repeated a couple of words from the rejected passage when the channel changed. Treisman interpreted this as a proof of limited processing of the rejected information. She proposed a modification of the filter theory: instead of an "all-or-none" barrier, the filter only attenuates signals from the rejected source. Treisman's article is usually referred to as the launching of the attenuation theory. However, this model also suggests a single bottleneck in all human information processing. Later, Treisman (1964a) strongly doubted the applicability of filter theory's channel metaphor in more complicated cases: Most of Broadbent's evidence was from dichotic listening experiments, in which two channels meant two ears of a subject. Treisman tested filter theory in conditions in which two passages with different content were handled as separate channels. She also tested whether two passages of different, equally familiar languages could be handled as channels of filter theory. She found severe interference and concluded that at least in those conditions filter theory could not be applied. She also developed the theory (Treisman, 1964b) by separating three different stages of the filtering (or attenuation) process: 1) discrimination of different channels, 2) rejection of irrelevant channels, and 3) semantic analysis of messages that are attended to.

In the 70's, evidence was presented that in certain conditions it is possible to divide attention between two or more tasks with little or no loss in performance compared with single task conditions (divided attention paradigm). This was said to be possible when there was "sufficient difference" among the tasks. (Allport, Antonis, & Reynolds, 1972; Kleiman, 1975; Rollins & Hendricks, 1980; Shaffer, 1975; Treisman & Davies, 1973). The core of the problem is, what is, in a qualitative sense, sufficient difference.

The analysis of qualitative differences among tasks inevitably leads to the analysis of the whole cognitive system. The resulting models are at least to some extent modular to be able to explain successful parallel performing of multiple tasks. In other words, there are assumed to be different cognitive structures that are employed by different kinds of tasks.

3.3.2.2 The selective-set paradigm (late selection)

Deutsch and Deutsch (1963) approached the phenomenon of attention in a different manner. They presented neurophysiological evidence about perceptional functions that are independent of attention. They argued that each sort of message has a preset weighting of importance. For Deutsch and Deutsch, attention was a concept that described the level of alertness. If the importance of a message is high enough for the current level of alertness, it will get a response (figure 9). In other words, the stage in which selection happens is not the perceptual level but the response level.

In figure 9, the bars illustrate single messages and the height of a bar indicates the importance of a message. Whether a message is able to cause an alert or not depends both on the importance of the message and the general arousal (illustrated with horizontal lines). In turn, if alerting takes place, the general arousal will rise. Thus, messages that occurred before alerting and were unable to cause an alert may be sufficiently important to generate a response after an alerting caused by a more important message.

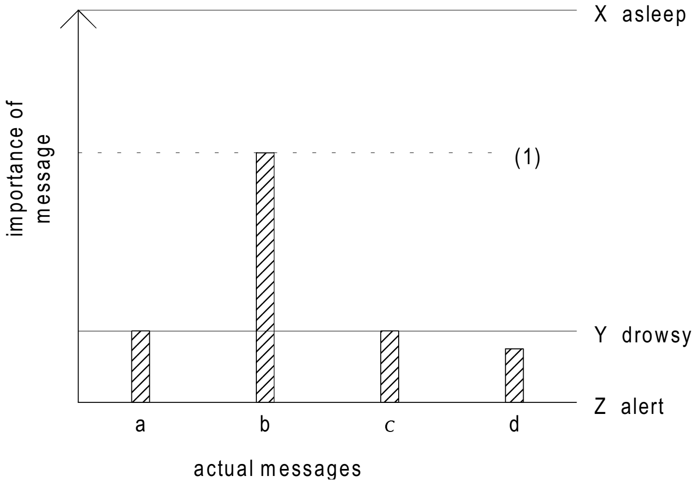

FIGURE 9 An illustration of the model of Deutsch & Deutsch (1963).

3.3.2.3 Early selection vs. late selection

The crucial difference in the explanatory power of early and late selection models is their relationship to the existing structures. In the modeling of information processing, the filtering paradigm relies on information presentation, whereas the selective-set paradigm is based on existing structures that are activated by an external stimulus.

Kahneman and Treisman (1984) discussed the traditional debate between the proponents of early and late selection models. They did not find the division between early and late selection models quite appropriate. Instead, they proposed an alternative approach, which can be seen as a synthesis of the two models. The authors compared perceptual processes with the opening of an object file, which is detected by physical characteristics. This implies that there exists a temporary storage of perceptual information, which activates the other stages of information processing according to whether it is a question of the creation of a new representation or of the updating of an existing one. The first case, a new representation, thus refers to conditions in which early selection models have been applicable. The updating of an existing representation refers to activating of an existing structure and has traditionally been explained within the selective-set paradigm.

3.3.3 Automatic vs. control processing

In the '70s, a major paradigm shift concerning the study of attention took place: the interest shifted from the limits of attention to automatization of processes (Treisman and Kahneman, 1984). The shift to capacity models is also a related phenomenon.

Walter Schneider and Richard M. Shiffrin (1977; Shiffrin and Schneider, 1977; Schneider, Dumais, & Shiffrin, 1984) have found the contrast between automatic and control processing essential concerning attention for several reasons:

- The nature of processing, or here, the level of automaticity, affects performance.
- Performance changes during a period of practice when automatic processes develop.
- When processes automate, the ability to control processes deteriorates.
- Control processes are limited by STM capacity unlike automatic processes.
- Control processing seems to cause modifications in memory.

In the development of automatic processes Schneider and Shiffrin separate two different kinds of practice: practice where stimuli and responses are consistently mapped (CM) and practice where they are variably mapped (VM). There is evidence that in order to develop automatic processes CM practice is superior to VM. (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977; Schneider et al., 1984)

The researchers who stress the role of automatization of processes see the question of limited attentional capacity as a result of competition between control processes. Automatic processes are of minor importance concerning limitations, since they have not been found to be resource consuming. Extremely complex dual tasks have been carried out by subjects in conditions where at least one of the tasks has required only automatic processing and the subjects have been consistently trained: A skilled typist was able to type from visual text in parallel with another verbal task (Shaffer, 1975). In another experiment, music students could shadow speech while sight-reading music (Allport et al., 1972).

Schneider and Shiffrin, the researchers who have been the most productive in the domain of automatization, have found it to be a key issue concerning attention and human capability. They also have discussed automatization of attention itself. They referred to search task experiments (detection) in their conclusions, but found the results applicable in other kinds of processes as well. Schneider and Shiffrin argue that:

In order to promote automatization of attention, the consistency of training has been found critical. However, consistency of one component is adequate – e.g., if the stimulus is consistent, no difference in performance improvement has been found whether the response is consistent or not (Schneider & Fisk, 1982). It is also important that the consistent, repeated task is always *completed*; the mere attempt is not enough.

It has been found that automatic processes have shown high transfer in the same class of stimuli; Schneider and Shiffrin conclude that this prescribes training that is consistent in real situations, too, while in reality, conditions in two different situations are rarely exactly alike.

Automatization is found to be faster the greater the dissimilarity between a target and a distractor is. However, unpleasant experiences of prior CM training and all VM training deteriorate the automatization process. Tasks requiring accuracy automate faster than those requiring rapid responding. (Schneider et al., 1984)

As a summary of the functions of control and automatic processes, it can be said that control process is an *instrument* in the process of automatization. But there are processes that cannot be automatized and must therefore be processed controlled. Control processes activate the nodes in memory thus enabling automatic processing. And in order to block or modify automatic processes, control processes are necessary.

As a typical example of automatic processes Schneider et al. (1984) state the execution of habitual behavior. Automatic processes can interrupt control processing and reallocate attention and resources.

On the other hand, no process at all, according to the definition of Schneider and Shiffrin, can be said to be purely controlled or automatic. All processes have at least some characteristics of both. So, when Schneider and Shiffrin call some process an automated process, they actually mean that the process is automatic to a high extent.

3.3.4 Capacity models

Structural theories explain the limitations in the ability to process multiple information units in parallel by assuming a bottleneck, in which all information is processed in a serial system. The dichotomy between early and late selection theories relates to the question of the stage in which this bottleneck takes place: in perception or in response. In capacity models, the idea about strictly serial processes is rejected. Instead, these models are based on the concept of mental resources that are used by a procedure to a greater or lesser extent (figure 10). (Wickens, 1984)

Capacity models are based on a dual-task paradigm. Human performance in multiple task conditions has been explained on the basis of both structural and capacity theories. However, these two traditions have developed quite independently (Wickens, 1980).

Structural theories describe divided attention in multiple task conditions as a process in which two or more tasks use certain mechanisms (structures). When the same mechanism is required by two or more tasks simultaneously, performance deteriorates compared with single-task conditions. Thus, these models divide information processing into sub-processes that are performed by mechanisms that cannot process several information units simultaneously. The mechanism is either used or not, and when it is used, it cannot be used by another unit. In divided attention, it is rather question of rapid switching between tasks than parallel processing.

According to capacity models, in multiple task conditions there is a question of resource allocation concerning each single task (figure 10) and between each task (figure 11).

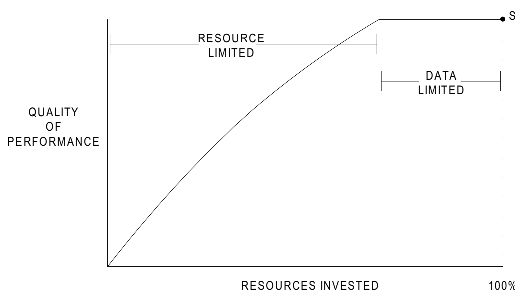

FIGURE 10 Performance-resource function (PRF). (Wickens, 1984)

Figure 10 presents a hypothetical performance-resource function (PRF). Wickens pays attention to the obvious, yet not profound investigated fact that in experimental situations not all available resources are invested in performing a task. The amount of resources invested is not easily measured in single task conditions, but Wickens presents a proposal that has been derived from the results of dual-task experiments.

According to PRF the quality of performance is a non-decreasing function of resources invested.[1] The quality of performance increases to the point where the task can be carried out perfectly or all resources are invested or both. Until the stage where the task is perfectly performed, the task is resource limited. If that stage is reached before all resources available are invested, the rest of the task is data limited (the task is performed perfectly whether more resources are invested in it or not.)

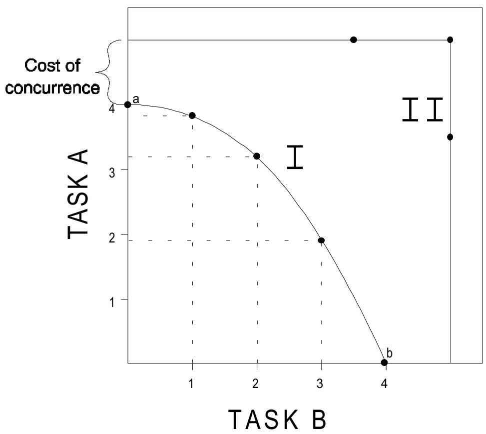

FIGURE 11 Two different kinds of Performance Operating Characteristic (POC) curves. (Wickens, 1992, p. 370)

The curves in figure 11 illustrate dual-task performance in different conditions. Points a and b indicate the single task performance. In curve I, the performance of task A deteriorates as a function of resources invested in task B, and vice versa. That is, when resources are removed from task A, task B is able to use them. Thus, the tasks share resources. Curve II illustrates single task performance. The difference between the starting point of curves I and II is explained as a cost of concurrence: the mere existence of another task reduces the ability to perform the primary task even if all resources were devoted to it. If the shape of a POC curve of two tasks resembles curve II, the performances of the tasks do not interfere each other. This could mean either that the tasks do not share resources or that both tasks are in the data limited area (see figure 10). In other words, the tasks are different enough or easy enough or both. There is empirical evidence of this kind of perfect dual task cases (Allport et al., 1972; Wickens, 1976; Shaffer, 1975). Wickens uses an expression *time-sharing* when referring to the phenomenon.

---

1 The idea is quite analogous with one of Rubin's (1993), according to which it is not appropriate to divide the audience into "passive" and "active" individuals since real cases fall between these extremes.

Multiple Resource Theory (Wickens, 1984) is introduced as a logical implementation of these findings. The theory proposes, instead of a single "bottleneck", an existence of several resource-like structures in human information processing. Wickens classifies the task according to the required sensory modality (auditory, visual) and the type of code (spatial, verbal). There is evidence of greater interference between tasks that require the same modality and use the same coding system (e.g. two tasks, both of which are visual and verbal) than between tasks that are different concerning either of these properties. But the evidence is insufficient to a conclusion that time-sharing between tasks that are different in this sense would be perfect. Therefore, Wickens assumes that there exists hierarchically above the distinct resources a common pool of undifferentiated resources, available for all kinds of tasks and stages of processing. (Wickens, 1984; 1992, pp. 378-382)

3.3.5 Attention and visual search

Visual search is often handled as a separate paradigm concerning attention. Knowledge about visual search can be applied so widely that the number of studies in the domain is remarkable.

Rabbitt (1984) considers visual search as an active process, in which we search for meaningful patterns of relationships in order to decide two things:

1. where to look and
2. what to look for next.

He also discusses the human way of learning and optimizing the control processes required.

There is much work done in the domain of visual search. The results from these studies are convincing enough for Rabbitt to argue, that we *know* that people

- effectively learn new search strategies,
- store them in LTM, and
- access the most appropriate one very rapidly from LTM when needed
  (Moray, 1978; Rabbitt, 1981).

According to Neisser (1967) the process of visual search can be divided into two separate stages: pre-attentional and attentional. In the pre-attentional stage subjects extract some global characteristics of a display. Neisser mentions (p. 89) figure, background and contour as such features. Then, in the attentional stage more detailed analysis takes place. Several studies have shown the attempt to find meaningful wholes as a general way of visual search. For example, the detecting of three-digit numbers deteriorated when the number of colors of display items increased (Cahil & Carter, 1976), but when shown a text in which alternate lines were printed in red and black, subjects were able to read the text in one color with no interference from the lines with the other color (Willows &

McKinnon, 1973). So it seems that the subjects try to detect some kinds of structures as a starting point for further search.

Some examples give reason to assume that the pursuit of meaning precedes the detection of details. We possibly try to classify an object first and search for details only after that. When a subject is shown a high-contrast black-and-white photo, the figure of which he or she is not able to recognize and then is given a cue (e.g. a verbal explanation), all the details fall into place at once (Frisby, 1979 p. 20). Some studies (Biederman, Glass, & Stacy, 1973; Biederman, Rabinowitz, Glass, & Stacy, 1974) show that when single objects of scenes are presented in conventional or meaningful relationships with each other, the scenes are both scanned faster and remembered better. Thus, previous experiences seem to guide the scanning of scenes.

Categorization is an important strategy for visual search. Rabbitt refers to the ingeniously constructed experiment of Corcoran and Jackson (1979) as evidence about categorization when a single, critical feature distinguishes the target from the distractors. The subjects of their study were to search for the character "Ø" among sets of either straight-line letters (e.g. X, K) or curved-line letters (e.g. O, C, G, S). The researchers noticed significant transfer when switching between different sets of the same type and even when switching from a set of one type to a set of another type (e.g. from a set of straight-line letters to a set of curved line letters). But, when switching to a set of all kinds of letters, the transfer was negative. Rabbitt (1967) and Neisser (1963) have noticed the same effect. Rabbitt draws this conclusion: categorization is effective when subjects can base their decisions on a single characteristic (here: / or O).

The process of categorization gets far more complicated when the number of features on which the decisions must be based increases. Rabbitt (1959) and Pollack (1963) have discovered that "*categorization times increase as a multiplicative function of the number of categories discriminated and the number of items within each of these categories*" (Rabbitt, 1984, p. 281).

Ingling (1972) discovered that it is easier to find a letter among digits than a letter among letters, and vice versa: it is easier to find a digit among letters than among digits. But there is also evidence that subjects succeeded as well when they were to find any letter among digits as when they were to find some particular letter among digits (Brand, 1971; Gleitman and Jonides, 1976; Jonides and Gleitman, 1976; Egeth, Atkinson, Gilmore, & Marcus, 1973). Thus, subjects recognized that a certain symbol was a letter or digit before they had identified which particular character it was.

Green and Anderson (1956) conducted an experiment, on the basis of which they argue that when symbols differ from each other in two dimensions, in certain conditions subjects can search for symbols with one property with no interference from the other property. For example, if they search for a red triangle surrounded by symbols of different colors, the range of shapes of distractors does not affect the performance. Rabbitt (1978) argued that that kind of serial search is possible only if the two dimensions are separable in the way that Garner and Felfoldy (1970; Felfoldy and Garner, 1971) defined it: some pairs of dimensions can be independently processed (e.g. color and shape), while other pairs are integral (e.g. color and brightness) and must be processed simultaneously.

If subjects optimize the search on the basis of two dimensions, they should first recognize which dimension allows the maximum reduction of the search. For example, if the number of possible colors is five and the number of shapes is two, the optimal strategy would be to first search for the color. Although there is evidence about such optimized search strategies, in complex situations this is seldom the way we act. Rather, it seems that when searching repeatedly for a complex item, subjects first compare it in a holistic way with the preceding item. If they are identical, a rapid response results. Otherwise, the analysis continues (Rabbitt, Cumming, & Vyas, 1977, 1979b). In other words, subjects place the finding of similarities first. But Fletcher and Rabbitt (1978) also found that the more subjects were practiced, the harder they searched for the identity of each item and the change between successive ones.

Rabbitt (1967) reported an experiment in which subjects practiced searching for certain targets among a certain set of background symbols. When the set of background symbols was then changed, negative transfer occurred. This indicated specific cue learning. In further investigations this learning was proved to be retained, with little practice, for four weeks. In other experiments (Prinz, 1979) the conditions were altered by replacing the target or background symbol set after practice with a new one or target symbols with background symbols or vice versa. The level of negative transfer was highest when target and background symbols were reversed.

According to Rabbitt, people change their searching[1] strategy when practiced. First, they try to remember which symbols were targets, and then they try to identify each item as an individual. In the second stage they learn to select optimal cues (Corcoran & Jackson, 1979; Neisser, 1963; Rabbitt, 1967). After extensive practice, the subjects do not need any particular cues but are able to use a wide variety of possible features of targets to detect them (Rabbitt, Cumming, & Vyas, 1979a; Kristofferson, 1977).

Bruce (1979) conducted three experiments concerning visual search. In them, the subjects searched for given faces among distractors. The most interesting findings were that search times were longer the more targets and distractors visually resembled each other, and that it was easier to detect politicians among actors than politicians among politicians. When searching for politicians among actors, visual similarity was no longer significant. On the basis of Bruces results, Rabbitt argues that people can independently recognize perceptual features of an item and classify it. This can be interpreted as evidence about the automaticity of semantic processing.

1 Rabbit (1968) expanded the domain of his assertions to cover more than only visual search by using the term *strategy of perceptual analysis.*

3.3.6 The visual search paradigm and the dichotomy between structural and capacity models

The visual search paradigm was handled separately here because it forms a more or less independent branch in the studies concerning attention. On the other hand, it provides a unique perspective to the domain, and discussing the other paradigms from its point of view is worthwhile.

In the visual search paradigm, the dichotomy between structural and capacity models can be seen as a dichotomy between models that rely on serial or parallel information processing (Egeth & Dagenbach, 1991). In the contemporary IP-approach, this dichotomy is often used as a basis when characterizing different stages in information processing.

Even in the 90's, Egeth and Dagenbach (1991) continue to find relevance in the model of Neisser (1967). They present Neisser's division of attention into pre-attentive and attentive stages as a typical approach of studies of the domain. According to them, this approach implies an assumption that the pre-attentive level is not capacity limited, thus enabling parallel processing. Allocation of resources takes place in the attentive stage, and the allocation of resources often results in processing that is characterized as serial. The pre-attentive level in visual search refers to early perception processes. Egeth and Dagenbach present results of their experiments that are interpreted to support this approach.

However, the dilemma of serial versus parallel processing is, despite extensive research done especially since the late 50's, far from solved. Townsend (1990) discusses with numerous examples from the history of research in the domain, how the same data can be interpreted as support for quite contradictory models. In certain cases, the possibility of contradictory interpretations is obvious. Especially in dual-task experiments, in which the performance of a certain task is poorer in dual-task than in single-task conditions, there are several possible ways to interpret the data: the same phenomenon can be seen, for example, as

- time-sharing between tasks in a serial process (filter theory)
- time-sharing between responses (late-selection models)
- a result of using two resource-consuming control processes simultaneously while neither of the applied processes is automated (automatization approach)
- allocation of common resources, inadequacy of resources (capacity theory)
- allocation of common resources between some sub-processes (multiple resource theory).

Thus, the debate is likely to continue while trying to create methods with which the nature of mental processes could be identified.

3.4 Attention models and theories in the design of multimodal information presentation

3.4.1 Filtering

Neville Moray, one of Donald Broadbent's students from the 50's and 60's, presents an interesting view of the last three or four decades of the study of attention in 1993. He once more defends filter theory as an applicable tool in the design of human-machine systems. Moray does not even try to deny the numerous results that conflict with Broadbent's theory. Instead, he takes an extremely practical approach: Moray rejects the old idea according to which theories of attention could serve as a tool to figure out the functions of human information processing in detail. He finds it adequate if we had a simple model that could predict human behavior in most cases on a level that does not give an over-optimistic idea about human performance. Or, as Moray expressed it, human performance yields at least the level that filter theory describes.

Moray's latest approach seems to be a reaction to studies that went deeper and deeper when trying to find a sound basis for a perfect human information processing model. Moray saw that the traditional approach did not lead to any applicable model. Or as Schweickert (1992) put it, Broadbent's theory was the last synthesis: the newer proposals are only like separate branches. However sophisticated a model was created, it was bound to be rejected sooner or later on the basis of some new experiments. To some extent, the tendency was cumulative; i.e., the new models only revised or complemented older ones. But the problem of revised models was that they made the original model more complicated and less useful in the practical design of human-machine systems. And when some simple, easily adopted model was revised a couple of times, the temptation to reject it and create a new one was naturally significant. So, Moray seemed to find it reasonable to return to the roots, to the filter theory of Broadbent, and argued that design of high quality can be based on it: it clearly covers most cases. Its superior usability outweights its deficiencies. On the other hand, filter theory has got support quite lately in a neurophysiological experiment (Mangun, Hillyard, & Luck, 1992). Early selection in modality-specific cortex was found clear concerning visual-spatial attention.

The revised versions of filter theory can be seen to soften the radical characteristic of Broadbent's original version. A good manifestation of this softening is that Treisman's theory (e.g., Treisman, 1960) is usually referred to as attenuation theory. The name of the theory clearly refers to contrast between this theory and its origin: rather than filtering information we are not focusing on, we attenuate it.

An interesting question is what are the conclusions about the refinements of filter theory. If the basic objective of constructing a model is to use it as a means to model human information processing, all refinements and conclusions about experiments should, of course, be taken into account. But if the target is to create applicable design recommendations, the conclusions about the worth of detailed work can be different. For example, it is likely that the conclusions in practical design decisions will be the same whether we assume that unattended information is rejected (filtered) or that unattended information is processed at a markedly lower level than attended information. In both cases, the conclusion is that it is not effective to deliver information when the attention of the user is likely to be somewhere else. Thus, Moray's opinion about the usability of Broadbent's theory is understandable.

In a design project, an interdisciplinary approach is vital (Kim, 1990) as already discussed in the introduction. A multimedia production project is such a complicated project that it needs expertise in many domains. However, as shown by Benson (1993)[1], the views of experts of different domains may vary markedly concerning the same issue. Therefore, an interdisciplinary approach should not be understood as a group of experts in different domains. Instead, not only the project group but also all participants should be – to some extent – interdisciplinary. They should all be aware of the guidelines or the applied principles of the domains of the other members of the group. For example, when adopting a cognitive model as a basis for human-computer interaction design, all members should be able to understand in a rough form the underlying theories. It is not realistic to require high expertise in different disciplines from all members of a large and heterogeneous design project group, but all must be willing to increase their expertise in various areas to be able to communicate in a group and work for common goals. In this kind of approach, Moray's (1993) idea is reasonable: an easily learned, simple, and in most cases relevant model, filter theory, could work as a model that is adopted by everyone. The refinements or extensions of filter theory (or whatever theories and models based on Broadbent's theory) could then be left for the specialist in cognitive psychology and could still be applied when necessary.

Another question is the relevance of structural models in design. Structural models seem to aim at constructing an overall idea or model of human information processing. In other words, their target is not only to help designers cope with the human part of human-machine interaction, but the research is also an end in itself. So, if we are satisfied with a model that is good enough for most design decisions and reject all efforts to try to go deeper, we have an extremely specific, although practical, model. But when the environment changes, like the technical environment of multimodal information processing, the model that worked in the old environment may no longer be valid in the new one.

The problem is quite contrary to the one that took place in the positivistic era. Then the challenge was to adopt the methods of exact sciences to the humanities by explaining human behavior on the basis of laboratory tests. The determinants of the human mind were supposed to be found in an inductive manner, by summing up details that were found in the quantitative data of

---

1 In a study conducted by Benson three different groups of experts were to analyze the problems in integrating text and picture. She found significant differences in the interpretations of these problems among groups of subject matter experts, professional designers and editors.

carefully designed experiments. That approach failed in many cases because it did not cope with such a complex subject as a human being. But in a practical or more synthetic approach more determinants are taken into account, but only of a specific setting.

3.4.2 The selective-set paradigm

The idea that the actual bottleneck in information processing takes place in the response stage is interesting from the point of view of human-computer interaction. The channels of information flow between a human being and a machine in a multimedia workstation, for example, are far from balanced. The output modalities of a computer are numerous and vast in capacity. Above all, the capacity of a display to present visual information is unbeatable in this interaction. But the means of the human counterpart to enter information or – in the vocabulary of behaviourism – to give a response, are few and narrow. From the point of view of late selection models, current technology makes the bottleneck of information flow even more tight compared with earlier stages of human-machine interaction: more and more information is presented to the user without providing the user with new tools to cope with the swelling information flow.

If the results of studies supporting late-selection models are combined with a view of the user as an active, intentional processor of information, suspicion arises about the tendency toward pictorially biased modalities in information presentation in user-interfaces. If it is desirable for the activity of the user to have a concrete manifestation, the designers should think seriously about the possibility of creating new devices in order to provide the user with a wider range of means to express herself or himself when interacting with – or with a help of – a computer.

The late-selection model of Deutsch and Deutsch (1963) is applicable in another respect, too. The model is an applicable framework for discussing whether an alerting is going to take place or not. In other words, the authors pointed out the critical considerations when trying to make a human being aware of an information unit (or receive a message, depending on the discipline): general arousal and the importance of a message. So, the task of the designer or the author of an application is:

- to heighten and maintain the level of alertness and
- to show the importance of a message for the user.

These conclusions are trivial and can be drawn from almost any theory concerning activation and motivation, but at least they provide a bridge to the other theories and show that the results can be quite equal whether we stress attention or learning as a theoretical framework.

If the tasks mentioned above are arranged hierarchically, they should be in an opposite order since the subjective importance or significance inevitably results in higher level of alertness. But it is not reasonable to reject the first one as a goal even if it is a natural consequence of the other because, according to this model, the higher level of alertness also makes it possible to get a response to messages with lower subjective importance due to lowered threshold (see figure 9). Thus, according to the late-selection model of Deutsch and Deutsch, presenting information that is important for the receiver can be seen as a means to heighten the level of alertness to be able to get a response to other messages with less subjective interest.

Although the model of Deutsch and Deutsch has clear deficiencies (Kahneman, 1973, p. 124) and has been revised (e.g., Norman, 1968), it was a significant attempt to take an alternative approach. Neisser's (1967) synthesis of early- and late-selection models is not separately discussed here since its implementation in design would probably be included in the implementations of the models it combines.

3.4.3 Automatization

The question of automatization is interesting when discussed from the points-of-view of user-interfaces and the user as an active learner in parallel. The advantages of automated processes are described as high effectiveness and low resource consumption, resulting in superb effectiveness/cost-ratio. But automatization takes place at the cost of active control. Especially from the point of view of modern comprehension about learning, automated processes can not be seen as a target in general. Human control is the central instrument when processing information in an active manner. The challenge of the designer is to classify user tasks in a detailed hierarchy and to set targets for the final processing level. In other words, the designer has to:

- divide user-tasks into atomic sub-tasks and
- decide, according to the nature of each sub-task, the level of desired cognitive activity concerning it.

As automatization of processing releases resources, it is reasonable to strive for that end on the following conditions:

- *The process, while not automated, shares resources with one or more other processes.* By automating the process, most shared resources are released for the use of the other processes.
- *The automatization process itself does not share resources with any process concerning the contents of a product.* Since automatization is not an end in itself, it is not reasonable to use resources that are shared with vital processes if the automated process is not expected to be in prolonged use. In other words, if the automated process is seen as a tool, instrumental to performing something more important, automatization can be seen as a process of preparing the tool. It is reasonable to prepare the tool if it is useful in several situations and for a long time. But if the tool is for specific task and the preparation needs resources shared with actual task, some other, existing tool should be used.
- *The process is distinct from the contents of a product.* In order to perform high level mental operations with the subject of the product, control of information processing is key. Automatization of mental processes that deal with the content hinders the

user from processing information in a free manner. And the ability to process in a free, unrestricted manner is important in order to construct knowledge. As an example, patterns of thinking processes can be seen as automatization of mental processes that prevent the user from thinking in a creative manner. But if the process has nothing or almost nothing to do with the content, automatization can be beneficial while releasing resources. For example, it is reasonable to automate the control procedures of a graphical user-interface by standardizing them and thus release resources from thinking about the meaning of each object.

- *No need to control the process is expected.* If automated, the process should work in exactly the same form in all situations in which it is applied. All modification and revision of the process is extremely resource consuming.

These requirements overlap with the conclusions of capacity models since Schneider et al. (1984) refer to consumption of resources when dealing with automatization.

3.4.4 Capacity models

Capacity models, especially the Multiple Resource Theory of Wickens (1984), provide a credible framework concerning attention. Wickens' theory is very flexible in the sense that it is quite abstract. Most studies can be interpreted from this point of view. But the same flexibility or abstractness is its weakness as well. Since it is an abstract framework and does not analyze information processing in detail, it cannot be the only basis for design recommendations. Rather, it provides a sound basis for empirical research. A designer is hardly interested in the fact that some processes have a common pool of resources and some have dedicated resources. As a basis for decisions a designer needs to know which particular processes share resources.

The abstractness of capacity models also decreases their explanatory power. Finally, all empirical findings can be explained to be in harmony with capacity models while capacity models seldom specify processes. If in dual-task conditions the tasks can be performed without interference, they are interpreted either:

- to have a common resource that is adequate for both parallel processes or
- to have dedicated resources.

The thing that the designer is interested in is which one of these possibilities is true. If the processes share resources, it is probably not wise to design such a combination of elements that requires execution of both processes in parallel: even if in laboratory conditions the subjects succeeded in sharing the resource, there is a risk that in other conditions, with other subjects, the result will not be as good. The designer would prefer a combination in which given user tasks have dedicated resources. For that practical design problem, capacity models have little to offer.

3.4.5 Visual search and signal detection

Signal detection, especially visual search, is, in contrast with capacity models, a paradigm that is extremely practical in nature. The related research usually consists of carefully designed laboratory tests and very usable conclusions based on extensive quantitative data. Models seem to be formed in an ambitious manner: they are strictly anchored to empirical data, and they are constructed inductively. But many of the primary results are applicable even without theoretical reasoning, on the basis of common sense.

The strength of the visual search paradigm concerning user-interface design is that design recommendations can be derived from the related research easily. These recommendations are related to visual information presentation and visual access, answering the question: How should information be presented to a user to ensure effortless reception of information? The research in the domain provides plenty of answers. Targets should differ from other items as much as possible – in color (Rabbitt et al., 1977, 1979b), shape (Neisser, 1963; Corcoran & Jackson, 1979) or semantic category (Ingling, 1972; Bruce, 1979). Similarities between successive displays are detected first (Rabbitt et al., 1977). These are but a few examples of practical design recommendations derived from the reports concerning visual search.

Data collected in laboratory conditions is often blamed for lack of transfer to real situations. Experiments conducted around the paradigm of visual search is, from this point of view, exceptionally applicable in user-interface design while test conditions are often very near to the conditions of everyday work with a workstation: in the experiments, information is usually displayed on a cathode ray tube or similar device.

The results of the research concerning visual search can, of course, be blamed for minor significance in user-interface design since they can usually be used only in the design of visual details. Thus, they can be seen as a tool for final revision only. On the other hand, ease-of-use and user satisfaction may often depend primarily on these details. If proper finishing results in user-satisfaction, it is surely worth paying attention to.

3.4.6 Summary of applicability assessments

The models and theories of attention have been objects of academic dispute for decades. They have been set against each other, and support for each has been searched for among empirical data. From the point of view of user-interface design many of these contradictions seem inappropriate. When practical conclusions for designer's use are drawn from these models, they seem to resemble each other astonishingly. Even Charles D. Wickens, the most significant developer of capacity models, found (Wickens, 1980; Wickens & Carswell, 1997) that structural models and capacity models actually handle the same phenomenon with different concepts: he proposed that processing structures actually define capacity. Similar conclusions can be drawn concerning many other theories as well. For example, automated processes is a framework that provides concepts that could be used by other paradigms of attention. When a process automates, it releases certain *structures* for the use of other processes thus enabling better dual-task performance. When people quickly learn new *search* strategies (Moray, 1978; Rabbitt, 1981), the new way of acting in a search task is automated.

The relationship between different models becomes even more evident when considering the possible recommendations for designers. The recommendations derived from different models would be in most cases in harmony with each other. Different approaches stress different characteristics. They can be seen as different points of view on the same object. They do not exclude, but rather confirm, each other.

The underlying model of the structure of human information processing seems to have developed along with the philosophy of science, not along with changes in the popularity of different approaches to the domain of attention. In the positivistic era, there were high expectations for psychology to explain mental processes as a sum of empirical findings. Structural models provided a suitable approach for those efforts. Now that the diversity and complexity of mental processes have been acknowledged, it has become reasonable to speak in a more abstract manner. Capacity models, for example, have represented that kind of approach.

After all, regardless of the school or era, the human being has not changed as a species. For example, if dichotic listening was difficult in the 50's, it is probably difficult today. Research on attention provides a vast resource of information about human mental processes that could be applied by user-interface designers.

3.5 Attention and redundancy

Designing redundant combinations of output elements can hardly be an end in itself. It is easy to argue on an abstract level the benefits of using multiple modalities and redundant messages. But real cases are never that simple. Every single case has its particular nature and has to be considered individually. As Hsia (1977) put it: there should be an optimal quantity of redundancy, neither too little nor too much. The question from this perspective is, in which conditions redundancy can be beneficial and in which conditions it can be assumed to cause deteriorating of communication. The challenge of a designer thus contains two aspects:

- to avoid such combinations of elements where the processing of one output element interferes with the processing of one or more other elements, and
- to design redundant combinations of elements where something worthwhile can be achieved with the combination.

These two should not be seen as consecutive stages of quality of combinations, i.e., it should not be interpreted that a designer has to complete the first task before the second one. In many cases, these two efforts can simply be contradictory and therefore some compromises have to be made. Typically, a picture and its caption do not fulfill the first requirement, but still is frequently a reasonable combination: it is impossible to focus sight in two locations simultaneously, and in this sense, a picture and a caption inevitably interfere with the processing of each other (aspect 1). However, a combination of a picture and a verbal presentation of the same item is often such a powerful way to present information (Bernsen, 1994), that it is reasonable despite the interference (aspect 2). In a multimedia environment, the interference can often be avoided by providing the verbal message in audio format.

The possible interference is interpreted here from the point of view of attention. Since most theories of attention handle human limitations, the use of these theories is primarily related to the first aspect mentioned above. The core of the question related to redundancy and attention is, how to construct combinations of output elements so that the processing of one element does not interfere with the processing of another, or so that the interference would be as little as possible. The same question can be expressed in several ways, depending on the approach, e.g.:

- What kind of combinations of tasks have the most effective time-sharing? (filter theory)
- What kind of tasks can be performed without sharing resources? (multiple resource theory)
- In which conditions is the ability to detect two independent signals maximal? (signal detection theory)

The possible findings and conclusions concern, of course, not only redundant conditions but all information presentation; in most of the related experiments the combined elements (tasks, stimulus) are totally unrelated (e.g., Allport et al., 1972; Moray, 1959, 1975; Shaffer, 1975; Treisman, 1960, 1964a, 1964b; Treisman & Davies, 1973; Wickens, 1976).

There are two reasons for handling redundant conditions separately. The first is the main motivation of the current work: if the combining of two elements without interference turns out to be impossible or very difficult, it would be hard to persuade any designer to adopt such a new design principle. But if we can argue that by following simple guidelines it is quite effortless to combine elements in such a way that they do not interfere with each other, we can continue the discussion about redundancy. In other words, the task is to use the theories of attention to show that it is possible to use redundancy in information presentation with no disadvantage. What the possible advantage could be, is discussed later.

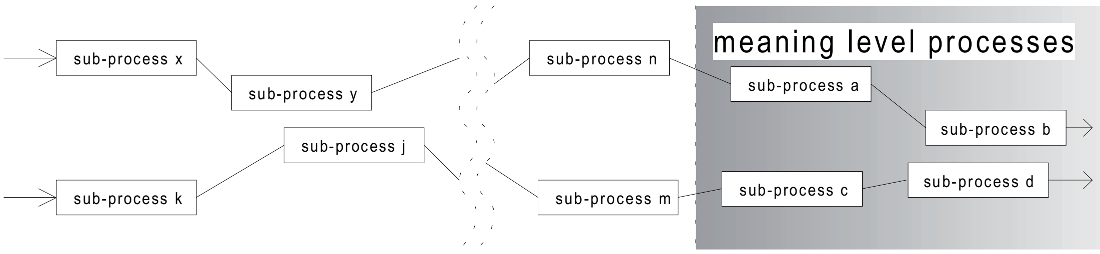

FIGURE 12. The sequence of parallel sub-processes in ideal dual-task conditions.

The second reason for concentrating on redundant conditions is that redundancy forms a special case among all possible kinds of relationships among human information processing tasks. Its unique characteristics can be illustrated with a chart, that is coherent with the way that mental processes are usually illustrated in the IP-approach: it highlights the flow of information through sub-processes (figure 12 and figure 13).

Figure 12 illustrates the processing of two unrelated mental tasks when the tasks do not interfere with each other: the two sequences do not share common sub-processes. Even the output (or responses) are separate. In redundant conditions (figure 13), the processes are separate until meaning level processing.

In the current work, the concept of *task* refers to the interpretation of a message. Thus, figure 13 illustrates a case in which two messages are received and processed separately until they actually blend together at the meaning level: they are no longer two separate tasks that compete for resources but a meaningful whole, whose origin is in two physically separate messages.[1] Thereby, once the meaning level in the processing of two redundant messages has been reached, there is no more concern about conflicts between two competing processes.

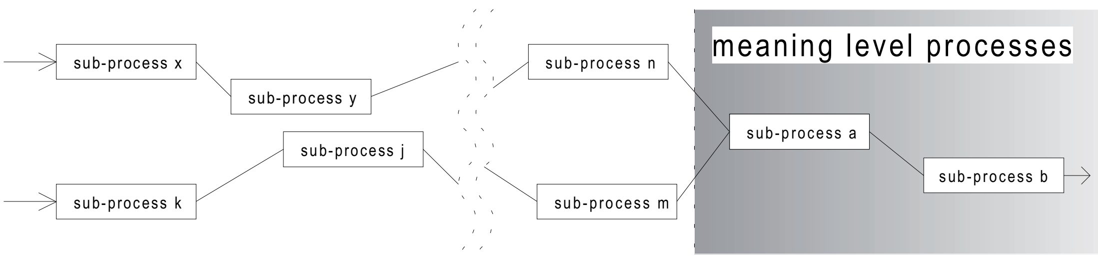

FIGURE 13. The sequence of parallel sub-processes in ideal redundant dual-task conditions

As shown, theories and models of attention provide an empirically verified framework for the analysis of multimodal information presentation. Research on attention is a sound basis for design recommendations and evaluation criteria concerning the user-interface. It provides a suitable conceptual

[1] The illustration highlights the sequential nature of the performance of different sub-processes. It should not be interpreted as a strictly linear process, even if the sequences are drawn - just for simplicity - as a path from left to right.

environment to analyze the quality of cognitive processes required in the receiving of redundant messages.

The major encumbrance of the applicability of research on attention is its IP-orientation. Even though the IP-approach provides concepts that are clear and illustrative, it is a mechanical view of something that is finally much too complicated to be completely explained with the means of science. If the true limits of the approach are recognized and acknowledged, the approach can be used in an appropriate context. Therefore, attention studies should not be seen as the sole framework of design or even of the current detail, redundancy in design. Rather, they provide one important point of view, a vital complement of less verified approaches.

4 THE CONTRIBUTIONS OF DUAL CODING THEORY AND INTERACTING COGNITIVE SUBSYSTEMS THEORY

In the present work, the models and theories of attention are the primary theoretical framework concerning cognitive processes in information reception. The strength of these theories here is the firm theoretical basis, the long history, and the clear analogy between the problems of information presentation in multimedia workstations and the problems that originally gave rise to the studies of attention.

Most literature concerning attention consists of empirical settings and theoretical conclusions. However, the theories and models are mostly on a quite general level, probably in order to cover as much empirical data as possible. The same problem was already discussed (p. 55) when comparing the applicability of structural and capacity models with each other. A highly conceptual model is easy to defend but difficult to apply. Therefore, two theories that provide quite detailed models of cognition are presented with discussion about their contribution to the functional definition of redundancy. The first of them, dual coding theory (DCT), is discussed since it has been used as a theoretical framework for educational applications in multimedia environments (e.g., Mayer, 1992). The second one, interacting cognitive subsystems-model (ICS-model), is worth discussion here because multimedia is its primary application environment. Both of these are IP-oriented, as is attention studies. Hence, it is possible to analyze them within a common conceptual framework. In addition, both provide useful concepts for the current needs.

4.1 Modalities in multimodal information presentation

Because both of the theories presented here stress the significance of the form of the message containing the presented information, a brief overlook of the basic concepts concerning modalities is necessary.

The word modality usually refers to sensory systems. In the current microcomputer environment, multimodal information presentation means that both visual (via monitor) and audio information (via speakers) is provided. The division between audio and visual information is clear and easy to make. Especially in a multimedia environment, this division seems natural since the same classification refers directly to the two most common presentation devices, visual display units and sound devices (loudspeakers or headphones connected to a sound card). But when discussing human information processing, the distinction between perceptual systems is far too crude to provide a basis for analysis of the consequences of message format selection. Therefore, a much more fine-grained classification is necessary.

A typical further division made after the division into audio and visual information is to divide both audio and visual classes into verbal and non-verbal sub-classes (e.g., Mayes, 1992). Allan Paivio (1986, p. 57) uses this kind of classification, enhanced with haptic, taste, and smell sensory systems (table 1). Since taste and smell have neither verbal properties nor presentation devices in a contemporary multimedia workstation, they are skipped here. Neither is the haptic system taken into account, as the focus of the current work is in information presentation. Braden (1992) used the same simple classification when illustrating different possible ways to combine two output elements in an audio-visual presentation (figure 14).

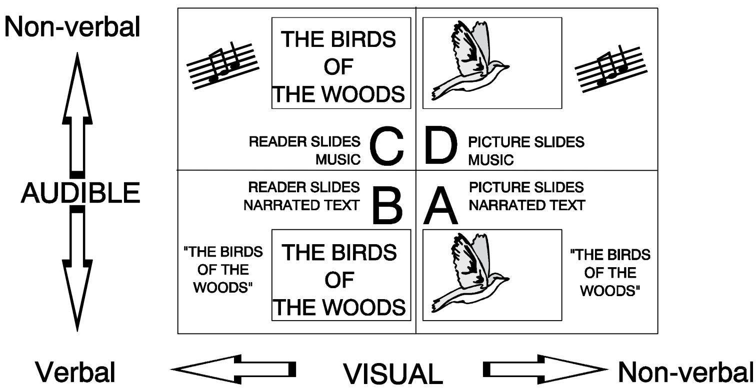

FIGURE 14 A 2x2 matrix concerning the classification of combined audio-visual messages. (Braden, 1992)

TABLE 1 The classification of Paivio (1986, p. 57) with examples. (Brackets and italics added by author.)

	Symbolic Systems
Sensorimotor	
Visual	
Auditory	
(Haptic	

	Symbolic Systems
Taste	
Smell	

Bernsen (1995) takes a long step forward in the classification of output modalities. In modality theory, Bernsen suggests a division of the concept of modality into sensory and representational modalities. By sensory modalities he means the division into modalities according to the sensory system applied, such as hearing and vision. By representational modality he refers to an elaborate classification of qualitatively different kinds of information forms. In his theory, different forms of output elements are classified according to whether they are linguistic, analogous with their referent, arbitrary, static or dynamic, and whether they are physically graphics, sound, or touchable information. The classification is coherent with the more common classification that only takes into account the sensory system and whether an information unit is verbal or not. However, modality theory goes much further, resulting in 28 different classes. For example, in Bernsen's classification, a case in which a text unit is in one location on a VDU (visual display unit) differs from a case in which the same text unit is in motion from one location to another because they differ with respect of their static/dynamic nature. Modality theory with its classification contributes the shift of focus from physical media to human information processing because each medium is used to present information in several representational modalities. For example, all ten graphical modalities are presented with a VDU.

When discussing information presentation to humans, it is obviously essential that the focus be on human mental processes instead of technical implementations. However, an astonishingly high number of research papers has been published concerning the effects of multimedia in, for example, learning.[1] This observation is made and strongly criticized by Clark and Craig (1992), who recommend a shift of attention in learning studies from technology used to instructional methods. For the same reason, in the present work different forms of the word "multimodal" are used instead of "multimedia" even when referring to applications that are run on multimedia workstations.[2] Thus, the conceptual basis of the present work is independent of the technology used. Additionally, since it is possible to present information via one sensory modality in several qualitatively different ways that require quite different kinds of cognitive processing, "modality" and related concepts refer here to representational modalities. Thus, multimodal presentation, according to the current definition, may be designed to apply to one single presentation device. For example, a combination of a narration and background music fulfills the criteria for multimodality.

1 A typical case is the approach of Najjar (1996). The conclusions about the "effects of multimedia" in learning are based on a comparison between a lecture and computer-based instruction.

2 The conception of multimedia and the related concepts is based on the broad reviews of Galbreath (1992) and Tolhurst (1995) that clearly suggest limiting the use of the word "multimedia" to a technological context. Dictionaries of the domain suggest a similar definition (e.g., Latham, 1995).

4.2 Dual coding theory

Allan Paivio's (1986, 1991) dual coding theory (DCT) has been used as a theoretical argument for using multiple modalities in information presentation. The inspiration for the development of the theory was the superiority of pictorial information in free recall experiments (Paivio, Rogers, & Smythe, 1968). The theory concerns two different kinds of mental processes, dedicated to handling verbal and non-verbal information. The verbal information units are called logogens and the non-verbal ones, imagens. The assumed benefit of combining verbal and non-verbal presentation is attributed to referential connections between logogens and imagens. The term referential connections means a process in which verbal representation of an object evokes corresponding non-verbal representation, or vice versa.[1]

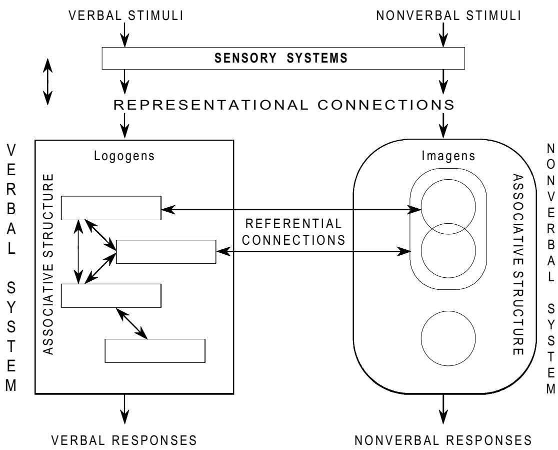

FIGURE 15 Information processing in two parallel sub-systems (Paivio, 1986, p. 67).

The empirical basis of the theory concerns memory, and the combination of, e.g., a picture (or some other non-verbal representation) and verbal representation is argued to result in better recall. Mayer and Sims (1994) developed dual coding theory further in their dual coding theory of multimedia learning. They stressed the significance of referential connections in the problem-solving process. Mayer and Sims can be said to have updated the theory in two respects: First, they shifted the focus from memory to problem-solving, thus expressing the recent paradigm shift of learning from positivism to constructivism. Second, they have chosen multimedia as the primary technological environment of their theory.

Dual coding theory has got much criticism. For example, Clark and Craig (1992) categorically reject the learning gains caused by using a combination of words and pictures in instruction. But it has to be remembered that the bulk of empirical research concerning dual coding theory concerns only recall.[2] Recall is an item that is easy to investigate empirically, but this kind of research provides quite a shallow view of mental processes.

Besides, there is, of course, much research concerning the combination of visual and verbal information that does not use DCT as a theoretical framework. Most of these studies measure recall and support the hypothesis

---

1 Anderson (1995) generalizes the classes by replacing the concept *verbal* with *serial-order*. He argued that more than just words can be processed serially (p. 128). Clark and Craig (1992) used the words *spatial* and *sequential*.

2 Most experiments that are interpreted to support dual coding have used verbal stimuli, but Hodes (1992) used a drawing test and still found support for Paivio's theory.

that relevant visual information facilitates the recall of verbal information (e.g. Findahl, 1971; Haring & Fry, 1979; Severin, 1967).

4.2.1 Dual coding theory and redundant messages

Even if the word redundancy is not usually used in the discussion related to the theory, the phenomenon itself is clearly present in both empirical and theoretical descriptions concerning dual coding. When the focus of interest is on referential connections between verbal and non-verbal sub-systems, it is obvious that, for example, picture-word combinations in experiments are meaningful, i.e., picture and text are closely related.

The foremost contribution of the dual coding theory to the argument for the use of redundancy is its empirical basis. Most evidence of dual coding and its benefits can also be seen as evidence for the benefits of presenting information in multiple formats, one of which is verbal. If the theory about referential connections between the relatively independent subsystems that result in significant benefits (as is argued by, e.g., Mayer, 1992) is accepted, the use of multiple modalities in information presentation is at least reasonable.

Additionally, dual coding theory provides a simple and applicable model of information processing that can be used as a conceptual framework when designing multimodal messages. Even if the model were at best only a fumbling illustration of reality, it would be useful in many decision-making cases in the design process.

In some respects, the applicability of dual coding theory should be questioned. First, the theory seems to suffer from extensive oversimplification concerning the stimulus objects of related studies; the research reports lack semantic analysis of the pictures used in experiments. Distinctions among different pictures are made in an unspecific manner, e.g., calling pictures *related* or *unrelated* to the text (or a single word). The subjective interpretation of a picture has not been paid attention to, i.e., the message has not been analyzed, or the analysis has not been reported. A picture contains so much information that its controlled use as a communication modality is problematic. However, the way Mayer and Sims (1994) used pictorial material (animation) reveals a reaction to this problem. The animated drawings in their experiments are made as simple as possible, thus minimizing the amount of information. And the smaller the amount of information in a picture is, the fewer the possibilities for unique, unexpected interpretations are. Another question is the appropriateness of this kind of approach, in which the task of the designer is seen as to control the cognitive processes of the user.

Third, as mentioned earlier, the bulk of the research that forms the experimental basis of the theory concern functions of memory. Free recall of single nouns is hardly an adequate expression of human cognitive processes. Even if Mayer and Sims have broadened the scope of the theory with their problem-solving experiments, a more diverse selection of experiments and approaches would be necessary.

4.3 Interacting cognitive subsystems (ICS) theory

ICS theory originally handled linguistic items in short-term memory (Barnard, 1985). Later, it was expanded to cover most cognitive, short-term activities that are relevant to multimodal human-computer interaction (Barnard & May, 1993a).

The model is constructed of highly independent cognitive subsystems, which all have their specific roles in information processing. Different subsystems function in parallel. The model thus rejects the idea of central processors and shared resources.

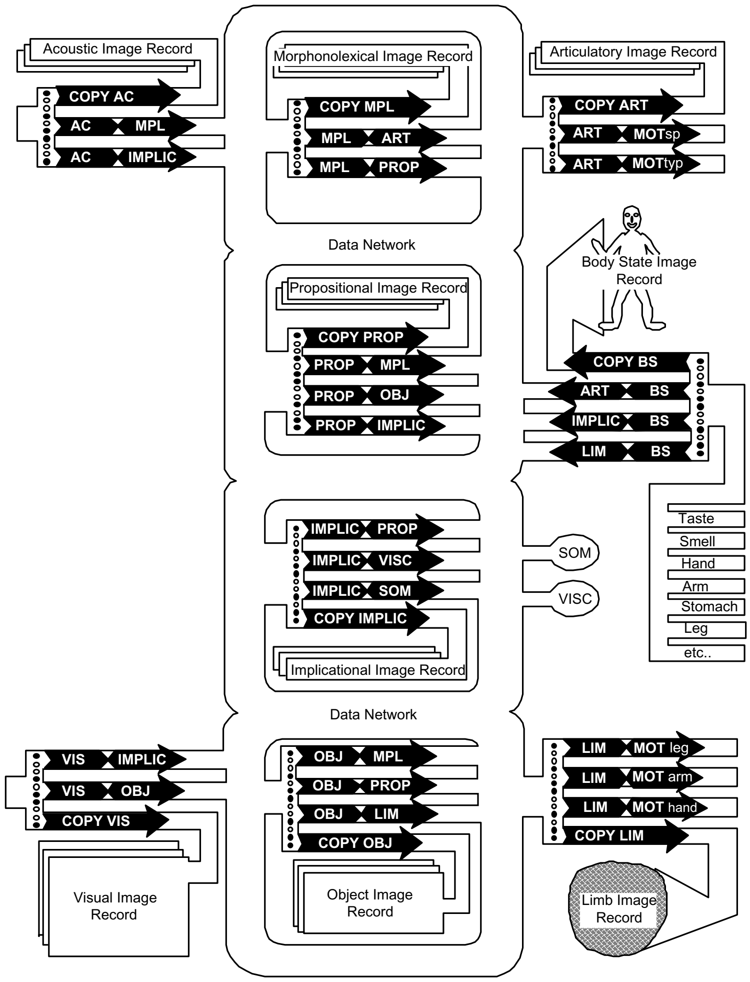

FIGURE 16 The overall architecture of interacting cognitive subsystems. (Barnard & May, 1993a)

In the model, subsystems are classified into the sensory and proprioceptive subsystems, the effector subsystems, and the central subsystems (in the middle column of figure 16).

All subsystems share some characteristics, as illustrated in figure 16. The dots in the input side of the subsystem refer to an input array. Copy processes store the input information in an image record. Other processes within a subsystem transform input information into a form that can be used as input in other subsystems. All the processes are parallel with each other, but each transformation process can handle one task at a time, i.e., transformation processes are internally sequential. Input into transformation processes may be either direct input from the subsystem or from the image record.

Different subsystems are specialized to handle different kinds of information. The sensory and proprioceptive subsystems (acoustic, visual and body state subsystems) provide the way for input to enter the system, while effector subsystems (articulatory and limb sub-systems) handle output information, as well as somatic and visceral response mechanisms (in the figure, SOM and VISC). Between input and output, there are four central subsystems: the morphonolexical subsystem handles linguistic and the object subsystem handles visual information, the propositional subsystem processes semantic information, and the implicational subsystem links information to a broader context.

4.3.1 The ICS-model and redundancy

The ICS-model is not an attempt to provide an overall description of mental processes. It is a functional model, constructed for the needs of human-computer interaction analysis. Its applicability is consciously limited by its creators, Phil Barnard and Jon May (e.g., 1993a), to short-term events. Therefore, there is no need to criticize its explanatory power concerning higher-order mental processes like problem solving. Instead, its structure makes it an interesting framework when considering the dynamics of integrating qualitatively different kinds of information.

If the basic assumptions of the ICS-model hold, the model provides a way to analyze the routes of different data streams within human cognition. As the model assumes that all other functions are parallel despite transformations, it can be interpreted that if two simultaneous messages are in an appropriate format so that the processing of them does not need the same transformations at the same time, they can be processed without interference. Hence, *if* the routes of data streams generated by each kind of representation could be reliably described on the basis of the ICS-framework, modality combinations that should be avoided could be easily detected. However, two facts make reality much more complicated: First, the complexity of cognitive processing ensures that cognitive models, however fine-grained, are always incomplete. Thus, the results of the analysis on the basis of ICS are only predictions that should be submitted to empirical findings.[1] Functional models like ICS are, of course, not even aimed at providing a complete description of cognitive processes. But another at least as difficult fact is that cognitive processes are unique, subjective processes. In the ICS-framework, subjectivity takes place above all in the output of the image records of each sub-system since the framework does not model the dynamics that generate the retrieval of information from an image record. The image record is, in ICS, only one possible source of information among others (like the sensory sub-systems); we can never know whether some sub-process is going to take information in the input information array directly from the input of the sub-system or from the image record of the sub-system. Thus, the models that are applied to predict information flow in cognition have a limit

---

1 Whether the plausible discrepancy between the model and observations should result in further developing of the model naturally depends on the degree of discrepancy and the consequences of further development to the clarity of the model. The problem is quite analogous with the one discussed earlier concerning the models of attention. The more a cognitive model is extended, the better might be its explanative power but the worse is its applicability. The balance between exhaustiveness and simplicity depends on the planned use of the model.

after which the model cannot be developed in more detail; a limit after which individual differences in cognitive processing prevent generalizations.

The most interesting feature of the ICS-model concerning redundancy is something that the authors call "blending of data streams" (Barnard and May, 1993b). The authors explain how distinct data streams can blend in different stages of processing. When two data streams enter the propositional sub-system and only one data stream results as output, the original streams have blended with each other, thus having only one common propositional meaning. This mechanism can be interpreted as a description of handling two redundant messages in human cognition. The kind of redundant phenomena that do not have an obvious similarity in meaning are blended at a more abstract level, in the implicational sub-system. Thus, the ICS model provides a way to classify qualitatively different kinds of redundant phenomena according to the level at which data streams blend with each other.

5 PILOT STUDY

To get information about the relevance of the proposed conceptual framework, a concise experiment was arranged. In the experiment, the proposed conceptual environment was exposed to use in user-interface analyses. The final experiment was preceded by a pilot study. The pilot study was carried out November 17-24, 1995.

A group of four undergraduate students from the Department of Computer Sciences was chosen as subjects for the pilot study. The students were busy with their field project, in which they planned a tutorial concerning basic skills in using a PC. The pilot study with the related lectures was designed to be one unit of their project; this experience was supposed to contribute to their skills in user-interface design.

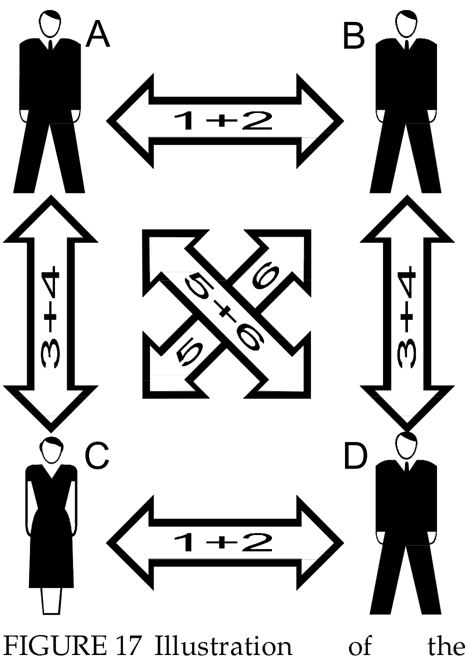

FIGURE 17 Illustration of the organization of the pilot study.

The process began with two sessions (lectures, 2 hours each, October 26, 1995), in which fundamental concepts and ideas of the proposed approach were discussed. In the actual test a pair of students analyzed two given nodes of hypermedia products. The products and nodes to be analyzed were chosen in cooperation with the group. At the same time (i.e. while choosing) explicit navigation instructions to each node were worked out and written down. As the students participated in the writing of the instructions, the clarity of them was guaranteed.

Six nodes were chosen. The students were paired in six different ways: each student worked with each other student for one session. Each student had three other students to work with, and each pair analyzed two nodes. This way each single student analyzed all six nodes (see figure 17).

The hardware configuration was as follows: PC CPU (486/66MHz) with audio card, speakers and CD-ROM-drive, monitor (21"), keyboard, mouse, video converter, video tape recorder and a microphone. Additionally, the microphone signal was amplified with an audio tape recorder's microphone pre-amplifier. The VTR was also connected to a television for monitoring (figure 18). With this equipment it was possible to save on videotape the events on the computer screen synchronized with the audio information (the conversations of the students).[1]

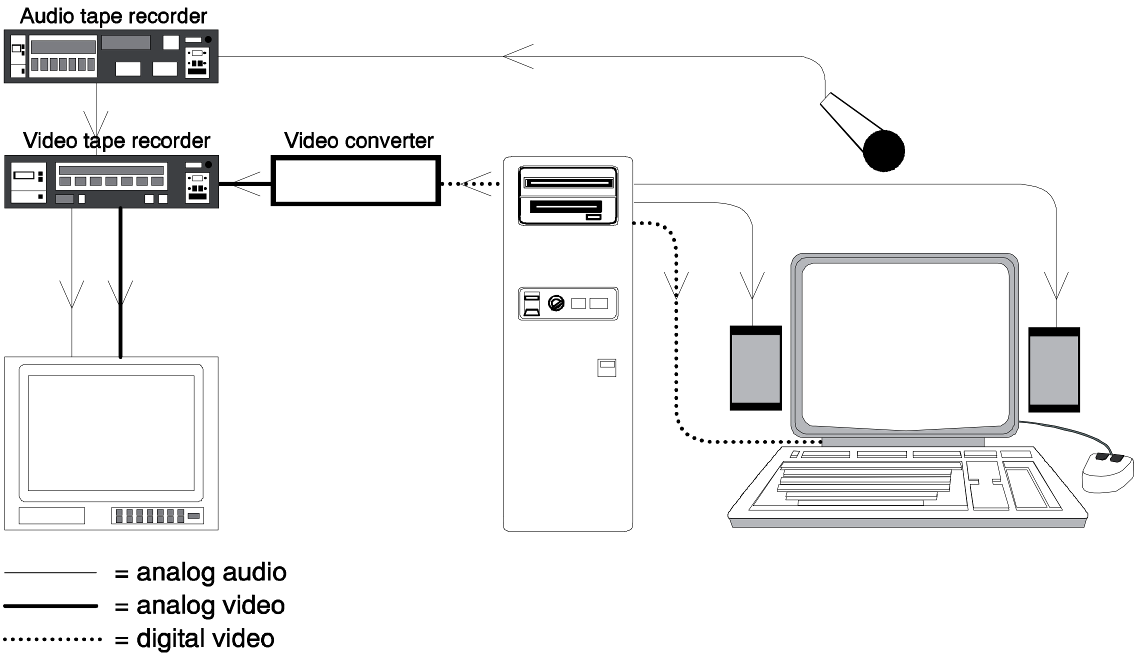

FIGURE 18 The hardware configuration.

The analyzed hypermedia products differed from each other in many ways. Below is a brief description of the products and the contents of the selected nodes.

1. A prototype of an introduction to the life and work of Finnish architect Alvar Aalto. This software is produced in the University of Jyväskylä by students who study multimedia. The analyzed node contained information about a chair designed by Aalto.
2. A famous American encyclopedia, MS Encarta. The selected entry was about the Ku Klux Klan.
3. An American multimedia product about fish. The focus was on crocodile fish.
4. A PC-tutorial. The selected node was about CD-ROM.

---

1 In addition, audio information of the application was recorded by mixing the signal from the output of the audio card with the microphone signal. This detail is not illustrated in the figure since the arrangement proved impractical concerning the balance of the two sources and was not implemented in the final experiment.

5. A multimedia-product about the anatomy of man. The liver was the focus of this node.
6. An introduction to the production of one small publisher in Jyväskylä (Teknolit). Information about a book concerning MS Access was selected.

At the beginning of the first two sessions the researcher was present to make sure that the task was properly understood and there were no technical problems.

The students had a form to fill in concerning each node (Appendix 1). The form began with a table in which they had to list and analyze each element of a node. The analysis was performed according to the Taxonomy of Modalities (Bernsen, 1995). Students had become familiar with the taxonomy in the preceding lectures.

The second part of the form was a table of connections among the elements of the node. In addition, the students were to interpret the redundant messages they found.

5.1 Evaluation of the pilot study

The role of the pilot study was to test and evaluate the technical arrangements, the relevance of the aspects of the analysis and the adequacy of the instructions. Therefore, the criteria for evaluation were based on these aspects.

5.1.1 Technical arrangements

The data was easily accessed in the videotape. After the first session it was noticed that the level of audio input was too low. That is why the microphone pre-amplifier was added to the configuration later (before the second session) and was connected directly to the audio input of the VTR. Thus, the audio information of the first session was barely sufficient.

The situation did not seem too artificial: the subjects' discussion sounded natural, lively and uninhibited. The only aspect that reminded the subjects of the experimental nature of the session was that the intuitive way to point at an object on the screen was a forefinger, but in order to indicate the object of discussion in the video tape, the subjects were asked to point to the objects with a mouse pointer.

5.1.2 Relevance of the aspects of the analysis

The first task of each session was to give a list of elements and to analyze them according to the Taxonomy of Modalities. The purpose of this analysis was to find out how obvious this kind of classification is and to provide a tool to describe an interface element briefly in a relevant way. This task was so difficult that subjects spent about half of the total session time on it. Additionally, this task, which was meant to be only a quick routine, was at the beginning of the form, which made it appear to be of primary importance. When the subjects were to begin the essential part of the task they were too tired to concentrate on difficult interpretations. In addition, the use of the Taxonomy of Modalities proved too difficult on the basis of the given instruction. The classification failed and did not contribute to the further analysis.

In the most important part of the task subjects were to interpret the relationships between the interface elements. They filled in a matrix in which they had to interpret the nature of the relationship between each possible pair of elements. They had five possible ways to label the nature of each relationship: distinction (D), complementing (C), substitution (S), emphasizing (E) and redundancy (R).[1]

It was soon recognized that it was hard, actually frustrating, to find any sense in interpreting relationships between elements that were not meant to have anything to do with each other. And since at least half of the relationships were of this kind, the whole task seemed somewhat embarrassing for the students. The result was that most relationships were interpreted as *distinction* or *complementing*. Interesting and meaningful relationships were lost among these obvious cases.

5.1.3 Adequacy of the instructions

It seemed at first that the introduction was understood and the work went on smoothly. Later it was found out that the subjects either trusted too much in their ability to cope with the task without the literal instructions they had or they were not motivated enough to concentrate on the task. There was also a delay of three weeks between the instructions (lectures) and the actual study. The delay may have caused too much confidence: the subjects remembered that they *understood* the point but in the test situation many things occurred that they had already forgotten. A brief repetition just before the analyzing session would have been of advantage.

One basic concept concerning the analysis was clearly not been understood by at least two of the students: a *message*. They did not use the word, and their interpretations of the relationships were not based on the messages of the elements but on some property they could not describe or name. This was easy to understand as a consequence of their background; interpreting the messages of – for instance – a photograph requires much more experience in semantic analyses than these students probably had.

---

[1] Note that the underlying model was in a very early stage. Even the form of the concepts was later modified.

5.2 Conclusions of the pilot study evaluation

5.2.1 Revisions to the technical arrangements

The hardware configuration was relevant to get the information needed. It would have been possible to get video information about the behaviour of the subjects by attaching a video camera to the configuration and saving its video signal in the same videotape in a PIP-format (picture in picture). But this kind of arrangement was not found worthwhile because the focus of interest was on the conversations, and actions related to the screen were performed with the help of a mouse (e.g., pointing, clicking). Enough information about the conversations was recorded via the audio channel, and the actions performed with a mouse were observable in a video recording of screen events. However, a PIP-window possibly would have covered important visual-only details such as body language or facial expressions.

The newest available software to encode video information was investigated to find out more sophisticated ways to treat the data. However, the investigated tool was not found appropriate for this purpose.

5.2.2 Revisions to the aspects of the analysis

For the needs of the final study the form had to be reworked completely to lead the subjects to pay most attention to the essential aspects of the task. Instead of mechanical, complete coverage and listing, the subjects should be encouraged to think and interpret, and do it aloud. Instead of analyzing separate details on their own, details always should be seen as a part of the whole. Therefore, in the instruction lecture as well as on the form, words like "essential" and "important" were found necessary for describing the level of completeness of the analysis.

It was found that in the interpretation task the selected elements usually should have a similar role in the interface. To be able to find suitable pairs of elements it is necessary to classify the elements first on some basis. For this purpose, the following roles of elements were found:

- Decorative; the element's major role is to make the interface look or sound pleasant. Background music or pictures, for example, might have been added to an application with no other purpose than to fill up a flat emptiness.
- Outlining; the role is to make the interface clear and illustrate the role of the other elements. For example, lines and frames that mark off a group of elements are outlining.
- Informative; the purpose of this kind of element is to provide some matter-of-fact information.
- Controlling; the purpose of this kind of element is to provide a way to control either the navigation in a product or the studying of one node of a product. Push buttons with arrows referring to sequential browsing are one of the most widely used elements of this kind.

One element may, of course, have more than one role. When a node has some kind of arrows on screen for navigation, their role is obviously controlling. But the more attention is paid to make the arrows look good, the more decorative role they have in addition.

5.2.3 Revisions to the instructions

The content of the instruction was relevant, but the way information was introduced was probably not the best possible. The level of abstraction was too high for the subjects to form intrinsic knowledge by hearing the oral instruction. It became clear that the subjects should have some practice in which they have to use the concepts they are supposed to understand. The implementation was an exercise in which subjects were to plan pairs of interface elements of the required quality.

The instruction should also have contained principles and examples about interpretation of messages of different formats. In the pilot study, the subjects did not quite understand the major role of verbal interpretations. If we are trying to compare two messages, of which one is verbal and the other is visual, we have to interpret the messages first in the same format. Usually, verbal interpretation is the only one that makes any sense. Of course, to demand a complete verbal interpretation of e.g. a photograph is absurd. But language is still the best way to express meanings comprehensibly.[1]

[1] The first version of this chapter was written when analyzing the data of the pilot study. Therefore, the ideas presented especially in this paragraph are quite contradictory with the ideas that arose later and are reported elsewhere in the present work. The most salient difference is the emphasis on the value of verbal communication in the older texts. However, this version is presented here with only minor changes compared with the first version in order to illustrate the evolution of the researcher's conceptions concerning the domain.

6 FINAL EXPERIMENT

The experiment was arranged utilizing the experiences from the pilot study. Major changes were made in the contents of the instruction and the analysis. As explained earlier, the role of the final experiment changed markedly when the data of it were analyzed. However, the original idea was first to teach the subjects the guidelines of the proposed framework and then to test how they could use the framework in the analysis of multimedia products.

The experiment was conducted May 22-24, 1996 at the University of Jyväskylä. That spring, the Continuing Education Centre of the university arranged a course concerning production of multimedia learning material. The participants were unemployed, and most of them had an academic background. The total number of students was 14, consisting of 9 female and 5 male participants.

On Wednesday morning, the subjects were given a lecture concerning the current approach. The lecture contained an orientation to cognitive models and their contribution to design as well as the central concepts with examples of each. Additionally, the students did a short group exercise after the lecture. The exercise required the subjects to design redundant combinations of output elements. It was implemented with paper and pen, and the results were presented and discussed in a separate session a few days later.

From Wednesday afternoon to Friday afternoon the participants analyzed four nodes of certain multimedia products in pairs. Each pair had two and a half hours to complete the task. The hardware was an ordinary multimedia PC with 17" display and 16-bit soundcard. The operating system was MS Windows 3.11. The hardware for data collection was similar to that used in the pilot study, with one exception: In the pilot study, sounds of the application were caught directly from the sound device and mixed with the signal from the microphone. Now, all the audio information was recorded through one microphone; the sound from the speakers and the conversations of the students.

This way, the quality of recorded conversations was guaranteed at the expense of the quality of the sound of the application. But, concerning the analysis of the data, the quality of the sound of the application was not important. The audio elements of the analyzed nodes were familiar for the researcher, and therefore they were easily recognizable even in a low-quality recording.

Each analysis session began with the opening of the first application and the given node. Concerning each application, explicit instructions for opening and navigation were given in written form. The instructions proved to be adequate, and no severe problems emerged in the opening routines. After finding the appropriate node, the subjects began their analysis. The subjects were aware that the main thing they had to do was to discuss. They were supposed to fill in a form that is presented in Appendix 2. However, the function of the form, rather than to provide a document about the results of the analysis, was to guide the analysis process and to make the subjects pay attention to appropriate aspects.

The first task was to list the elements of the node. It was up to the subjects to decide what they regarded as elements.[1] Different elements with a similar nature were encouraged to be handled as a group instead of listing each separately. For example, a visually grouped set of navigation push buttons is reasonable to handle as a group thus saving a monotonous repetition of identical analyses. Additionally, this kind of grouping shortens the list.

After naming an element, the subjects had to classify it according to its role in the node. The alternative classes were I (Informative), C (Controlling), De (Decorative), and Di (Distinguishing). In the instruction lecture the labels were clarified as explained in the discussion concerning the pilot study. It was also stressed that hardly any element has only one of these roles. The dominant role was asked for, but an element could be assigned to more than one category. The purpose of the classification was to stress the consequences of different roles of the elements in further analysis. In the pilot study, it proved quite frustrating to contemplate the relationships of elements that obviously did not have anything to do with each other. Now, those cases were rejected out of analysis by encouraging the subjects to limit further analysis (the second and the third tasks) about the relationships within each class.

The second task was to list some meaningful combinations of elements (see Appendix 2). The combinations were supposed to be labeled as distinct, complementing, substituting, emphasizing, or redundant. Finally, in the third stage, each redundant case was to be verbalized by describing each redundant message. The subjects were trained to perform the third task, which was the most difficult, with the help of examples in the instruction lecture.

When preparing the experiment, the researcher's conceptions about the phenomena in question still emphasized the role of verbal communication. Therefore, the examples given in the instruction lecture were quite mechanical and simple cases in which two messages were expressed verbally and then compared.

---

[1] The analysis of the concept of output element in chapter 6.5 is based on these decisions.

6.1 Post-test

Three months after the experiment, a simple post-test was arranged. In it, the subjects were asked to write a short description about the functions of an airship. This task was related to the core of the content of the first analyzed node in the experiment. The reason for arranging the post-test was an interest in the correlation between the observed deepness of the analysis and the understanding of the contents. It was assumed that a deep analysis results in understanding, i.e., understanding inevitably indicates deep analysis. Therefore, a positive correlation would have shown that the observed deepness of the analysis was real, thus showing the reliability of the observations. The results are reported and discussed in parallel with the results of the actual experiment.

6.2 The objects of the analysis: a description of the nodes

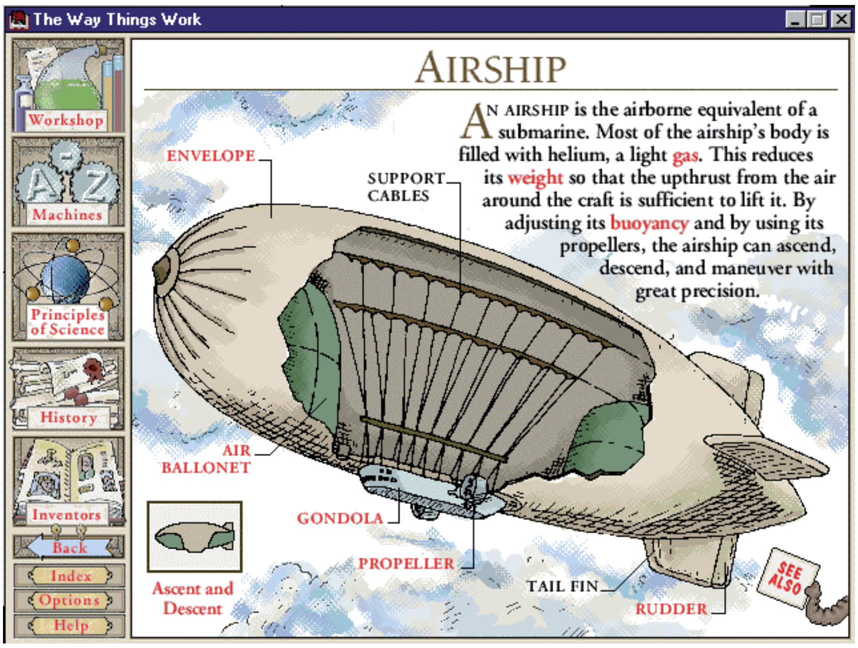

FIGURE 19 The screen design of "Airship".

Node 1. The first node was about airships; it was in a multimedia product "The Way Things Work." The focus of the node was in accordance with the name of the product: the structure and basic functions of an airship. The basic concepts were explained in written text. Different parts of an airship were labeled in a drawing, which contained information about the appearance of an airship and its inner parts. Some of the names of the parts were hot-words containing links to a pop-up text window with further information. The main text also had some links to text windows. But those windows did not have further links, so it was impossible to get lost from the airship node by following those links. The structure is illustrated in figure 20. The large circles illustrate the nodes, and the small ones refer to pop-up text windows that do not have any further links. The figure is naturally strongly simplified.

As can be seen from figure 20, the node with its dedicated links is very likely to form a mental representation of an independent whole.

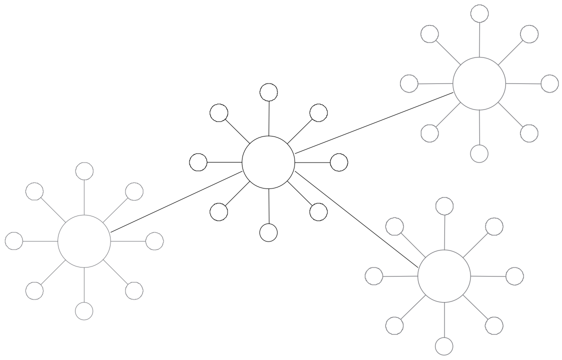

FIGURE 20 ”Airship” as a part of the structure.

Node 2. The second node concerned signal transfer from the ear’s sensory cells to the brain. It was a part of a Finnish multimedia product “Virtual Ear,” concerning our environment of sounds. The node to be investigated was the last part of a sequence of nodes that illustrated the function of the ear (“Journey to the ear”). Because this node was the last one of a sequence, it was necessary to follow the whole sequence to be able to understand the content. Therefore the subjects of the experiment were told to follow the sequence before beginning the actual analysis.

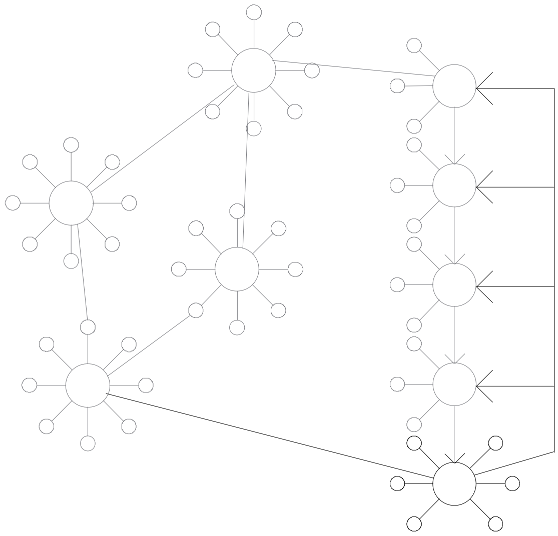

FIGURE 21 ”Sensory cells->brain” as a part of the structure.

The structure was not as clear as in the first case. The difficulty was that some nodes had a clear hierarchy while some others did not. The structure is illustrated in figure 21. The “Sensory cells->brain” node is drawn with a solid line. The sequence (“Journey to the ear”) is on the left-hand side of the picture, and was expected to be followed downwards. The solid arrows indicate the possibility of linking backwards through the sequence by choosing the name of the desired node from a menu. The small circles illustrate dedicated functions like external pop-up windows with text.

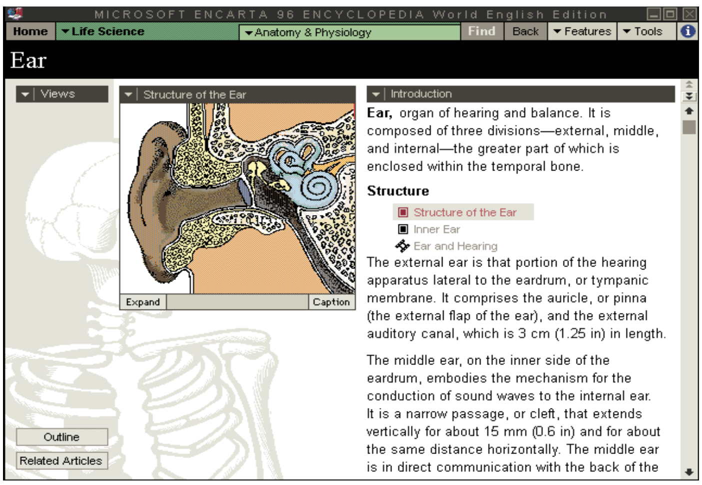

FIGURE 22 The screen design of “Ear”.

Node 3. The third node was from the MSEncarta ’96 multimedia encyclopedia. It was a survey of the ear. The structure of this encyclopedia was much more “hyper-like” than the former multimedia products. Despite the complicated structure of the product, the node “Ear” was easy to distinguish as one whole. In this product, only dynamic information (e.g. animation, video-clips, and sounds) was stored behind dedicated links. Most of the static material (written text, pictures) concerning one node, was presented on the same screen.

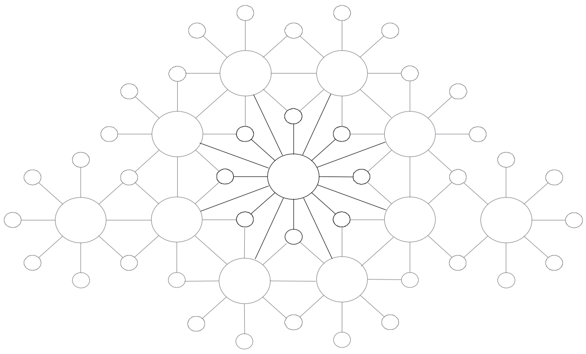

FIGURE 23 ”Ear” as a part of the structure.

To be able to put almost all the elements on the same screen, the designers have been forced to use scrolling properties. At least, using scrollable elements instead of pop-up windows makes it possible to make the structure clear. As illustrated in figure 23, most links provide a way out of the node. The figure, even if it simplifies the reality, shows the “hyper-character” of the product.

Node 4. The fourth node was from the multimedia encyclopedia MSMusical Instruments and reviewed the trumpet. The structure resembled that of the first product (The Way Things Work), but was somewhat simpler and clearer (figure 20 serves as an illustration of the structure). The layout of the main screen was extremely uncomplicated and carefully designed. A high-quality picture of a trumpet was spectacular. A white background highlighted the pure design of the graphics.

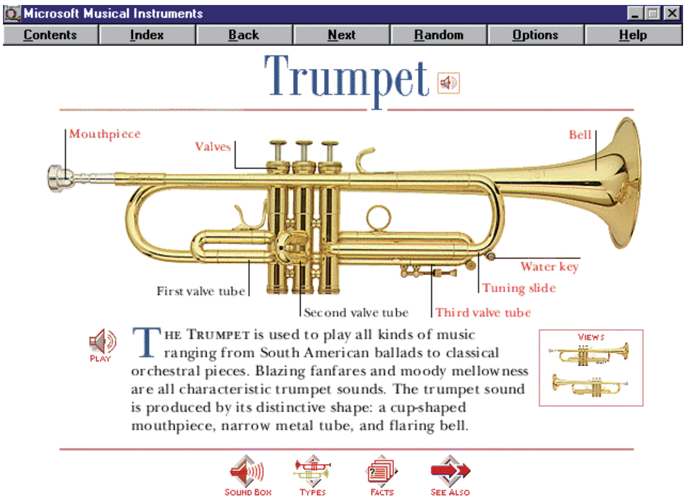

FIGURE 24 The screen design of ”Trumpet”.

6.3 Introduction to the descriptions of the analysis sessions

The following descriptions are based on the video recordings of the analysis sessions and the forms that the subjects had filled in. The tapes were first reviewed a couple of times. While reviewing the tapes for the first time, the researcher made his own remarks on the forms and gradually formed his idea about the contribution of the data to the domain of the current study. After that, the tapes were analyzed step by step, writing the following descriptions.

As mentioned earlier, the role of the data changed a great deal during the analysis of the data. The original idea was to test the comprehension of the subjects concerning the proposed model of combinations of output elements. However, the analysis of the data led to significant further development of the model.

The descriptions are arranged first according to the pair of subjects and then by node, i.e., the analyses of each pair are handled in sequence. The number codes of the pairs are the ones that were used in the organization of the experiment, but they are listed here in the temporal sequence in which the pairs conducted their analyses. The descriptions are headed with the pair code and the topic of the analyzed node.

As can be seen from the descriptions of the work done by the first pair (pair no. 2), only the analyses of the two first nodes were taken into account from them, because the recording of the third node failed, and the fourth product did not work properly and was changed for the other pairs. The fourth node was originally a Finnish multimedia product concerning the functions of a nuclear power station, but it was replaced by MS Musical Instruments.

The descriptions are the lowest level analysis of the data. Their nature and role is so different from the rest of the current work that they could have been attached to the work as an appendix. However, since the evolution of conceptions of the subjects is an essential phenomenon in the current work, the descriptions have been put here. It is possible to use this section as an appendix by skipping it, going straight to the summary of the analysis, and returning to the descriptions when they are referred to. The descriptions are bordered with gray vertical bars in order to illustrate their difference from the rest of the work.

6.4 Descriptions of the analysis sessions

Pair no.	2
Node	Airship (description on p. 77)

The session started with a rapid survey. The subjects were first concerned only about the screen elements. No attention at all was paid to the contents at this stage. The subjects were eager to begin with the formal analysis (filling in the form).

A remarkable finding was that this pair did not recognize the role of the node as part of a whole hyper-product. They did not find the role of hot-words and -spots to be controlling at all. At one point the first subject asked whether links are controlling, but the second subject rejected the suggestion and the first dropped the matter. Almost all the elements were considered to be in the first place informative.

Most relationships between the elements were found to be complementing. The subjects tried to find redundant relationships and actually found one good example, the animation about the ascending of an airship and the related sound. However, the subjects were too busy to write it down because they had already completed the listing of elements, and additional elements on the list would have expanded the number of different relationships, which, in turn, should have been analyzed. It is likely that they did not find this detail important enough as a part of the whole. This can be interpreted as a lack of vision about the aims of the node (or the whole product). How else can we understand that in a node that tries to explain and illustrate the function of an airship, the method of increasing and decreasing the weight of the vehicle is not found to be important enough to be listed? In the post-test, the subject who dominated the discussion could not explain the principles of the function of an airship, while the more restrained subject could.

These subjects found hot-words to be separate elements. Hot-words were red in color, which led the subjects to interpret that hot-words emphasized the details they were referring to. The only redundant phenomenon this pair noted was that there were alternate tools for navigation.

Pair no.	2
Node	Sensory cells→brain (description on p. 78)

The subjects failed to record the session on videotape, so all data available is on the filled-in form. Unfortunately, not very many perceptions were registered and the stored data does not contain new information about the way this pair interpreted user-interfaces. When compared with the previous node, one similarity was clear: this pair considered again all (despite one) elements as informative. This is an important observation when analyzing their interpretations of the relationships; if the roles of all the elements are seen as informative, it is impossible to make any deep analysis in multiple levels concerning meanings.

Again, only one redundant message was registered, and that one was quite obvious: the node contained the same text as narration and displayed text. Those elements were presented as alternatives; they could not be presented simultaneously.

Pair no.	2
Node	Ear (MSEncarta) (description on p. 78)

It became clear from the beginning that the user-interface of MSWindows'95 was not familiar to the subjects. Thus, some symbols (minimize, maximize, restore, and close, in the upper right-hand corner) were arbitrary in nature for the subjects.

The focus of the discussion was again on the layout of the user-interface. The subjects did not try to find out information about the ear. All the menus were clustered on the form as one unit without any deep investigation of the contents or roles of different menus. At first sight they looked (and sounded) alike to the subjects, and that was a sufficient basis for handling them as a whole.

The pair continued in the same manner as in the previous nodes; a quick and shallow overview and some obvious relationships but no deep analysis about the meanings of the messages. The informative role of the elements was considered primary; only arrows and some push buttons were seen to have a controlling role. Redundancy was considered, as in the first two nodes, synonymous to alternative (some navigation functions could be started either from a menu or from a hot-word).

This first pair (no. 2) was in a different position than the others: as mentioned, the whole session was not on tape. Additionally, the fourth product did not work properly and was rejected from this experiment and replaced with another one (MS Musical Instruments).

Pair no.	3
Node	Airship (description on p. 77)

As in all sessions, the researcher was present in the beginning to make sure that there were no technical problems. He suggested before the subjects started their task that an overview of the contents could help to figure out the whole. He even showed where they could find the essence of the contents (the hot spot which started a brilliant animation about the principle of ascending and descending of an airship). As a result, these subjects started to read the text field, but, having read a couple of sentences, began to judge it and most probably never read the text to the end. They also took a brief look at the animation, but the underlying attitude toward the task clearly was to evaluate the user-interface, and that is why they did not bother to pay attention to the contents – in other words, they did not understand how an airship gets higher and lower. Instead, they began to mechanically list the elements on the form almost immediately. Later, despite this lack of interest in the contents, the subjects were forced to understand the principle better when they translated (vocally) the texts of pop-up windows containing information about the basic concepts concerning an airship.

When listing the first element this pair defined in an implicit way the criteria of classification of the roles of the elements. The subjects discussed the push-buttons on the left hand side: "Is this informative or the window that emerges when I click this." "Their main purpose is to enable navigation." "They are quite decorative as well." After these comments they classified the element as controlling and put a check in brackets in the "decorative"-column also. Thus, the subjects seemed to have understood the multiple roles of elements and were able to evaluate the primary role. The listing of elements continued in the same way: first thinking about all the possible roles and then which of the roles dominated. This was a good basis for further analysis.

When this pair began to analyze the relationships between the elements, they started with the text field and the picture of an airship. As the text did not contain information about the appearance of an airship, the subjects classified the relationship as complementing. One subject put it in a pertinent way: "The text is in a quite different level than the picture." Then they discussed the relationship between the written name and the picture of support cables of an airship. After a few seconds silence, the first comment of the other subject was that the relationship was redundant. Later they found that the picture contained much more information: "For example, you cannot draw conclusions about the way the cables are tied up here from the text." According to the given instruction, the proper interpretation for the latter should have been emphasizing, of course. But the interesting point is why the first suggestion was to classify this relationship as redundant. Mechanically interpreted, we could argue that the text contained the name of the object and the picture contained information about the appearance. The work of this pair was so accurate that there is no reason to fault them for classifying this relationship as redundant too easily. They surely felt the relationship was somehow redundant. The picture must have given them more information than what they actually saw. They understood – even without the text – that the lines in the picture presented some kind of support cables. This can be seen on the form also, when they were to describe the redundant message. They wrote: "The picture contains the cables."

A similar phenomenon emerged even more clearly after this, when the subjects disagreed about some relationships. They first compared the level of redundancy in the relationship between the name and the picture of support cables and between the name and the picture of the tail fin – and disagreed. Each of them had a strong opinion, which indicates a clear intrinsic idea about redundancy. The disagreement resulted from different interpretations of the picture. The subjects seemed to understand this themselves when discussing the next relationships, namely between a picture and a textual explanation about the gondola, and thereafter, about the propeller. The male subject knew much more about airships, and that is why he had more tools to interpret the pictures. For example, when he saw a picture of the gondola he saw a cabin for people while the female subject did not even try to understand what this box was. She said "I did not even think that there are passengers." So, the textual explanation about a gondola was complementing for her since it contained mainly redundant information for the male subject. They formulated the redundant message like this: "The text contains the information given by the picture." (They used the same description to refer to both picture-text relationship cases: gondola and propeller).

Only the male subject returned the post-test paper. In it, he showed that he understood the principles concerning the contents correctly.

Pair no.	3
Node	Sensory cells→brain (description on p. 78)

In the literal instruction the subjects were asked to follow through the sequence of nodes concerning the structure and function of the ear and to try to understand the information. The actual analysis concerned the last node of the sequence, named "sensory cells → brain." Despite the advice, this pair was too impatient to follow the sequence. Due to this, the subjects missed vital information like basic concepts about the contents. Their ability to interpret the contents of interface objects was therefore restricted. For the same reason, they get lost in the structure and could not distinguish the node on which they were supposed to be focused from the rest of the product.

The most interesting finding concerning the relationships between the elements could be seen in the discussion about the animation and narration about the electrical functions of the sensory system of an ear. The subjects had classified the animation mainly as a decorative element and clearly found it difficult to argue its relationship with an informative element (the narration) as redundant. This indicates the informative bias in the assumed purpose of the product. It is much easier to verbalize or describe a meaning of an element that only concerns facts about an object. And the comparison of two elements requires the interpretations of meanings in the same modality (here: verbal written.)

The subjects found one detail in which the product's authors used redundancy in an effective way: an animation of a push-button going down when clicked and the simultaneous sound, both indicating a successful clicking. Although the role of this detail can be argued to be peripheral, the subjects clearly proved to have adopted the concept of redundancy.

There was some disagreement about the possible redundancy between certain control elements, namely back and forward arrows and the list of topics. The disagreement was probably due to partial misunderstanding about the function of the arrows, not the ability to interpret the messages of the elements in a proper manner.

Pair no.	3
Node	Ear (MSEncarta) (description on p. 78)

The first reactions of the subjects were irritation because of the abundance of details of the screen. The subjects criticized the organization of the information offered and did not (at least immediately) understand the icons, the messages of different colors used and many other illustrative details. The lack of previous experience with this product obviously caused most of the criticism, as the user-interface of MSEncarta is, anyway, internally consistent (this was even mentioned by the other subject). But the anxiousness of the subjects shows again how difficult it is to design an interface that appeals to our "natural" way of behavior, however fluid the graphics and sounds may be. Usually it is practice that makes these elements "natural" for us.

Feel of control – or the lack of it – was a salient concern. Unexpected, major consequences of an input act seemed frightening: "When you click one button, everything changes."

When trying to understand the relationship between a list of topics (hot-words) and the main text, the subjects found an analogy between this and the relationship between a table of contents of a book and the text of a book. This metaphor provided the subjects with a common mental representation about the structure of the node and proved a fruitful basis for the discussion about the relationships, although some disagreement about them remained after all.

The subjects proved again to have formed a relevant idea about redundancy: they felt that when the basic functions of the ear were presented in narration, text and animation, all of these elements have redundant relationships with each other, which is certainly the case. However, for one reason or another, the subjects did not find the implementation successful. Redundant messages were not found necessary at all. So, the possibility of using redundancy in an effective way remained unclear for them. These subjects thought the use of multiple media to be due to the need to show the technical competence of the authors.

As with the previous node, the sound related to the clicking of a push-button was also classified as redundant with the graphical counterpart.

Pair no.	3
Node	Trumpet (description on p. 79)

The clarity of the screen immediately created a favorable impression on the subjects. Perhaps the contrast between this node and the preceding one contributed to this response. The number of elements was minimal. However, the purpose of most of the dynamic elements proved unclear for the subjects until tried. This was not surprising, as the authors had used the same icon in order to activate information units of different modalities; clicking on an icon representing a loudspeaker meant in one place a vocal presentation of the word "trumpet" and in another place the same icon meant a sample of trumpet sound.

The subjects followed some links, and the more they investigated, the stronger became their positive attitude toward the product. Although neither of the subjects was, according to their own words, particularly musical, they concentrated on the contents much better than in the preceding task (ear). *It could be interpreted, that when the overall appearance was not frightening and the amount of information available on one screen was not huge, the subjects felt they controlled the functions of the machine and dared to grub in the world of the trumpet.*

The redundant phenomena that this pair of subjects found were obvious and thus did not evoke deep verbal analysis or discussion. Thus, further conclusions about the adopted concept of redundancy could not be drawn. However, some simple comments revealed that the idea of a redundant relationship was at least not naive like defining that the "same thing" is presented in the text and vocally. Even when as tiny a detail as the topic of the node, "trumpet", was presented in both of these ways and the relationship was listed as redundant, the other subject mentioned that the information concerning pronunciation is in the vocal element only. It is to be admitted that that comment could also have been due to the fact that this pair found the purpose of the vocal presentation as an instruction of the pronunciation, not as an alternative heading.

Pair no.	1
Node	Airship (description on p. 77)

The subjects seemed to have taken the task mainly as a mechanical classification. As soon as the appropriate node was open, they began to list the elements and classify them according to the given frame; no attention was paid to the contents. Even the push buttons of the node were rapidly classified into three groups according to their appearance only.

The atmosphere of the conversation reflected some kind of presumptuousness toward the product. Presumably the funny style of the product created a childish impression. This is a good example of what fatal consequences even a choice that seems to have not too much to do with the content, like the overall image or style of a product, might have.

This pair did not bother to investigate the node thoroughly enough to discover the most important section of the node illustrating the function of an airship, the animation. Thus, it was not a surprise that neither of the subjects had an idea about the item in the post-test. An excuse for this could be that the icons and texts of two hot spots that led to the animation did not illustrate too clearly what was behind them.

When the subjects began to think about the relationships between the elements, they first discussed the concept of redundancy, probably to ensure that their views about the concept were coherent. In that discussion they gradually drew nearer and nearer to the definition that was presented to them in the instruction lectures. "(1) What are the things that emerge twice... (2) that are overlapping... (3) what is the thing that is attempted to be presented in two ways?" From the sequence of these three comments of the subjects the whole concept is defined in three stages by raising the level of abstraction in every step. The first one can be associated with the definition used in information theory: the same information is repeated. This is the kind of definition of redundancy that most likely means something negative, something we should get rid of. The second definition takes into account the possibility of having two information units, the meanings of which are not identical but have *something* in common. Whether that is good or bad is not so clear. The third definition refers to fully intentional design action, in which the possibility of using different mental procedures in information processing is introduced.

A deep mental representation about the concept of redundancy emerged when one subject analyzed the relationship between an icon and a text in controlling push buttons. He said that the relationship is not redundant because the purpose of the button is not clear without the text. But he continued that if the icons remain the same throughout the product, the icons become familiar and the relationship between the icon and the clarifying text can be redundant. In this implicit way, he proved to have understood that as the construction of meanings on the basis of given information is a subjective process, so is the relationship between two information units individualistic. In other words, he seemed to have understood that the task is not to find out the relationship between the presentations of two information units but between the messages that are subjectively constructed.

The other subject stated that there is some redundancy in the picture of an airship and the text that contained information about the physical principles regarding the function of an airship. Unfortunately, she either was not capable of verbally arguing or did not find arguing essential, so it is difficult to interpret the sense of her choice.

The work of this pair shows that even if the concept of redundancy is theoretically learned, it cannot be applied without profound understanding of the contents. Meanings must first be constructed to be able to compare them with each other.

Pair no.	1
Node	Sensory cells→brain (description on p. 78)

These subjects followed the sequence of nodes called "Journey to the ear," as instructed. They concentrated hard on the contents, at least judging by the audio information of the tape; during the sequence either there was a complete silence or brief comments about the contents or the implementation were heard.

The discussion about the relationships between the elements began with a search for redundant messages. Again, a conceptual analysis of redundancy took place: The subjects investigated the relationship between a written text and a narration of the same text. The same subject, who analyzed in the previous task the proportion of the relationships between a text and an icon in a brilliant way, immediately pointed out the criteria for calling the current relationship redundant: "The tone of the voice did not add anything new." Later, he also mentioned that the style of the printed text could have had information that is not included in the narration, but that this typographical means was not used.[1] On the other hand, despite this obvious redundancy, the subjects agreed that redundancy would have been clearer if the two elements had been simultaneous. They used expressions like "I would have preferred to have them simultaneously. Then they would have been redundant." In other words, at least in this case, they found redundancy favorable.

The relationship between an animation and a simultaneous narration gave rise to an argument. One subject thought that the animation emphasized the narration. The other one found the relationship to be complementing.

Pair no.	1
Node	Ear (MSEncarta) (description on p. 78)

When listing and classifying the elements, the attitudes of the subjects toward the nature of different modalities became clear. Without reading the text field (except a rapid overview) one subject said: "This is information in its purest form." Probably, for him "information" means a collection of facts, and the role of text is usually to embody these facts. This kind of attitude restricts the possibility of interpreting messages on multiple levels and thus constructing combinations of messages (interface elements) with meanings in common. If, e.g. music and text deal with quite different levels of reality, multimodal redundancy cannot exist between these modalities.

---

1 The subject concluded that typographical means were not used because the style was interpreted as being very neutral. It could be argued, of course, that each style, e.g. font style, is a choice, especially nowadays when the font selection is usually enormous. Even "neutrality" is a choice that has a message.

Alternative controls such as different ways to close a window were mentioned to be redundant with each other, but they were not found to be essential enough to be noted. Actually, the alternative controls seemed to irritate the subjects, who did not understand the sense of them. The only reason that the subjects found for using alternative controls was in some programs to preserve consistency with older versions of the same product, but in this particular program that explanation could not be used. In other words, redundancy was not a desirable way (for them) to be provided with alternative ways to control the product. Alternatives were found to conflict with clarity.

Redundancy was recognized concerning the animation and the related narration, but it was not analyzed further or noted, because animation and narration were seen, or at least listed, as a single element.

These subjects spent a remarkably short time with this node. The enormous quantity of information may have made them give up. Many redundant messages were missed.

Pair no.	1
Node	Trumpet (description on p. 79)

In the classification task the subjects discussed the role of hot spots that activated links to pop-up windows containing more relevant information. After a discussion the subjects decided to classify the icons in hot spots as informative, although their obvious role was controlling. The subjects spoke about those icons as though they were the same unit with the emerging window. In other words, they so strongly associated the window with the element that activated it that they found it reasonable to handle them as one single element.

In the discussion about redundancy between the topic text and the vocal presentation of the same word, one subject noticed the significance of the character of the language used in this relationship: when English is used, the vocal presentation contains vital information about the pronunciation, but if the same product was in Finnish, the pronunciation would be obvious because each letter is practically always pronounced in the same way in Finnish. Thus, if the product were in Finnish, as the subject said, the relationship between these two elements would have been even more clearly redundant.

The disagreement about the relationship between the sound sample and the other elements was a typical example of difficulties in interpreting the nature of relationships. The subjects tried to find the one and only word to describe the relationship, but they actually found that there were at least two possibilities, and the task was to decide which one was dominant. The male subject thought that the only thing that the elements had in common was that they concerned the same object, a trumpet. So, he found these relationships to be complementing. The female subject felt that you can imagine something about the sound an instrument produces from its appearance. Thus, the picture and the sound sample would be redundant to some extent. The male subject illustrated his point by changing the node. He took some exotic, stringed instrument and argued that you cannot imagine what it sounds like on the basis of a plain picture. This was a brilliant example: why did he take an unfamiliar instrument? This illustrated his point of view; he takes the user as a 'tabula rasa,' an object of the information delivery act. In this sense, the trumpet was not the best possible example, since users in this case had so rich a mental representation of the instrument. When the users interacted with the multimedia product, they interpreted it actively, basing their work on their previous knowledge. And they found redundancy, not between a picture and a sound sample but between the meanings of the messages that were encoded in pictorial and sound formats.

This product evoked, again, the most positive reactions. At the end, even one subject who was very prejudiced and extremely critical admitted his fascination with the product.

Pair no.	6
Node	Airship (description on p. 77)

When the subjects began to carry out the task, they at once began to list and classify the elements. They were too hasty in filling in the form to be able to find out what this node was actually about. Even clear misunderstandings took place. For example, the subjects played the animation concerning the principle of reducing and increasing the mass of an airship, but did not find the sound that illustrated the streaming of the air to be informative! The post-test confirmed that these subjects did not understand the physical principles of the basic functions of an airship, although this was the main point of the node (this can clearly be seen in the name of the product, "The way things work"). Later, they read the text concerning the animation, but did not concentrate on it enough to catch the point. They even translated most of the text vocally into Finnish, but, unfortunately, left the essential part of the text for minor attention. However, the subjects understood the relationship between the animation and the text related to it: "In this text it is baldly described what happens in this picture," and the relationship was classified as redundant. This can be interpreted that the relationship was assumed to be as obvious as a picture and a caption. The content was not analyzed any deeper – the fact that these elements are located in the same window one on top of another seemed to have been the biased criterion.

The relationships between the main picture and some text elements were found to be complementing. The subjects got enthusiastic about this idea and were extremely satisfied with it. Unfortunately, their verbal analysis of the arguments for this decision was too scanty to be analyzed any deeper. So, it is hard to verify their mental representation of a complementing relationship. However, the delighted reactions indicated that such a representation exists and that it is quite strong. And when an interpretation that was consistent with the mental representation was found, the satisfaction was remarkable.

Hot-words that contained a link to a common node were classified as redundant. Some discussion took place about whether this is the case. The subjects seemed to have noticed the different function of redundancy compared with the redundant relationships found earlier: here, it was only about alternative ways of activating the same window.

Pair no.	6
Node	Sensory cells→brain (description on p. 78)

Just as in the investigation of the former node, the subjects appeared to have misunderstood that filling in the form was their primary task. The advice to follow the whole story before concentrating on the last node of it was probably either not understood or not found necessary or both. Thus, the central concepts were not learned and the ability to interpret the elements of the last node was not as good as it could have been. The confusion about the concepts was seen in some comments in which the subjects spoke about "sensory cells that swarm toward the brain."

The only redundant phenomenon found was the most obvious one: the written and narrated version of the same text. This was accepted without questions. This illustrates the somewhat superficial attitude of the subjects toward the task.

As the node did not overflow with different elements and the subjects did not analyze the relationships in a deep manner, not too much time was spent on this node.

One reason for the superficial attitude could be the time of the session; this was the last session in the afternoon, and exhaustion was reflected in the work of the subjects.

Pair no.	6
Node	Ear (MSEncarta) (description on p. 78)

Almost the first comments concerned the relationships between the elements. At that moment, the subjects had seen one picture and hardly had read more than the heading. The assessments about the relationships were, as in the former node, based on the layout and the first visual impression, not on the analysis of the content, not to mention the meanings.

There was exceptional quiet during this task. The animation about the function of an ear was investigated, as well as most options of this exhaustive information source about the ear. Despite all the multimodal representations, no redundant messages were found. This pair, like one of the former pairs, actually gave up and did not even try to analyze this node profoundly. They listed fewer elements than in the other nodes although the number of elements was the biggest of all. The pair was feeling very tired.

Pair no.	6
Node	Trumpet (description on p. 79)

Unlike in the earlier nodes, one subject began by trying to find out the contents of the text: she translated it completely and, after that, formulated the content in her own words: "This was about how the sound of a trumpet is produced." Once again, this product intrigued the subjects. Despite tiredness, they brightened up, and the product swept the subjects along.[1]

The subjects spent a much longer time with this product than with the previous one. They listed 16 different elements, which illustrates the thoroughness of the work when compared with the eight elements the subjects listed for the preceding node.

Despite the much more intensive work, the level of abstraction concerning the relationships between the elements remained low. Redundancy was search for and found in obvious similarities like the written and vocal forms of the topic and the main picture and the somewhat more detailed pictures of the same object. The subjects clearly found redundancy as a central concept of this study and felt that it was their duty to find redundant messages, however naive the findings might sound.

Pair no.	5
Node	Airship (description on p. 77)

When listing and classifying the elements, the subjects found it reasonable to handle elements in quite huge clusters. They divided the screen into four parts: text (referring to the main text), picture, navigation push buttons and what they called hyper-text (with this they seemingly refereed to hot-words and the related pop-up windows with text). This kind of oversimplification obviously restricted the possibilities for detailed analysis of the relationships between single user-interface elements. On the other hand, it provided a way to understand the subjects' mental representation about the overall structure of the node.

Unfortunately, the subjects did not find the animation illustrating the function of an airship. As a result, neither of them succeeded in the post-test despite their probable interest in technical subjects, i.e. they had a good likelihood of understanding the principles.

The first interesting discussion about the relationships took place when the subjects tried to decide whether the picture and the text were redundant or complementing. Their comments showed that they had clear mental representations about these concepts and that they could adapt them to use. "Those things do not actually appear in the picture… it is rather

[1] Although the purpose of this analysis was not to evaluate the products, it is impossible to ignore this repeated phenomenon. The discussion continues later in this report.

complementing than... You cannot see that it is filled with helium, for example... mostly complementing, just a little bit redundant." In order to get a better view of their thoughts, verbal argumentation concerning this little redundant phenomenon would have been necessary. Later, when hunting for redundant messages, they came back to this relationship and re-defined it as emphasizing. It was finally classified both emphasizing and complementing. However, the arguments for this decision were not too convincing.

The subjects classified the relationship between some push buttons ("History," etc.) as complementing. With this they referred to, not the actual push buttons, but to the nodes or pop-up windows linked to those buttons.

Later, when discussing the relationship between the picture and the push buttons, the other subject defined redundancy through negation: "Nothing is said twice." The conclusion about the relationship was "complementing."

At the end, the subjects once more tried to find some redundancy. They actually mentioned two such details, but did not find them worth noting: In the text an airship was compared to a submarine, and one subject interpreted this to refer to the appearance of the vehicle. But this was said to be too obvious. Another quite curious but smart finding was the accidental relationship between the hourglass ("busy"-icon of the pointer) and the noise caused by the function of the CD-ROM drive!

Pair no.	5
Node	Sensory cells→brain (description on p. 78)

The subjects acted according to the written instruction and investigated the contents in the recommended way: they followed the sequence of nodes containing vital information.

These subjects continued by handling elements as clusters, just as in the preceding node. The first comment when beginning to list the elements was: "Really, here are only three elements: picture, sound and text. And the push-buttons." The product's authors had clearly struggled hard to create a pleasant appearance, but all the decorative elements were completely ignored. This reflects the assumed purpose of the node: to deliver facts about the function of an ear. All the elements, the meaning of which is harder to interpret than the meaning of plain text, were ignored as of minor significance. Later, hyper-links with related pop-up windows and animation were added to the list.

The subjects discussed the relationship between the picture and the narration. They concluded that the relationship was complementing. The decision was based on negation, why the relationship was not redundant: "Neither of them tells the same thing alone." This was again a convincing proof of a strong mental representation of redundancy.

One problem concerning the definition of redundancy was very concrete in this node; the subjects wished to have both the text and the picture simultaneously visible. They discussed whether this was a redundant phenomenon at all as the messages could not be perceived at the same time. Whether a clear distinction between simultaneous and succeeding redundancy should be made is discussed later.

The narration and the written form of the same text were classified as redundant without any deeper analysis. The narration and the related animation were booked as redundant, as "the voice repeated the thing." The same kind of relationship was, of course, found between the written text and animation, because the written and narrated text had the same content. The narration, written text and the animation were thus seen to form a combination of three redundant messages.

These subjects closely argued and explicitly verbalized the basis for judging the relationship between the text and the animation as redundant. They analyzed which particular details were redundant: "The sensory cells bounded... and the signal transferred..."

Pair no.	5
Node	Ear (MSEncarta) (description on p. 78)

This pair of subjects, unlike most of the others, was not scared by the vast amount of information presented on one screen. The point of view of this pair was, as in the preceding task, biased toward facts. So, these subjects did not pay attention to the decorative points. Despite the fact that the animation illustrating the flow of signals in the hearing system was sparse and lacked visual decoration, it was regarded as superior compared with the corresponding animation in the preceding node.

Two relationships were classified as redundant without much discussion. The lack of verbal analysis reduces the informative value of these findings to this study. Perhaps the subjects overestimated their ability to interpret the relationships or underestimated the complexity of the task and did not find it necessary to verbalize their arguments in this case. The noted redundant relationships were between a picture and a written text and between an animation and a written text. The concentration on the task was not the best possible, judging by that the relationship between the animation and the simultaneous narration was not classified as redundant, although these elements certainly had more in common than the elements of the noted redundant relationships. In the discussion the relationship was rapidly mentioned as redundant and ticked off, but the checkpoint was in parentheses. Parentheses were added also to the checkpoint concerning redundancy between the main picture and text.

The subjects blamed this node for "cognitive overload," and perhaps the feeling of facing an impossible task reduced the accuracy of their work.

Pair no.	5
Node	Trumpet (description on p. 79)

The property of the user-interface that was the cause for the first comments was the simplicity compared with the previous node analyzed. This interface supported the habit of these subjects to make clusters of the objects; the visual objects with similar roles were in clear, distinct groups. Thus, the listing and classification of objects was carried out without effort.

When investigating the node, lively discussion took place. Disagreement stimulated the discussion and lead to very important analysis. It started in the discussion about the relationship between the written and spoken format of the topic:

– At least these are redundant... But can a picture and music be redundant?

– It is difficult.

– For example, a picture of an orchestra and the music it plays? Is it redundant or complementing?

– I would say it is complementing. Music is only a product of an orchestra, the sound is not an orchestra.

– In my opinion, it is redundant.

– The music is a sum of so many things. There are the composer, etc.

– But it is obvious that if there is a written text and it is read aloud, it is redundant. But it is because those media are so near each other... But if there are a picture and music...

– Anyway, they are complementing!

The subject who found it possible to have a redundant picture and music did not give up. Later, when they filled in the form and were to classify the relationship, the same discussion arose again. First, when thinking about the picture of the trumpet's valves and listening to the samples, the subjects agreed that the word "complementing" best describes the relationship. One subject mentioned that perhaps with more expertise about the trumpet the picture would have more content and redundancy could be possible. They also suggested that if there were animation about the function of a valve with a simultaneous sound sample, the two could have been described as having a redundant relationship. However, when analyzing the picture, sound sample and written text about different kinds of mutes, the possibility of redundant sound and picture was presented once more. All those three elements were finally defined as one redundant whole. And when noting the relationship between the main picture and the related sound sample, the other subject demanded it to have classified as redundant.

The analysis of this challenging interpretation task did not end with the listing. The subject who had obstinately defended his idea that even elements of very different modes can be redundant with each other continued to search for a verbal explanation for his seemingly extremely strong feeling that this must be the case. His experience of simultaneously seeing a picture of a trumpet and hearing its sound, was in some way so all-inclusive and satisfied his needs to experience the object, that he felt that this must be something more than just a complementing relationship. His attitude toward the different classes of relationships was revealed at the same time. Redundancy was for him something more, something worth aspiring to. And finally, he succeeded in verbalizing his arguments: "Music is a kind of... if you think about a *feeling*... if you take a poem that portrays joy or sorrow and take music in the background, is it complementing or..." The subject realized an important point: a message cannot always be described with words. Frequently, a meaning that is evoked by a physical message cannot be analyzed verbally. We can say that when combining two messages containing something affective in common, the result just feels good and produces a deep impression. *By merely arguing that the elements are complementing on the basis that, e.g., one element tells us that a trumpet looks like this and the other that a trumpet sounds like this, we end up with mechanical interpretations that do not support the kind of design that takes into consideration the human ability to experience things simultaneously on multiple levels.*

Pair no.	4
Node	Airship (description on p. 77)

First of all, there was a comment that revealed the expectations toward the product: "This is probably a good one since this is always put out."

The subjects had well understood the multiple roles of elements; when listing and classifying them according to role, they listed every possible role. E.g. hypertext was defined as informative, controlling and – to some extent – decorative. The subjects were oriented to the most important part of the task: they talked about the relationships already when listing and classifying the elements.

The attitude of the subjects toward the task was conscientious. They carefully read the text elements, e.g., the pop-up window containing the principles of regulating the height of an airship. However, some fatal misunderstanding about these principles took place. The subjects probably had some earlier idea about the function of an airship, an idea that was completely wrong. They did not concentrate on the contents enough to break that earlier idea down and adopt a new one. An explanation of this kind must exist, otherwise it is difficult to understand why neither of these subjects had any idea about these physical principles in the post-test. The result of the post-test was really surprising considering the accuracy of the work of this pair. Curiously, they even said that the pop-up window with the animation and sound was important. It could be said that they understood the idea of the structure of the node, but, unfortunately, misunderstanding took place at the worst possible moment concerning the content.

The subjects liked the back-button, which contained an arrow pointing left, the word "back," a little animation when clicked (the arrow made a tiny jerk toward the left) and an arbitrary sound effect. In other words, there were several elements with a common message within this control unit. This is actually a very clear example of redundancy, good also in the sense that the subjects found the result successful. But redundancy was not recognized. It is not likely that the subjects did not find the control elements important enough to be analyzed. They would, at least, have mentioned the finding vocally if they had recognized it. Instead, it can be interpreted that when they listed the elements and clustered all the control elements as a whole, they only searched for redundancy among the listed elements and did not even think of the possibility of dividing an element into sub-elements, among which relationships could be analyzed. Their thoughts were firmly locked into their first impression of how to look at the interface and what an element is.

This was the first pair that paid attention to the message of the texts' fonts. The subjects concluded that there was a hierarchy between the text units, and font selection illustrated the level in which a text unit was in that hierarchy. However, they did not consider the nature of a font choice to be informative (containing information about the level of a text).

When leafing through the pop-up windows with written text, the subjects longed for narration. "Narration completes the comprehension, somehow," one subject said. Many other subjects had found the narration irritating. So, *the lack of redundancy in this case implies a lack of freedom to choose the modality in which the message is received*.

The subjects were oriented to search for redundancy. They, or at least one of them, were prepared for this task by having formed a simple, easy-to-remember verbal definition of redundancy. When searching for it, she asked several times, which elements are "overlapping." This expression illustrates the nature of the subject's mental representation of redundancy.

The clustering of elements, e.g., the handling of the whole pop-up window concerning the function of an airship (text, animation, hot-spots, etc.) as one single unit, restricted the possibilities for accurate, detailed analysis of the relationships. Thus, the interpretations were shallow, and the listed relationships were obvious. The only relationship that was defined as redundant was the relationship between the main text and the main picture, in which the picture was seen as containing the written names of the parts of an airship around the picture. Most of those words were hot-words containing a link to a pop-up window with textual explanation of the concept. The same pop-ups could be activated from the hot-words of the main text. These alternative routes to the same pop-up windows were found redundant.

An important finding took place when the subjects thought about the relationship between the picture and the related text: "The text can never explain the picture exhaustively." There were good opportunities to analyze the power of a picture and the power of a verbal explanation, but, unfortunately, a slight hurry prevented the subjects from staying longer around this item.

Pair no.	4
Node	Sensory cells→brain (description on p. 78)

When "Journey to the ear" was followed, clear signals about an overload of information were heard. The product was experienced as extremely "heavy" (as the subjects described it). Illustrating animations were longed for, despite the high-quality animation concerning the actual moving around in various parts of an ear. Actually, one redundant element was suggested: when the narration contained information about the transfer of vibrations, the subjects wished to have an animation of them. After a while, they got just what they wanted when they found the animation about the item. However, the subjects did not sound delighted. Probably the poor graphical quality of the animation and the fact that it could not be followed in parallel with the narration resulted in quite lukewarm reactions.

In this node, the combination of narration, animation, and the picture, which also served as a background for the animation, was found very important. Above all, the animation and the related, simultaneous narration made an impression on the subjects. The 3-D effects were also noticed and found as signs of high quality design, but those effects did not take place in the node that was supposed to be analyzed. The relationship between the picture and the animation was classified as emphasizing and substituting, but the discussion revealed the latter one as an accident concerning the formal listing only. It was meant to be complementing, and actually the complementary nature of the relationship was found to be dominant.

The interpretations were hardly argued at all, just stated. So, this material is inadequate in order to verify the adopted concepts. For example, when the narrated and written texts were found redundant, the subjects only said that they are "the same." Later, the subjects talked about the narration and said that the opportunities to make the narration expressive (tone, rhythm, etc.) were not used, i.e., the narration was blamed for being monotonous. Perhaps this kind of implementation of a narration prevented the subjects from seeing the possibility of having information in narration that is difficult to express by means of other modes. But it can also be argued that the subjects discussed that very thing but in an unexpected context: not when comparing the written and narrated form of the same text but when criticizing the dull implementation of the narration.

Although the product's authors probably had invested many resources in its visual attractiveness, the overall judgment was that this product was biased toward information and dull. The listed elements were all classified as informative, and only one element (a picture) was found to have another kind of role additionally (distinguishing).

Pair no.	4
Node	Ear (MSEncarta) (description on p. 78)

Immediately upon finding the node to be investigated, the subjects opened the animation window and played it. The unavoidable comparison between this and the corresponding animation in the previous node ended up in an assessment according to which this product was superior to the previous one. The clarity of this node and the split nature of the previous one made an impression. The composition of animation, narration, and sound was noted as one unit and the relationships within this unit were described by classifying the whole unit as complementing, emphasizing, and redundant. Some reasons for the choices were given, and the subjects almost succeeded in defining the detailed criteria for redundant and complementing relationships:

– There are at least complementing relationships.

– Yes, and also redundancy takes place as it (refers to the phenomenon in question) is narrated and shown in animation (the Finnish word "redundanssi" was not used, but rather the synonym, which could also be translated "repetition").

– Is it redundant (the synonym used again) or complementing? Imagine that if you would not hear the text you would not necessarily understand. Or vice versa; it would be quite blunt without the animation.

(The final judgment was that the relationship was merely complementing, but there was some redundancy as well.)

It can be interpreted that there was an explicit criterion for whether the relationship was redundant concerning the facts. As the function of an ear was clearly presented with the combination of animation and narration, can the function of an ear be understood only on the basis of one of the elements (narration or animation)? The subject supposed that

1. the function probably cannot be understood on the basis of animation only
2. the narration by itself would make the presentation uninteresting.

But what would the appropriate conclusions have been? It was not argued that the narration did not contain all the vital facts, so let us assume it did. Then we can say that animation emphasized some important parts of the narration. And when emphasizing, both elements would inevitably have something in common, i.e. something redundant.

The comment, according to which plain narration would be blunt, takes into consideration more than just facts. However, the ability to analyze messages beyond the level of facts seemed to be out of reach.

The way in which the subjects treated this unit illustrates their delight with the design of this detail. Different elements were found to form a seamless whole that presented the information in an effective way. The reason for the positive reactions was probably that they started with this animation, and the first impression was therefore formed on the basis of the most appropriate part of the node. Even the sound related to the pull-down menus that irritated the other pairs almost without exception, was found a pleasant detail.

The subjects were so eager when investigating the node that it took quite a long time for them to realize that they had got lost when following the links of the hypertext. That was not a surprise since most of the elements of the user-interface looked alike except for the topic of the node. The heading was the only element that indicated the current location in the product, and thus, the subjects were lost until one of them paid attention to the heading field.

Pair no.	4
Node	Trumpet (description on p. 79)

The product was familiar to one of the subjects. The first reaction was that different media were effectively used, which was seen as a result of the nature of the product; sounds are in a central role. In many products, it is sounds that have caused irritation because of inappropriate use.

The user-interface was found to be extremely illustrative. When first discovering the node, the subjects began to think about the relationships and first met the typical question as to whether the text and the related music were in a redundant or a complementing relationship with each other. Also, a substituting relationship was searched for because the subjects realized that they had not recognized such a relationship so far. When searching for a substituting relationship, the word "surprising" was used to describe the object.

When the subjects were leafing through the node, they rapidly made several clever findings about the product and the relationships of the elements. They really seemed to have adopted the desirable point of view, e.g., there were icons with text (hot-spots, functioning as push buttons), and the definitely intentional redundant relationship between them was recognized very easily. The interpretations were suddenly even too bold; one of the subjects looked at the picture and the explanation of one part of a trumpet (the water key, in a pop-up window) and said without reading the text that the relationship was redundant. She even argued that "...there are enough sounds in the product, so you do not necessarily have to read the texts..."! The other subject led the discussion back to reality and the relationship between the picture and the text by saying that the role of the text is to focus, thus actually expressing the same idea as Bernsen (1994).

Once again, this product tempted the subjects to investigate the product more than would have been necessary concerning the formal task. The profound investigation was useful for the interpretations; the subject had time to think about the user-interface in a deep manner. The accuracy of filling in the form was not the best possible. For example, they forgot to classify the listed elements according to role. But that can be seen as a positive symptom of the eagerness with which they interpreted the product and discussed it. They simply did not have the patience to concentrate on the form.

Sound was seen to play a central role in the product, of course. As one subject put it: “Sound is the most powerful element here – even more powerful than the picture.” The strength of using sound was analyzed gradually going deeper and deeper.

– You cannot know what a piano sounds like if you do not hear it.

– That is the point of redundancy here.

In other words, the subjects intuitively experienced something that they wished to call redundancy, but could not verbalize the arguments at once. But later, they found a criterion for whether the relationship between a picture of a trumpet and a sound sample is redundant. One subject asked the other:

– When just hearing the sound, can you see the trumpet in your eyes?

The conclusion was that the interpretation depends on your previous knowledge about the instrument. If the instrument is unfamiliar to you, the relationship is complementing since the picture gives you information about the appearance and the sound sample about the sound. But if the instrument is familiar, you use your previous experiences and the bare sound evokes a visual mental image about the object. In that case, the additional picture is redundant to a high extent. And, since the picture and the sound are in a reciprocal relationship with each other, the picture also raises expectations about the sound if the instrument is familiar.

In this case, the subjects interpreted the relationship between the picture of a trumpet and the related sound sample from their personal point of view and classified it redundant. But when considering the arguments, they found three different things that made the relationship redundant for them:

1. that they could recognize the instrument on the basis of any of these two elements
2. the feeling
3. the style

Thus, they could analyze several dimensions in which the elements formed a more or less redundant whole.

Pair no.	7
Node	Airship (description on p. 77)

These subjects worried about the sufficiency of time available for the session. They counted how many minutes they had per task. So the way they worked was hasty. The classification of elements occurred on the basis of first sight without verbal analysis. One of the subjects gave a couple of verbal arguments about the details, but the subjects did not agree immediately. When the subjects then caused the application to crash, they were not too eager to try the functions any more. The constant threat of losing time affected the work.

When classifying the relationships between the elements the subjects started with the relationship between the push buttons and the related sounds. The decision about the nature of the relationship was based on one example, a button with an animation and a sound of bubbling liquid in a test-tube (referring to a "workshop"). In the picture, there were also other details from the workshop. The relationship was classified as "emphasizing" because "the sound does not tell, for example, where the liquid bubbles." So the elements could not be redundant. The relationship between the drawer-like buttons and the related sounds were also classified as emphasizing, probably because of the analogy with the preceding case. The decision lacked vocal analysis. The relationship between the topic and the text under it was seen as complementing: "It is like a nose in the head." The metaphor was to the point and indicated a relevant mental representation about the concept.

The point of view when contemplating the relationship between the picture and the text showed that the subjects had a mental representation not only about complementary cases but about the whole scale. The first reaction was to investigate whether the text contained something about the appearance. A disagreement forced the subjects again to clear verbalizing; one subject found the relationship redundant because of the existence of the propeller in both elements, while the other subject did not find the picture of a propeller obvious enough to be able to call the relationship redundant. In other words, both of them searched for overlapping information from two elements to be able to call the relationship redundant. In the end, when they had to describe the redundant messages they found, they once more stated their personal opinions about what was redundant and what was not. Only the opinion of one subject was noted (main text and main picture). The other subject only said that the relationship between the sound and the animation in a push-button ("workshop") was redundant if anything.

Unfortunately, the subjects did not find the animation concerning the function of the vehicle. As a result, they did not have an idea about the principles of the functions of an airship in the post-test.

Pair no.	7
Node	Sensory cells→brain (description on p. 78)

The subjects followed, as instructed, "Journey to ear". For one reason or another, the first impression was extremely negative ("Totally trifling shit!") It could be interpreted that the subjects felt themselves to be underestimated, as the most sophisticated animation clips were about the most obvious details while the product failed to illustrate the core of the content. In other words, the animation seemed an end in itself. It was used when it was easy to use, not to illustrate something that is difficult to explain verbally. The most resources seemed to be spent on the visual excellence, not on supporting understanding. The same kind of interpretations could be made of the comments of the other pairs, too.

The visual design even caused embarrassment. The subjects tried to find the meaning of the shape of the main frame of the screen that contained the picture or animation and a list of nodes. The frame had one round corner and three sharp. The conclusion was that it is a decorative phenomenon and "has no meaning." The subjects felt that this detail (and many others) was designed from the point of view of an artist or a professional designer, and it was not even meant to be understood by ordinary people.

The product did not motivate the subjects to profound work: the classification of the roles as well as the classification of the relationships was performed rapidly and in a shallow manner. The only discussion containing arguments was about the relationship between the narration and the written version of the text. It was classified as redundant as soon as the subjects discovered that the texts were identical.

Pair no.	7
Node	Ear (MSEncarta) (description on p. 78)

The huge amount of information was confusing for the subjects. On the other hand, most icons (like picture and video hot-spots) were either familiar or were instantly interpreted in a desirable way.

The subjects got lost in the encyclopedia by just trying what happens when clicking this and that. The way in which information was organized was not clear for them.

The classification of elements according to their roles was rapid and shallow, but this was the only pair that found the meaning of the background picture (a skeleton, indicating that this node was about anatomy). So its role was booked as informative as well as decorative. The role of the sound of the drop-down menus was found only decorative. The sound irritated the subjects and that is the probable reason that the similarity in the meanings (and the roles) between the animation of the dropping menu and the related sound was not noticed. Later, when analyzing the relationship between these elements, the sound was found to emphasize the visual dropping of the menu.

A disagreement about the relationship between the text and the related picture about the structure of the ear forced the subjects again to argumentation. The session was on Friday afternoon and one subject especially seemed to be tired. This might have hindered more profound analysis; the conclusions about the arguments did not reach very deep. The relationship was finally classified as redundant. The other subject clearly found the picture so obvious that it could be interpreted as redundant with the related text. The instruction was learned well; the relationship between the heading ("Ear") and the picture was booked as complementing because the heading did not have a predicate. The instructions said that a meaning can be expressed with a complete sentence, and the judgments about possible redundancy should be based on these verbal interpretations about the meanings. Now the instruction was followed too "literally" by applying it to the physical form of the message instead of the interpretation. In other words, the message of the heading "Ear" could be something like "This node contains information about the ear." And that interpretation was supposed to be compared with the meaning of the picture. Probably, something in common would have been found.

Surprisingly, the subjects found the former product about ear clearer than this one. Although the visual excellence of the preceding product did not mean that it would have been more illustrative, the overall appearance must have been clearer because of the enormous number of different elements in this product.

Pair no.	7
Node	Trumpet (description on p. 79)

The subjects started with listing the elements, but then began to investigate the product. Again, this product stimulated the subjects and encouraged them to trials. The subjects were delighted and expressed it clearly in their conversation.

The formal task was performed in a hurry. The result was a very typical set of classifications. Because of great agreement there was no verbal argumentation until the subjects thought of the relationship between the main picture and the main text. A very usual combination, the combination of a picture and a text unit, evoked again the most fruitful conversation about the relationships. As a result, the relationship was classified as redundant because there were a couple of statements about the appearance of a trumpet in the text.

6.5 Summary of the analysis

The subjects took the analyzing task seriously and worked in an accurate manner. Several little squabbles indicate that they found the task worth arguing about. Two pairs were tired, probably because of the time of day, and did not give their best all the time. Both observations about tiredness concerned the last session of the day and the same product (Virtual Ear).

TABLE 2 Number of elements listed

	Time of session		Number of elements listed / node			
Pair	Day	At	1	2	3	4
2	We	16-18.30	11	9	14	-[1]
3	Th	8-10.30	16	12	14	9
1	Th	10.30-13	11	8	9	10
6	Th	13-15.30	14	8	8	16
5	Fri	8-10.30	4	6	10	8
4	Fri	10.30-13	8	5	4	9
7	Fri	13-15.30	11	9	12	10

6.5.1 First node ("Airship")

The first node was in some aspects a warming-up task. The subjects were enthusiastic but their view of the task was still shallow. Their view of the nature of the task was usually revealed at the beginning of the session. Some subjects tried to fill in the form rapidly either because the formal analysis was seen as the primary task (pair no. 2) or they worried about the sufficiency of time (7). One pair (1) expressed their opinions about the design so eagerly, especially in the beginning, that the task seemed to be rather an evaluation than an analysis for them.

All the subjects seemed to have felt sufficiently informed about the task. They began immediately, without questioning the appropriateness of the task. They took the classification as given and evaluated the example products. They did not question the appropriateness of the classification concerning given examples.

The design of the first node was very rich. Despite its logical structure, the richness inevitably led to a vast quantity of different elements for the subjects to analyze. In practice, the subjects developed two strategies to cope with the material:

1 Pair number 2 had a different product.

1. Clustering of elements into functional wholes. This way the subjects could at least argue that their work covered all the details. The disadvantage of the strategy was that it neglected many tiny details, the importance of which is hard to say. For example, most pairs clustered the push buttons of the first node as one element. As a result, they missed the clearly redundant relationships between the related icon, the animation and the sound. Without any doubt, the details of the push buttons were not of great importance concerning the contents of the node, but the object of the analysis was the user-interface, not the content. Somehow, the clustering of elements limited the perceptions of the subjects. One pair (4) discussed the push buttons, analyzed them verbally, and liked the way those push buttons were implemented but did not even mention the word "redundancy."
2. Leaving out less important elements. This strategy was widely used in the third node. The problem of the strategy is that in it you have to rank the elements. Ranking was usually based on the quantity of facts that the element contained. Thus, this strategy stressed facts and usually ignored other aspects.

Some kind of clustering inevitably takes place in listing perceived elements. Basically, it is a question of what exactly constitutes an element. This is discussed later, but in table 2 we can see that the degree of clustering was very different with different pairs; the number of elements listed varied between four and sixteen. However, the pair that listed only four elements also left some details completely.

6.5.1.1 Conclusions based on the post-test

The subjects' levels of previous knowledge about airships and the physical principles of the functions of an airship were very different. The post-test measured their knowledge after their investigation of this multimedia product, but in order to draw conclusions about possible learning effects, a pre-test would have been necessary. However, nothing indicated that the method of changing the mass of an airship would have been familiar to any subject before viewing the multimedia product.

As explained in chapter 6.1 (p. 77), the purpose of the post-test was to provide a way to evaluate the reliability of the observations. Four pairs (2, 3, 4, and 6) found and investigated the animation about the ascending and descending of an airship, and pair 4 was found to pay most attention to the crucial details concerning the content of the post-test. Therefore, it was expected that pairs 2, 3, 4, and 6 would have stood out of the rest of the pairs and that pair 4 would have been the most successful. However, the result was not that clear. Only two subjects had, according to the post-test, learned the principle of changing the mass of an airship. Those subjects were from pairs 2 and 3. But the most unexpected result was that the subjects of pair 4 did not, according to the post-test, have any idea about the essence of the contents. Possible reasons for this have already been discussed in the analysis session of that pair (p. 96).

Even if the post-test failed to serve as a proof for reliability as it was intended to do, it is reported here for other reasons. It should be noted that this node was the first one and conceptions concerning redundancy were still very shallow (see next chapter). It was not yet understood that the detection of redundant meanings requires deep analysis of the meanings of single messages. This, in turn, was quite in harmony with the observation that the subjects only paid attention to the surface structure, thus neither being able to detect redundancy nor being able to construct knowledge, i.e., learn.

6.5.1.2 Redundancy

When analyzing the first node, most pairs had a shallow mental representation of redundancy and the other concepts concerning the relationships. On the basis of the instruction lecture, most subjects seemed to think they immediately got an appropriate view of the concept. This node made some of the subjects notice that it is not that simple. The best way to make the subjects aware of the complexity of the object was to force them to face disagreements with their partners. These disagreements led to successful discussions in three pairs (3, 1, and 7). One pair had already searched for a formal definition of redundancy (1) and made enormous progress in that process. The typical verbal expressions concerning the mental representations of redundancy were something like ”overlapping” (1, 4) or “twice” (1, 5).

Due to the shallow view of redundancy, the focus of most subjects was on finding a clear repetition of the same information. It seemed easier to find this kind of phenomenon by comparing two written text elements (pair 4) or written text and a picture (7, 3). In both of these cases, it is a question of successive redundancy. Probably, if the subjects used words like ”repetition” and “twice” to illustrate the meaning of redundancy, those words oriented their interest toward successive redundancy. The clearest case was pair number 4. The subjects were strongly oriented to find redundant relationships. They discovered the animation about the functions of an airship with related text and sound, which was a very good example of using simultaneous redundancy. They even found this window very important. They also discussed animated push buttons with sound and expressed their satisfaction with the design. But they neither mentioned nor listed any redundant relationships concerning these details. Their view of redundancy was still so superficial that they could not recognize it in a context where they did not particularly search for it. They did not search for it in these contexts because they had clustered the elements and searched for redundancy only between the clusters, not within them. Pair number 2 also found the animation window after having completed the listing. They recognized redundancy but did not note it because they were afraid about the sufficiency of time if they made the listing more detailed to be able to note the relationship found.

After all, the most clear redundant message found was the one discovered by pair 5, namely the noise of CD-ROM drive and the hourglass pointer. This is an interesting case, because the noise of a disk drive can hardly be seen as a message since it is accidental. For the receiver, the original purpose (or the lack of it) of a signal is not important. What matters is the meaning that the signal evokes.

6.5.2 Second node ("Sensory cells->brain")

The layout of the second node could be seen as either simpler or richer than that of the previous one; the subjects, who did not pay attention to decorative elements, listed only a few elements. The product was produced in a design school, and that was easy to see. Most of the subjects stayed in the layer of facts and did not bother to discuss the meaning of decorative elements.

The visually dominant element, the background picture, was listed by all except two pairs (4 and 5). Despite the fact that most pairs noticed it and classified it as at least decorative, in the discussions it was only mentioned, not analyzed any further. One subject said aloud what could be seen as an attitude of the other subjects also: "It is only decorative, so it does not mean anything." It was curious that even if anyone could see that vast resources had been spent on the appearance of the product, the node was interpreted as mainly informative. The meanings of the messages of many graphical details were probably so distinct from the other elements that they were seen as quite separate items that had nothing to do with the topic.

6.5.2.1 Redundancy

All pairs except 3 and 5 noted only one redundant message, the written text and the narration of the same text. Because typographical devices in the written text and intonation, etc. in the vocal version of the same text were sparingly used for expressing meanings, the elements were usually said to repeat the same thing. Two pairs, 1 and 4, paid attention to the possibility of using the unique means of narration (intonation etc.) and written text (e.g., typographical devices).

The most interesting findings came from pairs 3 and 5. The subjects in pair 3 found a good, though familiar, example of redundancy: the animation and the sound of a push-button. It was neatly implemented, and therefore it was surprising that this example was neglected by all the other pairs. The same subjects regarded the relationship between the animation and the related narration as redundant, as did pair 5. The subjects in pair 5 even hesitated to list the relationship between the written and vocal versions of the same text as redundant, probably because they found the case so different from the one with simultaneous narration and animation. Because simultaneous narration and animation was considered a successful combination, a clumsy, even unnecessary repetition of the same information was difficult to put in the same category. The subjects in pair 1 would have preferred the written and vocal versions of a text to be presented simultaneously. Seemingly, simultaneous and successive redundancy should be separated in the conceptual analysis at an early stage.

The use of narration and a simultaneous, illustrating animation is so obvious that one pair (4) even suggested it in when following the "journey to the ear." An animation took place just after the suggestion was made, but it was not simultaneous with the narration and therefore a disappointment for the subjects.

6.5.3 Third node ("Ear")

Consistency with the rest of the product is an obvious feature of a node of an encyclopedia. The strength of consistency is revealed when one uses the product regularly. The standardized user-interface becomes familiar, and human functions in the interaction process develop into automatic routines. The point of view was different when only one node was analyzed without the subjects having previous experience with the product. The subjects would have needed more tools to figure out the whole, consisting of a vast quantity of information on one screen. In this case, the screen was filled with icons and small text. Most of the icons appeared to be arbitrary for the subjects. Even the MSWindows'95-style push buttons for the most primitive procedures like minimizing, maximizing and closing a window were unfamiliar for some subjects.

Some pairs clearly found this node too complicated to be analyzed in the way they analyzed the previous node and actually gave up (6, 1, and probably 5). The subjects clustered the elements as much as they could, and many redundant combinations were missed because of that.

The dispersion of attitudes toward the node was remarkable. While most of the subjects found this node overloaded with information and, as mentioned, even gave up the task, one pair (4) was extremely impressed by the node. The differences between the work of this pair (4) and the rest of the subjects are interesting. The subjects in pair number 4 did not begin to fill in the form mechanically. They started with exploring the node in an intuitive way and did not immediately try to search for something they thought they were expected to find. Above all, they accidentally started with the animation about the function of an ear. Just a few minutes earlier they had analyzed the corresponding animation of the previous product and had a good opportunity to compare these two different kind of implementations. They found this one superior, and that finding served as a first impression of the node. It is likely that the first impression had a major influence on the way they interpreted the product. They even liked the sound of the drop-down menu, the one that irritated most of the others. The focus of the work had turned from the formal filling out of the form to discussion; pair four, despite all their enthusiasm, listed only four different elements.

Undoubtedly, the animation with the related narration was the core of the node. Furthermore, the implementation was as clear and illustrative as it could be concerning the facts. There was only the essential information and nothing else. Colors were used sparingly and drawings were very simple.

The comment of one subject in pair 2 best describes the overall attitude toward the implementation compared with the earlier node. He said that the pictures were worse in Encarta, but the whole was clearer. Only pair 7 found the earlier one clearer, but this interpretation seemed to be based on the quantity of different elements.

6.5.3.1 Redundancy

Three relationships were most often noted when the subjects searched for redundant messages: the alternative controls (1, 2), the main picture and text (3, 5, 7), and the animation with related narration (1, 2, 3, 4, 5). These were good examples of different kinds of redundancy. The first represents alternatives or repetition; the second, successive redundancy; and the third, simultaneous redundancy. The second one, picture and text, is here classified as successive because both elements are visual and therefore impossible to be focused on simultaneously.

The combination of animated push-button and the related sound was classified as redundant by one pair (3). It is likely that the other pairs did not find it possible to go to this detailed level since the number of elements was so big. A similar relationship, the one between the animated drop-down menu and the related sound, was not noticed at all. A possible reason for that, besides the one that was mentioned above concerning push buttons, is that the sound irritated most of the subjects. The irritation was understandable, at least with some pairs, because the sound properties of the computer did not work properly all the time, and the carefully designed, soft dropping effect turned into a horrible knock.

Because most of the subjects felt this node was overwhelming, it did not provide the best possible basis for deep analysis. Therefore, only one pair made noticeable progress in constructing a mental representation of redundancy and the hypothetical related dimension. Pair 4, which was the one whose attitude toward the product was clearly more positive than the attitude of the other pairs, had the motivation required for this kind of analysis.

Until this stage, most of the pairs had proved to have formed a mental representation of redundancy that served as a tool to see user-interface elements of different modalities as a whole if the meanings of the different elements have something in common, thus forming one, immaterial object. Exceptions are pairs 2 and 6, which still searched for repetition of the same information. Pair 7 was one step further; the subjects understood that the target was deeper analysis and tried to apply mechanically all the advice they had got in the instruction lecture, but were too tired to take the risk of relying on their personal interpretations. It is not surprising that all these pairs with shallow representations and mechanical interpretations (2, 6, and 7) worked as the last pairs in the afternoon. All of them indicated clearly that they were tired.

It cannot, of course, be concluded that four-sevenths (8 out of 14) of the subjects had formed an appropriate mental representation of redundancy. Theoretically, it is possible that only one of the subjects succeeded and the other one just approved the interpretations, in which case only 4 out of 14 subjects (two-sevenths) would have succeeded. But at least all the subjects of pairs 1, 3 and 5 were eager to express their opinions and reacted immediately if their partner suggested something with which he or she did not agree. Actually, the case could be the opposite. Pair number 2 was said to have given superficial interpretations. However, the female participant was so shy and silent and seemed to withdraw from formal decision-making concerning entries on the form that it is possible that she had important thoughts that did not come out. This kind of assumption does not lack evidence, as she had perhaps the most accurate and appropriate report about the functions of an airship in the post-test, even if nothing indicated in the discussion that she had paid exceptionally close attention to the content. It is likely that when her garrulous co-worker talked and talked, she had plenty of time to think.

6.5.4 Fourth node ("Trumpet")

When summing up the analyses of the fourth node, it is obvious that this node, without a single exception, inspired the subjects and appealed to them. In the afternoon sessions (pairs 6 and 7), the subjects woke up and began to work in a quite different manner than before. Suddenly, all symptoms of working merely out of duty vanished and seemingly real interest in the product appeared instead. The circumstances were fruitful for analyzing in a deep manner.

The secret of this product is discussed earlier, and the exceptionally appropriate use of sounds was found as one hypothetical reason for its appeal. The kind of hardware configuration that is usually called a multimedia workstation simply provides suitable tools to present the content of this product. It is hardly chance that this content has been chosen for one of the first multimedia reference products. This has been an obvious way to demonstrate the sound properties of current multimedia systems.

Should we then reject the analyses of this node because the content was exceptionally suitable for the hardware and actually tailored for it? Rather, we should take this as an opportunity to have a node in which information is not compressed in a constrained manner in the multimedia format.

There is another possible reason for the appeal of this node, too. The subjects spoke about clarity, which can be interpreted to refer to two different properties:

1. the reasonable amount of information (visual elements) on one screen
2. the layout of the screen.

The screen consisted of a blank white background and a few elements that were grouped according to their role. The user of this product is able to figure out the whole without constantly scanning the screen. The contrast with MSEncarta is remarkable. Basically, it is a question of the structure, i.e., how different information units are linked to each other.

6.5.4.1 Redundancy

The data concerning the fourth node did not show any drastic change in the subjects' understanding of the basic concepts. Rather, the analysis of this section confirmed the earlier findings.

Pairs 6 and 7, who had quite a shallow conception before this particular task, continued at the same level despite their eagerness. Pair 2 did not view this node at all. Pairs 3 and 1 continued at the same level as earlier. Perhaps the slightly presumptuous attitude of one subject in each of these pairs prevented them from going further when they had already formed some mental representation. The most interesting findings came from the conversation of subjects in pairs 4 and 5. Both of them, or at least one subject of each, had an even stronger need than earlier to express the impression that two elements, even if they are of different modalities, can form a seamless whole in which the whole is more than just a sum of its parts. With both of these pairs, the culmination was the analysis of the relationship between a high-quality picture of a trumpet and the related sound sample. Both of them wanted to call the relationship redundant, but were not able to argue mechanically by forming verbal interpretations of the meanings of both elements and then comparing them. Instead, both understood that the common meaning is at a level that cannot be verbalized. Both of them spoke about feeling. In this way, they had named one layer behind the layer of facts. The overlapping of these elements, sound and picture, took place at that level.

6.5.5 A formal definition of an output element

A message from a designer to a user is embodied in a user-interface output element. Because of its concreteness, the output element was chosen as the primary object of analysis in the experiment instead of the message. The subjects of the experiment were not given any formal definition of an output element. Because they had to list the elements, they first had to form at least an implicit definition of the concept. The concept is briefly discussed here as one conclusion of the experiment.

The way the subjects understood an element appeared quite obvious and thus the lists of elements of different pairs were coherent. However, the number of listed elements varied a lot. The reason for this variance was basically not that the different subjects would have understood the concept "output element" in different ways. Instead, two other differences in the way the pairs worked resulted in a different number of classified elements. First, different pairs grouped the elements in different ways. Instead of analyzing the relationships between each single element and every other element, the subjects grouped elements that had both physical and contextual similarities. The criterion for grouping thus varied, resulting in different kinds of lists. However, controlling elements like a set of push buttons (backward, forward, etc.) and menu commands were grouped by all. Another factor that varied among the pairs was the thoroughness with which they performed the formal task, i.e., did the subjects find their task was to list every single element, or did they leave out some they found less important.

According to this experiment, the distinguishing of an output element does not seem a problematic task. However, even if the way the subjects understood the concept was coherent, a formal definition is necessary for complicated cases.

The definition is based on the observations of the subjects. That way it is possible to find a definition that is as close as possible to the existing intuitive definition. Thus, defining a concept is, in this case, constructing a formal expression about a phenomenon that is already an object of mental representation.

In figure 25, the concept is approached through visual presentation. The drawing is abstract in nature, and the borders do not represent a VDU or anything else concrete. The purpose of the illustration is to clarify the role of an element as a part of the user-interface. The arrows in the drawing refer to contextual relationships.

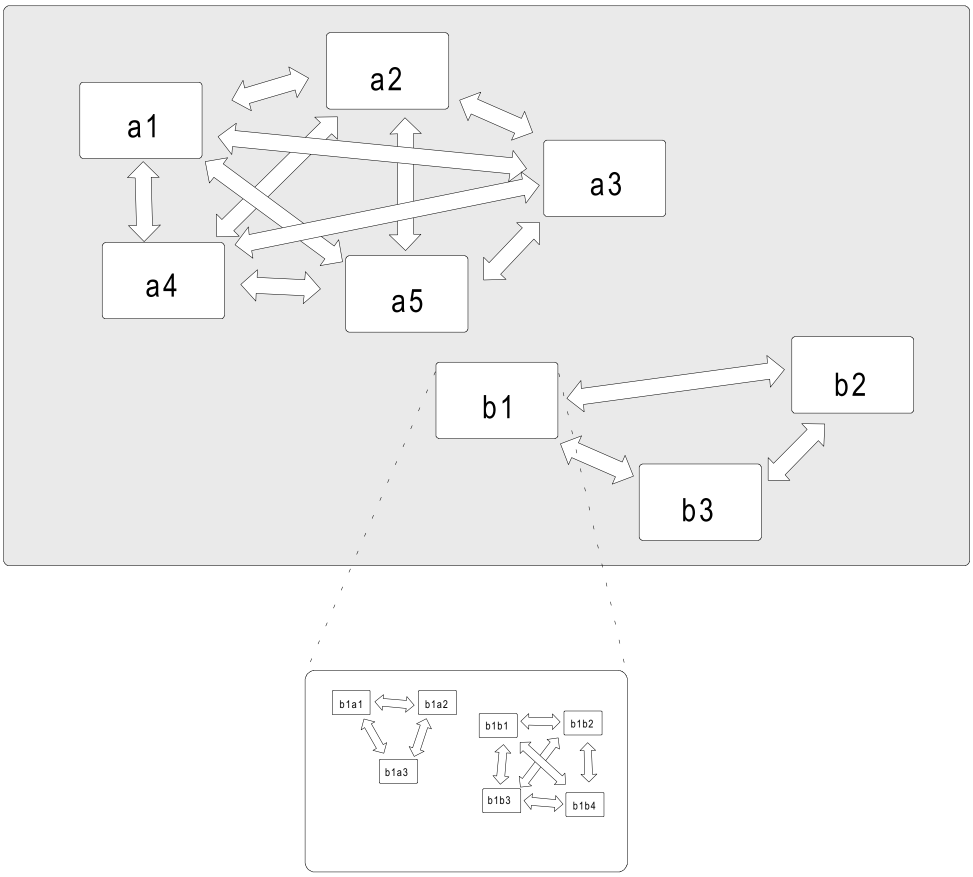

FIGURE 25 The relationships between output elements and the hierarchical structure.

In a user-interface, the elements form a hierarchy as illustrated in figure 25. When the actual elements of the user-interface are distinguished (a1, a2 etc.), they can be further divided into sub-elements (e.g. b1a1). Sub-elements can be further split, in principle to the level of the pixel (in a VDU) or the shortest distinguishable sound. In this definition, sub-elements are elements of their parent element. Only the highest level elements are elements of the whole user-interface. For example, in figure 25, b1a1 is not an element of a user-interface but an element of b1, which, in turn, is an element of the user interface. Sub-element b1a1 has no contextual relationship to any first-level elements other than its parent element b1. Its informative value in the whole interface is bound up in b1. In summary, it can be said that a user-interface output element is a distinguishable audible or visual unit produced by some device of a computer, a unit that has an independent informative value concerning the content of the product that launched it.

In the experiment most pairs handled groups of elements as elements. It is reasonable to group elements that share characteristics that are the criterion for their relationship to the other elements. In other words, members of a group share many features and differ in few, and the differing features do not govern their relationship with their environment. But a group is only an abstract object, and grouping in user-interface analysis is only a way to treat a set of elements in a similar way. Therefore, the formal definition of an output element does not need any extensions concerning groups.

7 DISCUSSION

7.1 Conclusions and a critical view

The preliminary definition of redundancy with the related conceptual environment proved applicable in user-interface analysis. It can be used in the context of theories of attention and thus be applied in parallel with these cognitive models. Additionally, the experiment broadened the preliminary definition to levels in which messages are difficult or impossible to verbalize. However, these conclusions should be exposed to criticism for several reasons. Some critical notes are presented here concerning the use of literature and concerning the experiment.

7.1.1 Critique about the use of the literature

The theories of attention and the related IP-oriented cognitive studies are broadly handled in the current work. The major problems in using them concerned ontology, epistemology, and the prevailing paradigm.

The problems of ontology and epistemology are already discussed in the introduction. The problem was caused by the contradiction between physical and phenomenological approaches as in the traditional mind-body problem. The problem was met in about the same manner as that of Jackendoff (1989) in his book "Consciousness and the computational mind." Jackendoff simply divided the mind into physical and phenomenological parts. Likewise, in the current work the IP-approach is used for the analysis of phenomena with a clearly physical nature, while a broader view of mental events is searched for among less materialistic approaches. However, this kind of division into physical and non-physical phenomena concerning mental events can reasonably be blamed for artificiality. It is to a high extent common sense that "mind" and "body" are inseparable and that the boundary between them is vague. But the debate about this controversy is probably endless (unless scientists someday succeed in creating an artificial consciousness). Concerning the current work, a much more immediate concern is the way that research reports of different eras are used. The use of literature of different paradigms was earlier defended by arguing that they are re-interpreted within the prevailing paradigm. This is certainly not a strict "kuhnian" idea. Interpretations are – according to T. S. Kuhn (1970) – "incommensurable" (p. 149) since the prevailing paradigm influences observations. Thus, a researcher of the behaviouristic era discovered different items than a researcher with a constructivistic approach, even if both had observed a similar event. It can be asked how a phenomenon could be reported if it is not discovered. As can be seen, the use of literature of a rejected paradigm is not simple. In the current work, that problem is salient.

7.1.2 Critique concerning the experiment

The current work is not an empirical study in the strict sense of the word. The focus is on conceptual analysis, but a concise experiment was found a reasonable supplement in order to make the theoretical framework easier to apply in practice. However, as reported earlier, the role of the experiment proved much more central than expected. The findings were so interesting that the subject would have deserved a more thoroughgoing empirical setting. Now, the small size of the sample reduces the reliability of generalizations. Generalizations are of course not the most important aims of qualitative analysis, but the results of the experiment undeniably encourage the author to acknowledge the clear similarities in the evolution of conceptions of different subjects.

The methodological choices were made in the terms of the data. A phenomenographical approach to data-analysis seemed appropriate, especially since different categories of conceptions began to take shape at an early stage. But as categorization is clearly one form of quantification, the phenomenographical approach itself reduces the possibilities of paying the most attention to the qualities of single cases. Here, the author tries to compensate for the inevitable loss of data caused by categorization by reporting each case in great detail. But the interpretation of that report is a demanding process. The reader has to *interpret* the researcher's *interpretations* about the subjects' *interpretations* concerning the user-interfaces of the multimedia products. As interpretations are always subjective in nature, the appropriateness of a meaning resulting from three-stage interpretation can be questioned. On the other hand, referring to the semiotic models of communication (chapter 1.2.3.2), the essence of scientific communication, like writing a research paper, is not a mechanical transfer of meanings. Any kind of reporting of any kind of research finally results in subjective knowledge that is constructed by the reader. The essence is that we as authors or readers of a report understand this.

7.1.3 The temporal aspect in the definition of redundancy

The results of the experiment clearly showed the significance of the separation of successive and simultaneous redundancy, which means that the temporal aspect has to be taken into consideration. Likewise, the analysis of redundant phenomena from the point of view of IP-models revealed the need to extend the preliminary definition of redundancy to take the temporal aspect into consideration.

The role of the theories of attention and other IP-oriented theories in the current work is above all to provide a means to analyze the processing of redundant messages. The analysis is necessary when trying to avoid cognitive conflicts in which the processing of one message interferes the processing of another. The possible conflict was earlier described as time-sharing or sharing of resources, depending on the model. Especially in the concept of time-sharing, the temporal aspect is obviously present. A simple solution for the problem of cognitive conflicts would be to avoid presenting multiple elements simultaneously. However, this kind of conclusion can not be made for several reasons. First, the nature of the application usually requires simultaneous messages. Second, even in the most detailed IP-models the timing of processing has a minor role; thus, it is impossible to define the adequacy of delay between two successive presentations. Third, simultaneousness seemingly has, in certain conditions, an important role in information presentation; this is discussed next.

In the experiment, the most expressive redundant combinations of elements were simultaneous. On the basis of the experiment, more can not be said than that those combinations resulted in better user-satisfaction, which is one criterion of usability. On the other hand, the most irritating combinations were successive and were seen as unnecessary repetition. Besides user-satisfaction, successive redundancy has, in certain conditions, clear advantages in cognitive processing. Dual coding theory provides a formal explanation for this (see chapter 4.2). However, these explanations (e.g., Mayer, 1992) only concern problem solving with an instruction that combines narration and animation. In a very simple reaction test with both audio and visual stimuli the simultaneous condition was found the most effective (Treisman & Howarth, 1959).

The illustration of the nature of a virtual object (figure 26) also supports the efforts to construct simultaneous combinations of redundant elements. As we get information about a real-life object in multiple forms simultaneously, so a virtual object should affect its environment in many ways at the same time. For example, a push button would lose its sense as a virtual object if the clicking of the mouse, animation of a sinking button, and a suitable audio effect took place sequentially and not simultaneously. In real life, we instantly recognize phenomena in which, e.g., what we hear and what we see do not synchronize. Until a child understands a physical explanation, it is quite mysterious why thunder and lightning are not simultaneous. The child has difficulty understanding that they are caused by the same physical event. We are used to simultaneous redundant multimodal information concerning our natural environment.

7.2 Creating virtual objects

The main motivation of the current study was earlier described as to define a conceptual basis for the needs of multimodal information presentation. The conceptual framework has arisen from the interest in the relationships among output elements of a user-interface. The relationships are discussed concerning their contribution to communication. The underlying motivation has been an attempt to see an application with its user-interface as a meaningful whole, consisting of meaningful wholes. Therefore, the atomic level of the approach is neither a pixel nor some other physical unit, but a message.

A multimodal communication environment makes it possible to create multimodal combinations of messages that can be interpreted to form a single whole. But should this whole be called a message or a combination of messages? This problem concretizes the core of the nature of creating virtual environments consisting of virtual objects. The comparison between a real-world object and a virtual object is illustrated in figure 26. The drawings illustrate the formation of a mental representation on the basis of external and internal events. The physical origin of the process that results in the formation of mental representation is a single physical object. That object affects its environment in multiple ways. For example:

- It absorbs part of the light that falls on it and reflects the rest.
- If it or some part of it moves, it causes changes in the material that surrounds it.
- It fills an identifiable portion of space.
- It might emit particles in its environment.

From the point of view of a human being, the first property causes a visual perception. The second, assuming that the movement is vibration and the frequency is between 20 and 20000 Hz, produces an audio perception. Because of the third property it is possible to identify it by touching it. The fourth

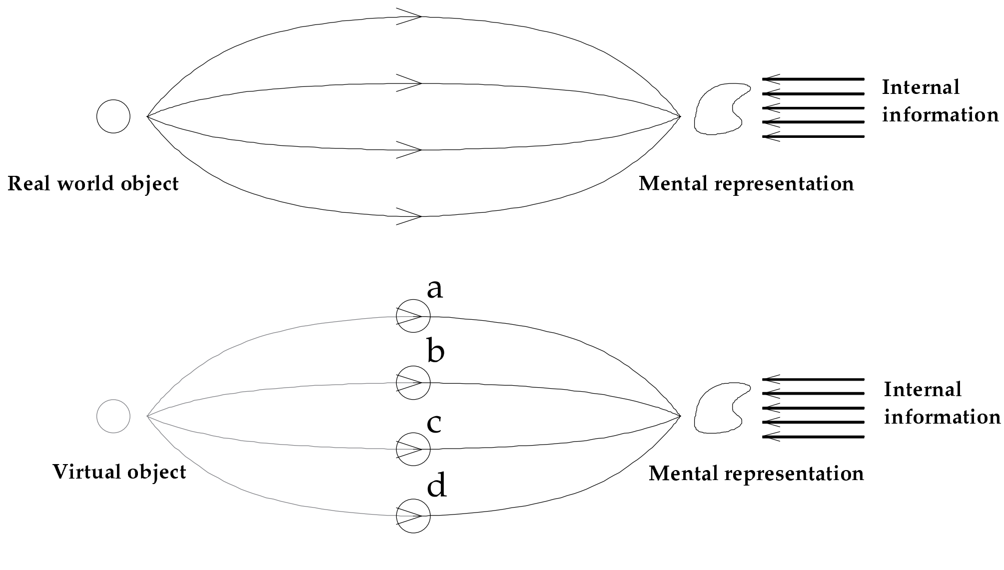

FIGURE 26 Formation of mental representation on the basis of real and virtual object.

property mentioned might cause, in certain conditions, a perception of smell.

The physical effects of the object in its environment are illustrated in the drawing with four arrows. These external information sources, together with existing mental structures (in the drawing labeled as internal information sources), cause a mental representation of the object.

The lower drawing illustrates a situation in which a similar mental representation results but the physical object is missing. In this case, the primary physical origin is not one single object but several qualitatively different output elements (a, b, c, and d). Each of these presentation elements must be produced individually. The physical property that combines them is that they are, for example, activated simultaneously as a result of the same user action.

An ambitious designer of a virtual object would probably wish to use all available media in order to provide as natural an impression as possible. The primary motivation to naturalness might be, e.g., user satisfaction. But from the perspective of the preliminary definition of redundancy (chapter 2), the interesting question is the cognitive consequences of using multiple media when presenting the properties of a virtual object. According to the preliminary definition, different presentation elements can easily be interpreted to be clearly in a complementing relationship with each other. For example, if the real life object that is imitated in a virtual environment is a dog, the VDU provides information about the appearance of a dog and the sound device about its voice. But as found in the experiment, in some conditions, the combination of elements that are complementing at the cognitive level, may finally be much more than just a sum of its parts. This finding requires an elaboration of the model illustrated in figure 5 (on page 34). The model was defined to concern facts.[1] On the basis of the results of the experiment it seems reasonable to expand the model by adding a third dimension to it. In that dimension, the layer of facts is only one level.

The need to analyze real-life objects in a deep manner, i.e., the need to go beyond the physical and logical facts, has resulted in different models in different disciplines.[2] In education, learning objectives have traditionally been split into cognitive, affective, and psychomotor levels. Bloom (1972) developed this taxonomy for the needs of planning and evaluating education. It is a clear example of an attempt to broaden the conceptual basis of the analysis of the interaction between a human being and his environment. Bloom's taxonomy can be seen as a rough outlining of the third dimension of the proposed model (figure 5), or at least an attempt to name more levels than just facts. The attempt

---

1 In order to be more exact in the use of concepts, the preliminary model can be seen to handle *physical* and *logical* facts.

2 The idea about multidimensional wholes may have surprising applications. Lang and Friedstad (1993) formed a bold hypothesis that is based on theories according to which both verbal information and positive emotions are processed more by the left hemisphere and, in turn, negative emotions and visual information by the right hemisphere. Lang and Friedstad get some support in their experiment for the hypothesis that memory for emotionally positive messages is visually biased and for negative messages verbally biased.

is respectable especially since it is done in the behaviouristic paradigm. Thus, only perspectives that can be seen to have solid connections to overt behaviour could be taken into account.

Figure 27 illustrates the current view of the contribution of Bloom's taxonomy to analysis of a real-life entity at more than one level. By identifying and naming three levels the model is a much better basis for designing a presentation of the entity than a description concerning just facts (the nearest counterpart of facts in the figure is the cognitive level.) The figure illustrates how the taxonomy, however, only concerns some thin slices or sections of the entity. By increasing the number of layers it is possible to get a better and better view of the whole. But finally, that approach can not lead very far since basically it is a question of classification, and classification inevitably means reduction. If it were possible to define a whole new perspective, we could see the whole profile of the object instead of a set of sections. The new perspective would mean the definition of the third dimension discussed above.

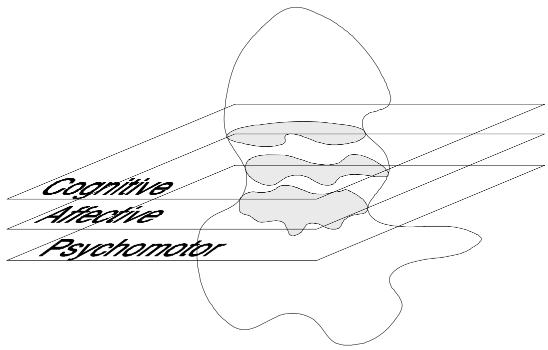

FIGURE 27 Three layers derived from Bloom's Taxonomy.

As mentioned, Bloom's model is constrained by its behaviouristic basis. Conscious experience, understood in the way it is handled in consciousness studies, could be a promising framework in formulating the new perspective and dimension within the contemporary paradigms. Likewise, the implicational level in the ICS-model, if further developed, could be the new perspective. The problem of using the current version of the ICS-model is that the properties of the implicational level are not yet fully articulated. Or the discussion about the dimension could lead to analysis about the relationship between science and art and their roles in outlining reality. However, the results of the experiment suggest so clearly expanding the two-dimensional model that this can not be rejected even if the expansion is likely to make the model much more complicated.

Although it is not yet possible to define a third dimension in the model of relationships between output elements, the contribution of the approach to the discussion about redundancy can be illustrated as in figure 28. The figure illustrates two objects and the multiple levels that define their properties. The objects illustrate mental representations of concrete or abstract entities. The tiny white area on the top of each refers to the layer of factual information. The illustration stresses the conception according to which

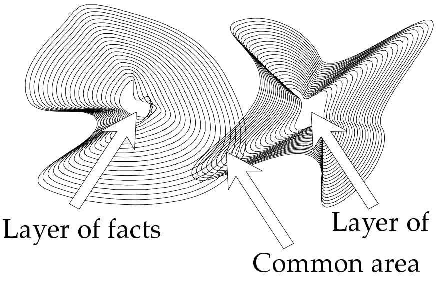

FIGURE 28 Two representations with common area outside the layer of facts. (Pirhonen, 1997)

this layer is only the tip of an iceberg. In the illustrated case, the relationship would be classified as *distinct* if only the uppermost layer is taken into consideration. But if the interpretation of the underlying mental representations goes deeper, clear similarities are found at some level or levels. Some degree of redundancy takes place in a form that cannot be explicitly defined or described verbally. If the real-life objects that have evoked the representations illustrated in figure 28 are two output elements, they can be seen, according to the illustration, to form a meaningful whole. That whole is – in turn – a representation of a virtual object.[1]

The example above about a virtual object was a dog. Creating a virtual dog is, in a certain sense, quite easy; the designer tries to imitate a real-life object. The process of creating such a virtual object can be seen as a mechanical reproduction of reality. The designer takes a photo of a dog and records its bark. The material is then simply stored in digital form and linked or embedded to an appropriate application. In fact, the process is, of course, much more complicated. Just the taking of a simple photograph of a dog implies an enormous number of decisions that all affect the result. A designer's skills in using a picture as a communication tool define the quality of the result.

In the illustration (figure 26), cases that were compared with each other were the formation of a mental representation of a physical real-life object and the formation of a mental representation on the basis of a virtual object. But if the virtual object lacks a real-life counterpart, there is no physical object that could be imitated or that could be reproduced. In that kind of case, the origin of the presentation is the designer's mental representation of the object. It is a question of the skills of the designer to express with the available means something that has a highly abstract and subjective form in her or his mind. On the other hand, this can also be generalized to cover the cases in which a virtual object does have a concrete counterpart. Finally, the designer's mental representation of the object rules her or his decisions in the design process. Referring to the example above about the complexity of taking a photo, the mental representation of the object of a photo rules the decisions in that situation. The whole repertoire of photographic means of expression is – or at least should be – in the use of the designer to express something essential about her or his mental object whose physical counterpart is called "a dog."

There is a danger that the illustration (figure 26) may be interpreted as a slightly modified version of the mathematical theory of communication. It has to be admitted that it illustrates transference of information with its one-way arrows. The arrows cannot even be two-headed since there is usually no interaction between the designer and the user. But the essential difference between this illustration and the process models of communication is that it is not a question of an attempt to copy a meaning from the consciousness of the sender (designer) to the consciousness of the recipient (user). The illustration and the underlying idea emphasize the role of the activity of the user when working

---

[1] In the lower drawing of figure 26, objects of figure 28 refer to the mental representations evoked by, say, output elements a and b.

with external information on the basis of internal information. While in process models human qualities are handled as constraints, in this model they are a vital resource. From the perspective of process models it would be a great success in communication if we someday had direct access to the physical structures that store mental representations. From the point of view presented here, it would mean skipping the substance of human communication, active interpretation.

7.3 The design process as communication

Finally, the essence of the process of presenting information in a multimodal format seems to be quite analogous with everyday communication between human beings. Basically, however sophisticated the information technology used to deliver information is, a multimedia product is a collection of messages from a designer to a user. In other words, it is a question of mediated human-human communication. Both *how* and *what* information is presented inevitably reflects the mental life of a designer. Technical devices are tools that enable a designer to express her or himself. The skills in utilizing the available technology are certainly of great importance. But in a more central role are general communication skills. How could a designer who does clumsy writing with a typewriter or with a pen, or always becomes misunderstood because of his shortcomings in verbal communication, write fluent and understandable text with a multimedia development application? Or how could a designer who took totally unimaginative and boring photos on his family's holiday in Hawaii suddenly take fascinating and inspiriting photos only because they are taken for a multimedia product?

New information technology contributes to communication by providing a powerful way to present huge amounts of information. The central question is, for what reason is the information presented. There are at least three possibilities:

1. Information is presented for the senses of a user. A sound device is used because the user has ears. A VDU is used because of the existence of eyes. The amount and the format of presented information depend on the content of the presentation and the capacity of the devices. Technical expertise is central.
2. Information is presented for the cognition of a user. Detailed information about the cognitive system is utilized in order to make the presentation effective. The amount of information is rationed out and encoded on the basis of the cognitive models used.
3. Information is presented for the consciousness of a user. Information presentation is seen as a challenge to express meanings, many of which cannot be directly presented with the devices available. The active, interpretative role of the user is therefore essential. Skills in expressing oneself and communication are required.

The classification above should not be seen as a trivial hierarchy of the values of different approaches. Each of these may result in failure or success. It may even be difficult to say on the basis of a completed application which – if any – of these approaches has been applied. For example, the fourth product of the experiment (MS Musical Instruments) may be a result of any of these. It is possible that all the information available is just encoded in the most obvious form in an attempt to use the devices effectively (approach 1). Or there might have been some underlying cognitive model according to which different elements were combined impressively (approach 2). Or, as well, there might have been some creative individual in the application development group who has successfully conducted the production according to his or her idea about the nature of musical instruments (approach 3). The result is, as could be seen, successful, regardless of the approach. The second product (Virtual ear), did not lack creativity and aesthetic values. It can be seen that the designers wished to express more than just physical facts, and the approach is nearest to the third one. But this certainly human approach did not automatically result in high quality. In this case, the designers presumably lacked the expertise that is central to the first two approaches: First, the available technology was not used too skillfully. Second, cognitive structures were not taken into account when combining the elements. Neat details did not compensate for a clumsy whole. The third node (MS Encarta) serves as a good example of a product in the production of which there seem to have been resources available to pay attention to both technical and cognitive aspects (approaches 1 and 2). The somewhat boring outcome may have resulted from pressure to create a very pertinent, informative, and respectable encyclopedia. The first product (The Way Things Work) is difficult to analyze in a similar manner. For all the subjects of the experiment, using this product was more or less a question of "warming up." Therefore, it is not fair to handle it with equal criteria. The product seemed to have successfully combined different points of view; physical facts were presented in an easily understandable form, in many cases by skillfully combining different elements. Humor and aesthetics were widely used. Perhaps a certain ambiguity of the structure and the resulting difficulties in the navigation spoiled the chance to inspire the subjects.

These three approaches should be seen to complement each other. Each of them should be taken into consideration in order to design a high-quality multimedia product. The more technical (approach 1) and psychological (approach 2) expertise as well as creativity and communication skills (approach 3) available in a multimedia project, the better. However, a phrase like "high-quality multimedia product" is not appropriate here since the objective of the current work is not to evaluate actual products. As could be seen in the experiment, an evaluative attitude is quite obvious and unavoidable when analyzing existing products. But the evaluations, since performed in an intuitive manner and without explicit criteria, are very subjective and all-inclusive. Even in the analysis of the subjects' work, an evaluative attitude can be seen in the tone of related sentences. Products that inspired the subjects were drawn apart from the insipid ones.

The prioritization of the analyzed products according to their quality is inappropriate, not only because the subjects were not supposed to evaluate, but also because each multimedia product can only be evaluated in the light of its own objectives. The criteria that are common to all multimedia products are few. For example, as obvious a criterion as increasing the motivation of a user can be complicated; the motivation should be directed to the primary task. If a multimedia product is only a tool for performing some other task, too inspiring a multimedia product can be harmful. Another concern is that motivation can be reached with numerous means. Cheap, loose tricks can entertain for a while. Most of these tricks are probably used to impress a customer in a commercial demonstration. In the long run, the means that are used for motivating should be firmly connected to the content. Perhaps the humor in the first product that was analyzed (The Way Things Work) was found loose or inappropriate concerning the content. The relationships between the humorous messages and the messages that concerned the functions of an airship were mainly *distinct*. The results might have been better if the designer would have tried to make the relationships more *redundant*, thus supporting the creation of knowledge concerning the content on multiple levels.

7.4 The user as an active knowledge constructor

In figure 26, the shapes that signify real world or virtual objects and the resulting mental representations differ from each other. In the early process models of communication this would have been interpreted as a result of noise – disorder or interference in the communication channel. In those models, the target was to reproduce the original meaning. The activity of the receiver meant, in that view, preparedness to receive information units. In the current view, it is admitted that meaning can never be transferred from the consciousness of one person to the consciousness of another person. This is not a constraint or an undesirable fact with which we simply have to cope. Rather, this illustrates the creativity and activity of a human being. The contribution of the current work to design is therefore *not* to support *effective information transfer*. Instead, the target is to *present information in a form that is applicable for the users' needs*. There is an important difference between these two approaches. In the first one, the purpose of the communication via multimedia is to influence and control; multimedia presentation is expected to cause more or less permanent changes in the mental structures of the user. The expertise on human mental life is used to perform these changes effectively. In this approach, redundancy, in the sense it is presented in the current work, is one way of having a stronger effect on the user. On the contrary, the second approach leads to a more user-centered view. Information is *provided* for the user, who either uses it or not in his or her constant knowledge construction process.

The approach that stresses the activity of the user might result in overflowing user-interfaces. As much information as possible is presented, and the user is expected to choose the appropriate pieces. The approach in which the user is a passive receiver of information, may, in turn, result in much clearer user-interfaces since the attention of the user is then easier to control. However, it should not be thought that this is necessarily the case and that the relying on the activity of the user results in confused user-interfaces. If the user-interface is filled up with elements, can we assume that all the information that is coded in those elements is really presented for the user? Referring to the classification on page 122, it can be said that overflowing user-interface information is presented to the senses of the user. As this is seldom enough, we have to turn attention from the quantity of output elements to their quality. The designer should not assume that a certain piece of information is presented for the user when it is coded as a user-interface element since other elements can drown it out.[1] The key is to piece the whole user-interface together and to analyze each single element in relation to each other and the whole.

An alternative viewpoint to active learning and overflowing user-interfaces is to analyze a single node of a multimedia product as a hyper-space. For example, a node that has plenty of screen-elements that compete for the attention of the user with each other functions like hypermedia. In this approach, single screen-elements are nodes. The user focuses her gaze on one element and receives information. She has thus actively chosen the "node." Some cue, internal or external, makes her choose the second one. A screen-element can be designed so that it includes a cue to the succeeding element. These cues represent links in this approach. Thus, the number of elements does not necessarily define whether a node is clear or confused. Even a screen with a huge number of elements may be usable when the design supports construction of meaningful structures by providing appropriate cues to "gaze-navigation," relying on the user's intention to actively construct meanings. This challenges especially those who are responsible for the lay-out of a screen.

7.5 Assessments of the significance of the work and guidelines for future study

If research work is divided into applied, theoretical and meta-theoretical research, the focus of this thesis is on the theoretical level. Meta-theoretical aspects are taken into consideration, e.g., when discussing the bases of different approaches. The underlying pursuit toward applicability as well as the extremely practical needs that have generated this work indicate an applied approach. However, since neither the objective nor the result was detailed design recommendations, the main result of the work can be characterized as theoretical. However, the value of the theoretical work should in this case be assessed according to its applicability.

As reported, the proposed model of the conceptual framework proved understandable and – in many cases – inspiring. Some subjects were able to apply it at a much higher conceptual level than they were trained for. This is a promising result concerning the applicability of the framework. Additionally, the proposed model can be interpreted on the cognitive level within the IP-

1 Kohda (1991) discussed the problem when outlining a "semi-meaningful interface."

paradigm. Therefore, theories and models of attention and other IP-oriented approaches can be utilized with the proposed model of redundancy when analyzing multimodal information presentation. However, the object of the analysis in the experiment was a concrete set of presentation elements. In the creation of something new, e.g. when designing a user-interface, the origin is a highly abstract mental object that first forms a mental representation. The proposed multidimensional analysis of the relationships between individual elements in the creation stage requires firm, intrinsic principles. In order to find out how this would work in practice, an investigation of a real multimedia project would be needed. Participants in this project should be guided from the beginning to analyze their own work according to the proposed criteria. After that, the proposed principles could be evaluated according to their applicability in a real case.

Even if the primary environment of the proposed framework is multimedia, due to the theoretical nature of the model, it can be applied much more broadly. Whenever considering information presentation, regardless of the technological environment, the proposed framework can be used in analysis. Actually, concerning communication studies, conceptual analysis of redundancy is specifically called for (Hanson, 1992).

The main constraint of the applicability of the current work is, however, the constraint of science. When trying to reach better and better quality in multimodal presentations, analysis is an applicable approach to a certain limit. The limit can be localized in a taxonomy in which multimodal presentations are divided into three classes according to their quality:

1. In the lowest class are the presentations that have identifiable defects. The defects have been indicated by the means of science, e.g., some IP-model. For example, if the user is exposed to two verbal tasks simultaneously and the applied cognitive model assumes that linguistic processes are never parallel, the defect is identified.
2. The products of the intermediate class have utilized the available cognitive models religiously. Therefore, no defects can be shown. The products have been made "correctly." However, the result *might* feel blunt, perhaps mechanical and flat.
3. At this stage the limits of science have been reached. From now on, on the way to the highest class, human creative resources that are – at least so far – beyond the domain of science, have to be deployed. It is difficult or impossible to put into words the qualities that are taken into consideration at this stage, and therefore those qualities cannot be completely analyzed.

The classification above does not mean that analysis is not a way to high-quality design. Instead, the classification stresses the importance of expertise in cognition in design; without that expertise a multimedia project is not likely to attain its primary goals. However, the use of the current models does not automatically lead to success. It has to be kept in mind that cognitive models are only models, not reproductions of reality. They are based on statistical analyses, thus describing averages, never a single case.

Multimodal presentation in a multimedia environment is an extremely powerful way of conveying meanings. But, due to its complexity, it is also resource consuming. Investment in multimodal presentation is worthwhile if it enhances communication. According to the current work, multimodal redundant presentations are likely to have that property.

The proposed viewpoint with a conceptual environment concerning the relationships among output elements can also be seen as a promising tool to enhance communication within a multimedia project group. Redundancy is not only an *appropriate* concept concerning information presentation. It is also easy to be adopted and can therefore serve as a unifying link between different views in an interdisciplinary multimedia project group. This would make it possible to get more benefit from the synergy within the group, thus engaging more human resources and resulting in better quality of the actual product.

YHTEENVETO

Käsitteistö luo näkökulman käyttöliittymäanalyysiin

Tietokoneiden käyttöliittymiin kohdistuva mielenkiinto sovelluskehityksessä on jatkuvasti kasvamassa. Aikana, jolloin tietokoneet kuuluivat vain pienen teknisesti orientoituneen asiantuntijapiirin arkeen, tärkeintä oli saada kone laskemaan tehokkaasti. Käyttäjän ja koneen vuorovaikutuksen laatu ei ollut keskeinen kysymys. Nyt, kun tietokoneet ovat tulleet myös tietotekniikkaan perehtymättömien ihmisten työhön ja vapaa-aikaan, yhä tärkeämmäksi ovat nousseet ihmisen ja koneen vuorovaikutukseen liittyvät ongelmat. Kun toisaalta laskentanopeuden kasvu koneissa on ollut huima, osa koneen tehosta on voitu kanavoida käyttöliittymän parantamiseen itse perustoimintojen siitä kärsimättä.

Käyttöliittymiä lähestytään tässä tutkimuksessa käyttöliittymän eri elementtien välisiä suhteita analysoiden. Keskeinen käsite on redundanssi, jolla viitataan suhteeseen, jossa kaksi tai useampia käyttöliittymäelementtejä ovat välittämänsä viestin suhteen identtisiä; samansisältöinen viesti lähetetään esimerkiksi sekä kuvan että tekstin muodossa. Tutkimustehtävänä oli ensinnäkin määritellä redundanssi siten, että se voisi toimia käyttöliittymäsuunnittelun tukena. Koska tarkastelun kohteena on kahden käyttöliittymäelementin tai – viestinnän näkökulmasta – kahden viestin välinen suhde, toisena tutkimustehtävänä oli löytää teoreettinen viitekehys samanaikaisten viestien vastaanottamisen problematiikkaan. Perinteisessä ihmisen ja koneen vuorovaikutuksen kehittämisessä ja tutkimisessa ihmisen ominaisuuksia on mallinnettu kognitiivisen psykologian IP-malleilla (information processing). Näiden mallien metaforana on tietokone: ihmisen kognitio samastetaan sähköiseen informaation prosessointiin. Lähestymistapa helpottaa informaatioyksiköiden kulun mallintamista ihmisen ja koneen välillä. Lähestymistavan ongelmana on kuitenkin ihmisen roolin typistäminen. Tässä tutkimuksessa ihmisen ja koneen vuorovaikutukseen on otettu IP-malleja laajempi näkökulma. IP-malleja on kuitenkin käytetty analyysissä niiltä osin kuin ne tukevat tutkimusongelmien selvittämistä.

Lähdettäessä määrittelemään redundanssia käyttöliittymäsuunnittelun tarpeisiin piti ottaa huomioon se, että käsite on jo entuudestaan käytössä monissa eri konteksteissa. Koska sillä on eri konteksteissa hyvinkin erilaisia, jopa ristiriitaisia konnotaatioita, käsitteen nykyistä käyttöä on aluksi kartoitettu. Hyvin laajasti redundanssi-termiä käytetään tarkoituksessa, jonka termi on saanut informaatioteoriassa. Siinä redundanssilla viitataan tarpeettomaan ylimäärään. Suomenkielisenä vastineena käytetään silloin usein sanaa *toiste*. Redundanssi on jotakin, mistä pyritään pääsemään eroon. Tälle, hyvin kielteiselle sävylle aivan päinvastainen sävy on siinä redundanssin käytössä, jota esiintyy joukkoviestinnän tutkimuksen yhteydessä. Erityisesti televisioviestinnässä redundanssi on tärkeä suunnitteluperiaate. Siinä yhteydessä redundanssi esitetään tarpeellisena, jopa välttämättömänä ominaisuutena rakennettaessa mielekkäitä viestikokonaisuuksia kompleksiseen viestintäympäristöön.

Ensimmäinen tutkimustehtävä, redundanssin käyttöliittymäanalyysiä tukeva määrittely, on toteutettu kuvaamalla se käsiteympäristö, johon redundanssi tässä käyttötarkoituksessa kuuluu. Tässä ympäristössä redundanssi asettui teoreettiseksi ääritapaukseksi dimensiolla, joka kuvaa kahden viestin herättämien merkitysten samankaltaisuutta. Täydellistä redundanssia ei tämän määrittelyn mukaan tavata todellisissa tapauksissa; kahdella viestillä on aina, viimeistään subjektiivisella tasolla, eroavaisuuksia. Todellisuudessa ei siis voidakaan puhua redundanssista tai sen puutteesta, vaan redundanssin asteesta. Toisena ääripäänä samalla dimensiolla on *distinct*-suhde, jossa viesteillä ei ole mitään yhteistä. Dimension puolivälissä on toinen dimensio, jonka ääripäihin sijoittuvat *complementing*- ja *substituting*-suhteet. Näissä kaksi viestiä täydentävät toisiaan: edellinen viestin vastaanottajan odotusten mukaisesti ja jälkimmäinen odotusten vastaisesti.

Kehitetyn käsitteistön käyttökelpoisuutta testattiin empiirisesti käyttöliittymäanalyysitehtävässä, johon osallistui 14 multimedian suunnittelukurssin opiskelijaa. Opiskelijoille opetettiin aluksi käsitteistö ja sen käyttö. Sen jälkeen he analysoivat pareittain neljää annettua multimediakäyttöliittymää. Kuvaruudun ja kaiuttimien kautta tuleva informaatio tallennettiin istuntojen aikana videonauhalle, jonka ääniraidalle tuli myös opiskelijoiden keskustelu. Koehenkilöt täyttivät istunnon aikana myös analyysin tueksi laadittua lomaketta. Kerätty aineisto on analysoitu fenomenografisen metodin mukaan: koehenkilöiden käsityksiä esitellyistä käsitteistä on analysoitu ja luokiteltu, ja huomiota on kiinnitetty myös käsitysten muuttumiseen ja kypsymiseen istunnon aikana.

Keskeinen empiirisen osuuden tulos oli, että käsitteistö osoittautui helposti omaksuttavaksi ja että käsitteistön käyttö tuki käyttöliittymäanalyysiä. Käsitteistö ja sen avaama näkökulma paljasti käyttöliittymistä piirteitä, joita olisi muuten ollut vaikea havaita ja verbalisoida. Yllättävin piirre tutkimusaineistossa oli kuitenkin se säännönmukaisuus, jolla käytetty käsitteistö kypsyi parituntisten istuntojen aikana: Redundanteiksi luokiteltiin aluksi ilmiasultaan läheisesti toisiaan muistuttavien käyttöliittymäelementtien suhteet. Pian kuitenkin siirryttiin viestien tulkinnassa uudelle tasolle ja kyettiin tulkitsemaan kahden, eri koodijärjestelmän mukaisen elementin merkityksiä verbaalisti ja voitiin näin verrata merkitysten redundanssin astetta. Seitsemästä parista kaksi eteni vielä pidemmälle. Näissä koehenkilöt luokittelivat suhteeltaan redundanteiksi joitakin sellaisia elementtejä, joiden merkityksien yhtäläisyyksiä he eivät kyenneet verbalisoimaan, mutta kokivat ne muuten niin voimakkaina, että halusivat kirjata suhteet redundanteiksi. Tämä havainto johti siihen, että syntyi tarve kehittää koko käsiteympäristöä ja sen taustalla olevaa ajattelua moniulotteisempaan suuntaan. Kun alkuperäinen ajatus luokittelusta oli mekaaninen tehtävä, jossa käyttöliittymäelementit tulkitaan verbaalisti ja verrataan näiden verbaalisten esitysmuotojen sisältöä, tutkimusaineisto antoikin aihetta laajentaa tulkintaa tasoille, joita on vaikeaa tai mahdotonta verbalisoida mutta jotka saattavat olla jopa merkityksellisempiä kuin verbalisoitavissa oleva asiasisältö.

Koska redundantit käyttöliittymäelementtikombinaatiot merkitsevät käytännössä usean samanaikaisen viestin lähettämistä käyttäjälle, työssä on otettu tarkastelun kohteeksi myös ihmisen kyky suorittaa samanaikaisesti useampaa kuin yhtä informaation prosessointiin liittyvää tehtävää. Jotta kannattaisi edetä miettimään redundanttien viestien hyötyä, täytyy ensin yrittää ymmärtää, mitä haittaa monen viestin samanaikaisuudesta voisi olemassa olevan tutkimustiedon perusteella olla. Toinen tutkimustehtävä olikin sellaisen teoreettisen viitekehyksen löytäminen, jonka sisällä monen samanaikaisen tehtävän esittämiseen liittyviä ongelmia voi analysoida. Esille otetut mallit ja teoriat analysoidaan työssä redundanttien viestien käsittelyn näkökulmasta. Pääosa käsitellyistä malleista liittyy pitkään attentiotutkimusperinteeseen, joka on noussut ajankohtaiseksi usein juuri ihmisen ja kompleksisen teknisen ympäristön vuorovaikutuksen ongelmia tutkittaessa. Analyysin kohteena olleet kognitiiviset mallit osoittautuivat käyttökelpoisiksi multimodaalisten käyttöliittymien suunnittelun tarpeita silmällä pitäen. Vaikka useimmat mallit olivat vanhoja ja niiden alkuperäiset sovellusympäristöt aivan muuta mitä tällä hetkellä ymmärretään multimedialla, samoja havaintoja ihmisestä informaation prosessoijana voitiin soveltaa myös nykyaikaiseen tekniseen ympäristöön.

Aiempi aiheeseen liittyvä tutkimus on keskittynyt etsimään esimerkiksi muistin tai ongelmanratkaisun kannalta tarkoituksenmukaisia keinoja informaation esittämiseen, jolloin jossain tapauksissa redundantit viestit on todettu tehokkaiksi. Näissä tutkimuksissa on kuitenkin helposti nähtävissä liiallista yksinkertaistamista. Esimerkiksi muistamista koskevissa koeasetelmissa on asetettu vastakkain tapaukset, joissa kuva ja teksti ovat joko samansisältöisiä tai erisisältöisiä. Tämän tutkimuksen valossa näin jyrkkään luokitteluun ei ole perusteita. Redundanssia ei voi pitää myöskään itsetarkoituksena. Sen sijaan se voi toimia käsitteellisenä kiinnekohtana suunniteltaessa monitasoisia virtuaaliobjekteja, joita ei voida pitää enää osiensa summana vaan sovellussuunnittelijan mielikuvien monitasoisina ilmentyminä. Esimerkiksi virtuaalinen koira ei ole mikä tahansa koiran kuvan ja äänen yhdistelmä, vaan mielikuva, joka on herätetty juuri tietyn kuvan ja tietyn äänen samanaikaisella esittämisellä. Keskeistä virtuaaliobjektien luomisessa onkin sovellussuunnittelijan kyky ilmaista itseään multimediaympäristössä.

Monialaisissa sovelluskehitysprojekteissa törmätään helposti kommunikaatiovaikeuksiin, jotka johtuvat eri alojen edustajien joskus hyvinkin erilaisista näkökulmista ja käsitteistöistä. Koska tutkimuksessa kehitetty käsitteistö ei ole sidottu minkään tieteenalan erikoissanastoon tai lähestymistapaan, sen voidaan katsoa palvelevan monialaisia sovelluskehitysprojekteja tarjoamalla välineitä ryhmän sisäiseen kommunikointiin. Koska käsitteistö koskee erittäin oleellisia käyttöliittymään liittyviä ominaisuuksia, käsitteistön käyttö suuntaa myös sitä käyttävän ryhmän tarkkaavaisuutta tarkoituksenmukaisesti.

REFERENCES

Allport, D. A., Antonis, B., & Reynolds, P. 1972. On the division of attention: a disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology 24*(2), 225 - 235.

Anderson, J. R. 1995. *Cognitive psychology and its implications*. New York: W. H. Freeman and Company.

Bahrick, H. P. & Gharrity, K. 1976. Interaction among pictorial components in the recall of picture captions. *Journal of Experimental Psychology: Human Learning and Memory 2*(2), 103 - 111.

Barker, P. & King, T. 1993. Evaluating interactive multimedia courseware - a methodology. *Computers & Education 21*(4), 307 - 319.

Barnard, P. J. 1985. Interacting cognitive subsystems: a psycholinguistic approach to short-term memory. In A. Ellis. *Progress in the psychology of language, vol 2*, chapter 6. London: Lawrence Erlbaum, 197 - 258.

Barnard, P. J. & May, J. 1993(a). Cognitive modelling for user requirements. In P. F. Byerley, P. J. Barnard, & J. May (Eds.), *Computers, communication and usability: Design issues, research and methods for integrated services*. (North Holland Series in Telecommunication). Amsterdam: Elsevier, 101 - 146.

Barnard, P. J. & May, J. 1993(b). Real time blending of data streams: A key problem for the cognitive modelling of user behaviour with multimodal systems. Esprit Basic Research Action 7040: The Amodeus Project, document UM/WP10.

Barnard, P. J., May, J., & Blandford, A. 1992. To model or not to model? That is the question. Esprit Basic Research Action 7040: The Amodeus Project, document UM/WP2.

Basil, M. D. 1994. Multiple resource theory I: Application to television viewing. *Communication Research 21*(2), 177 - 207.

Bennett, J. L. 1972. The user interface in interactive systems. *Annual Review of Information Science and Technology 7*, 159 - 196.

Bennett, J. L. 1979. The commercial impact of usability in interactive systems. In B. Shackel (Ed.), *Man-computer communication. Infotech State-of-the Art, Vol. 2*. Maidenhead: Infotech International, 1 - 17.

Bennett, J. L. 1984. Managing to meet usability requirements. Establishing and meeting software development goals. In J. Bennett, D. Case, J. Sandelin, & M. Smith (Eds.), *Visual display terminals*. Englewood Cliffs (NJ): Prentice Hall, 161 - 191.

Benson, P. J. 1993. Problems in Picturing Text. In *Art, science & visual literacy*. Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 107 - 117. (Eric Document Reproduction Service No. ED 373 717)

Bernsen, N. O. 1994. Why are analogue graphics and natural language both needed in HCI? In F. Paterno (Ed.), *Design, specification and verification of* *interactive systems*. Proceedings of the Eurographics workshop, Carrara, Italy, 8 - 10 June 1994, 165 - 179.

Bernsen, N. O. 1995. A toolbox of output modalities: Representing output information in multimodal interfaces. Esprit Basic Research Action 7040: The Amodeus Project, document TM/WP21.

Biederman, I., Glass, A. L., & Stacy, E. W. 1973. Searching for objects in real-world scenes. *Journal of Experimental Psychology 97*(1), 22 - 27.

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. 1974. On the information extracted from a glance at a scene. *Journal of Experimental Psychology 103*(3), 597 - 600.

Blattner, M. M. & Glinert, E. P. 1996. Multimodal integration. *IEEE Multimedia 3*(4), 14 - 24.

Bloom, B. S. 1972. Taxonomy of educational objectives: The classification of educational goals. New York: McKay.

Braden, R. A. 1992. Examining Visual Verbal Relationships. In *Art, science & visual literacy*. Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 229 - 238. (Eric Document Reproduction Service No. ED 373 717)

Brand, J. 1971. Classification without identification in visual search. *Quarterly Journal of Experimental Psychology 23*(2), 178 - 186.

Broadbent, D. E. 1958. *Perception and communication*. London: Pergamon.

Brown, D. W. & Schneider, S. D. 1992. Young learners' reactions to problem solving contrasted by distinctly divergent computer interfaces. *Journal of Computing in Childhood Education 3* (3-4), 335-347.

Brown, J. R. & Cunningham, S. 1989. *Programming the user interface: Principles and examples*. New York: John Wiley & Sons.

Bruce, V. 1979. Searching for politicians: an information-processing approach to face recognition. *Quarterly Journal of Experimental Psychology 31*(3), 373 - 395.

Cahill, M. C. & Carter, R. C. 1976. Color code size for searching displays of different density. *Human Factors 18*(3), 273 - 280.

Card, S. K., Moran, T. P., & Newell, A. 1983. *The psychology of human-computer interaction*. Hillsdale (NJ): Lawrence Erlbaum Associates.

Chalmers, D. J. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies 2*(3), 200 - 219.

Chalmers, D. J. 1996. *The conscious mind. In search of a fundamental theory*. New York: Oxford University Press.

Clark, R. E. & Craig, T. G. 1992. Research and theory on multi-media learning effects. In M. Giardina (Ed.), *Interactive multimedia learning environments: human factors and technical considerations on design issues*. NATO ASI Series F: Computer and System Sciences, 93. Heidelberg: Springer-Verlag, 19 - 30.

Collins, A. M. & Loftus, E. F. 1975. A spreading-activation theory of semantic processing. *Psychological Review 82*(6), 407 - 428.

Cope, P. & Simmons, M. 1994. Some effects of limited feedback on performance and problem-solving strategy in a Logo microworld. *Journal of Educational Psychology 86*(3), 368 - 379.

Corcoran, D. W. J. & Jackson, A. 1979. Flexibility in the choice of distinctive features in visual search with blocked and random designs. *Perception 8*, 629 - 633.

Craik, F. I. M. & Lockhart, R. S. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior 11*(6), 671 - 684.

Deutsch, J. A. & Deutsch, D. 1963. Attention: Some theoretical considerations. *Psychological Review 70*(1), 80 - 90.

Drew, D. G. & Grimes, T. 1985. The effect of audio-visual redundancy on audio and video recall in television news. Paper presented at the annual meeting of the Association for Education in Journalism and Mass Communication (68th, Memphis, TN, August 3 - 6, 1985). (Eric Document Reproduction Service No. ED 257 130)

Dyck, J. L. 1995. Problem solving by novice macintosh users: The effects of animated, self-paced written, and no instruction. *Journal of Educational Computing Research 12*(1), 29 - 49.

Edwards, A. D. N. 1992. Redundancy and adaptability. In A. D. N. Edwards & S. Holland (Eds.), *Multimedia interface design in education*. NATO ASI Series F: Computer and system sciences, 76. Heidelberg: Springer-Verlag, 145 - 155.

Egeth, H. & Dagenbach, D. 1991. Parallel versus serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance 17*(2), 551 - 560.

Egeth, H., Atkinson, J., Gilmore, G., & Marcus, N. 1973. Factors affecting processing mode in visual search. *Perception & Psychophysics 13*(3), 394 - 402.

Ekman, P. & Friesen, W. V. 1969. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica 1*(1), 49 - 98.

Felfoldy, G. L. & Garner, W. R. 1971. The effects on speeded classification of implicit and explicit instructions regarding redundant dimensions. *Perception & Psychophysics 9*(3A), 289 - 292.

Findahl, O. 1971. The effect of visual illustrations upon perception and retention of news programmes. Stockholm: Swedish Broadcasting Corporation. (Eric Document Reproduction Service No. ED 054 631)

Fisher, S. S. 1990. Virtual interface environments. In B. Laurel (Ed.), *The art of human computer interface design*. Reading, MA: Addison Wesley, 423 - 438.

Fiske, J. 1990. *Introduction to communication studies*. London: Routledge. 2nd ed.

Fletcher, B. & Rabbitt, P. M. A. 1978. The changing pattern of perceptual analytic strategies and responce selection with practice in a two-choice reaction time task. *Quarterly Journal of Experimental Psychology 30*(3), 417 - 427.

Frisby, J. 1979. *Seeing: Illusion, brain and mind*. Oxford: Oxford University Press.

Galbreath, J. 1992. The educational buzzword of the 1990's: Multimedia, or is it hypermedia, or interactive multimedia, or...? *Educational Technology 32*(4), 15 - 19.

Galitz, W. O. 1993. *User-interface screen design*. Wellesley, MA: QED.

Gardner, H. G. 1985. *The mind's new science*. New York: Howard Gardner.

Garner, W. R. & Felfoldy, G. L. 1970. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology 1*(3), 225 - 241.

Gibson, J. J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Glaser, G. & Strauss, A. 1968. The discovery of grounded theory: Strategies for qualitative research. London: Weidenfeld and Nicolson.

Gleitman, H. & Jonides, J. 1976. The cost of categorization in visual search: Incomplete processing of targets and field items. *Perception & Psychophysics 20*(4), 281 - 288.

Green, B. F. & Anderson, L. K. 1956. Color coding in a visual search task. *Journal of Experimental Psychology 51*(1), 19 - 24.

Griffin, R. E. & Gibbs, W. J. 1992. International icon symbols: How well are these symbols understood? In *Art, science & visual literacy*. Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 132 - 141. (Eric Document Reproduction Service No. ED 373 717)

Hanson, L. 1992. The concept of redundancy in television learning research: Questions of meaning. *International Journal of Instructional Media 19*(1), 7 - 13.

Hanson, L. 1993. Perceptions of between-channel redundancy in television messages. In *Art, science & visual literacy*. Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 174 - 184. (Eric Document Reproduction Service No. ED 373 717)

Haring, M. J. & Fry, M. A. 1979. Effect of pictures on children's comprehension of written text. *Educational Communication and Technology Journal 27*(3), 185 - 190.

Hazari, S. I. & Reaves, R. R. 1994. Student preferences toward microcomputer user interfaces. *Computers & Education 22*(3), 225 - 229.

Hodes, C. L. 1992. Effect of visual mental imagery on speed and accuracy of information retrieval. In *Art, science & visual literacy*. Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 252 - 259. (Eric Document Reproduction Service No. ED 373 717)

Holst, S., Churchill, E., & Gilmore, D. 1997. Transporting honeybears: A cognitive analysis of the effects of interface manipulation style on a constraint-based planning task. In M. J. Smith, G. Salvendy, & R. J. Koubek (Eds.), *Design of computing systems: social and ergonomic considerations*.

Advances in Human Factors/Ergonomics 21B. Proceedings of HCI International '97, San Francisco (CA), Aug. 24-29 1997. Amsterdam: Elsevier, 169 – 172.

Hsia, H. J. 1971. The information processing capacity of modality and channel performance. *AV Communication Review 19*(1), 51 - 75.

Hsia, H. J. 1977. Redundancy: Is it the lost key to better communication? *AV Communication Review 25*(1), 63 - 85.

Ingling, N. W. 1972. Categorization: A mechanism for rapid information processing. *Journal of Experimental Psychology 94*(3), 239 - 243.

Jackendoff, R. 1989. Consciousness and the computational mind. Cambridge (MA): MIT Press. 2nd pr.

Jackson, D. F., Edwards, B. J., & Berger, C. F. 1993. The design of software tools for meaningful learning by experience: Flexibility and feedback. *Journal of Educational Computing Research 9*(3), 413 - 443.

James, W. 1891. *The principles of psychology*. (Vol 1). London: MacMillan.

Jonides, J. & Gleitman, H. 1976. The benefit of categorization in visual search: Target location without identification. *Perception & Psychophysics 20*(4), 289 - 298.

Kahneman, D. 1973. *Attention and effort*. Englewood Cliffs (NJ): Prentice Hall.

Kahneman, D. & Treisman, A. 1984. Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando (F): Academic Press, 29 - 62.

Kim, S. 1990. Interdisciplinary cooperation. In B. Laurel (Ed.), *The art of human computer interface design*. Reading, MA: Addison Wesley, 31 - 44.

Kleiman, G. M. 1975. Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior 14*(4), 323 - 339.

Kohda, Y. 1991. Toward the semi-meaningful interface beyond the direct manipulation interface. In J. Motiwalla (Ed.), *Multimedia information systems '91*. Proceedings of Jurong, Singapore Jan. 7.-11. 1991, 85 - 95.

Kristofferson, M. W. 1977. The effects of practice with one positive set in a memory scanning task can be completely transferred to a different positive set. *Memory & Cognition 5*(2), 177 - 186.

Kuhn, T. S. 1970. *The structure of scientific revolution*. Chicago (ILL): University of Chicago Press. 2nd ed., enl.

Lang, A. & Friestad, M. 1993. Emotion, hemispheric specialization, and visual and verbal memory for television messages. *Communication Research 20*(5), 647 - 670.

Latham, R. 1995. The dictionary of computer graphics and virtual reality. New York: Springer-Verlag.

Laurel, B. 1990. *The art of human computer interface design*. Reading (MA): Addison Wesley.

Laurel, B. 1991. *Computers as theatre*. Reading (MA): Addison-Wesley.

Levie, W. H. & Lentz, R. 1982. Effects of text illustrations: a review of research. *Educational Communication and Techonology Journal 30*(4), 195 - 232.

Levin, J. R. & Lesgold, A. M. 1978. On pictures in prose. *Educational Communication and Technology Journal 26*(3), 233 - 243.
Longman dictionary of the English language. 1984. London: Longman Group.
Lovie, A. D. 1983. *Attention and behaviourism - fact and fiction*. British Journal of Psychology 74(3), 301 - 310.
Mangun, G. R., Hillyard, S. A., & Luck, S. J. 1992. Electrocortical substrates of visual selective attention. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artifical intelligence, and cognitive neuroscience*. Cambridge (MA): MIT Press, 219 - 243.
Marton, F. 1978. *Describing conceptions of the world around us*. Mölndal: The Institute of Education, University of Göteborg, reports 66.
Marton, F. 1988. Phenomenography: A research approach to investigating different understandings of reality. In R. R. Sherman & R. B. Webb (Eds.), *Qualitative research in education: Focus and methods*. London: Falmer Press, 141 - 161.
Massaro, D. W. & Cowan, N. 1993. Information processing models: microscopes of the mind. *Annual Review of Psychology 44*, 383 - 425.
May, J., Barnard, P. J., & Blandford, A. 1992. Using structural descriptions of interfaces to automate the modelling of user cognition. Esprit Basic Research Action 7040: The Amodeus Project, document UM/WP21.
Mayer, R. E. 1992. The instructive animation: helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology 84*(4), 444 - 452.
Mayer, R. E. & Sims, V. K. 1994. For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology 86*(3), 389 - 401.
Mayes, J. T. 1992. The 'M-Word'. In A. D. N. Edwards & S. Holland (Eds.), *Multimedia interface design in education*. NATO ASI Series F: Computer and system sciences, 76. Heidelberg: Springer-Verlag, 1 - 22.
Maynard, J. 1975. *Dictionary of data processing*. London: Newnes-Butterworths.
Mercadal, D. 1990. *Dictionary of artificial intelligence*. New York: Van Nostrand Reinhold.
Miller, R. B. 1971. Human ease-of-use criteria and their trade-offs. TR00.2185. Poughkeepsie (NY): IBM Corp.
Miller, S. C. 1993. The effects of redundant visual, auditory, and tactile stimuli on a minimum information tracking task. In *Art, science & visual literacy*. Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 143 - 152. (Eric Document Reproduction Service No. ED 363 297)
Moray, N. 1959. Attention in dichotic listening: affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology 11*(1), 56 - 60.
Moray, N. 1975. A data base for theories of selective listening. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. London: Academic Press, 119 - 135.

Moray, N. 1978. The strategic control of information processing. In G. Underwood (Ed.), *Strategies of information processing*. New York: Academic Press, 301 - 328.
Moray, N. 1993. Designing for attention. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent*. Oxford: Clarendon Press, 111 - 134.
Najjar, L. J. 1996. Multimedia information and learning. *Journal of Educational Multimedia and Hypermedia 5*, 129 - 150.
Neisser, U. 1963. Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology 76*, 376 - 385.
Neisser, U. 1967. *Cognitive psychology*. New York: Appleton-Century-Crofts.
Norman, D. A. 1968. Towards a theory of memory and attention. *Psychological Review 75*, 522 - 536.
Paivio, A. 1986. *Mental representations. A dual coding approach*. New York: Oxford University Press.
Paivio, A. 1991. Dual coding theory: retrospect and current status. *Canadian Journal of Psychology 45*(3), 255 - 287.
Paivio, A., Rogers, T. B., & Smythe, P. C. 1968. Why are pictures easier to recall than words? *Psychonomic Science 11*(4), 137 -138.
Palmer, S. E. & Kimchi, R. 1986. The information processing approach to cognition. In T. J. Knapp & L. C. Robertson (Eds.), *Approaches to cognition: contrasts and controversies*. Hillsdale (NJ): Lawrence Erlbaum, 37 - 77.
Paul, S. 1992. Pictorial redundancy, processing time, number of propositions, and comprehensions of compressed speech. *International Journal of Instructional Media 19*(1), 71 - 78.
Peirce, C. S. 1958 / 1966. *Collected papers of Charles Sanders Peirce*. Volume VII Science and philosophy and volume VIII Reviews, correspondence, and bibliography. Edited by Burks, A. W. Cambridge (MA): Belknap Press.
Peled, Z., Peled, E., & Alexander, G. 1992. A taxonomy for computer software adoption policy. *Journal of Educational Computing Research 8*(1), 81 - 100.
Petre, M. & Green, T. R. G. 1990. Is graphical notation really superior to text, or just different? Some claims by logic designers about graphics in notation. Proceedings in the Fifth Conference on Cognitive Ergonomics, Urbino, Italy, September 3 - 6, 1990.
Pirhonen, A. 1997. Meaningful whole or a jumble? The problem of designing a combination of output elements. *Management 2*(8), 62 - 65.
Pirhonen, A. 1998. Analysis of the concept of redundancy concerning the design of multimodal combinations of output-elements. In Werner, B. *APCHI 98*. Proceedings in the 3rd Asia Pacific Computer Human Interaction, July 15 - 17 1998, Shonan Village Center, Japan. Los Alamitos (CA): IEEE, 273 - 278.
Platt, J. R. 1964. Strong inference. *Science 146*(3642), 347 - 353.
Pollack, I. 1963. Speed of classification of words into superordinate categories. *Journal of Verbal Learning and Verbal Behavior 2*(2), 159 - 165.
Popper, K. R. 1959 / 1980. *The logic of scientific discovery*. London: Unwin Hyman.

Prinz, W. 1979. Locus of the effect of specific practice in continuous visual search. *Perception & Psychophysics 25*(2), 137 - 142.

Rabbitt, P. M. A. 1959. Effects of independent variations in stimulus and response probability. *Nature 183*(4669), 1212.

Rabbitt, P. M. A. 1967. Learning to ignore irrelevant information. *The American Journal of Psychology 80*(1), 1 - 13.

Rabbitt, P. M. A. 1968. Repetition effects and signal classification strategies in serial choise-response tasks. *Quarterly Journal of Experimental Psychology 20*(3), 232 - 240.

Rabbitt, P. M. A. 1978. Sorting, categorization, and visual search. In E. C. Carterette & M. Friedman (Eds.), *Handbook of perception 9*. New York: Academic Press, 85 - 134.

Rabbitt, P. M. A. 1981. Cognitive psychology needs models for changes in performance with old age. In J. Long & A. D. Baddeley (Eds.), *Attention and performance IX*. Hillsdale (N J): Lawrence Erlbaum Associates, 555 - 573.

Rabbitt, P. M. A. 1984. The control of attention in visual search. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando (F): Academic Press, 273 - 292.

Rabbitt, P. M. A., Cumming, G., & Vyas, S. 1977. An analysis of visual search: Entropy and sequential effects. In S. Dornic (Ed.), *Attention and performance IV*. Hillsdale (NJ): Lawrence Erlbaum Associates, 363 - 386.

Rabbitt, P. M. A., Cumming, G., & Vyas, S. 1979(a). Improvement, learning and retention of skill at visual search. *Quarterly Journal of Experimental Psychology 31*(3), 441 - 459.

Rabbitt, P. M. A., Cumming, G., & Vyas, S. 1979(b). Modulation of selective attention by sequential effects in visual search tasks. *Quarterly Journal of Experimental Psychology 31*(2), 305 - 317.

Reese, S. D. 1983. Improving audience learning from television news through between-channel redundancy. Paper presented at the annual meeting of the Association for Education in Journalism and Mass Communication (66th, Corvalis, OR, August 6 - 9, 1983). (Eric Document Reproduction Service No. ED 229777)

Resnick, L. B. 1989. *Knowing, learning, and instruction*. Hillsdale (NJ): Lawrence Erlbaum.

Rollins, H. A. & Hendricks, R. 1980. Processing of words presented simultaneously to eye and ear. *Journal of Experimental Psychology: Human Perception and Performance 6*(1), 99 - 109.

Rubin, A. M. 1993. Audience activity and media use. *Communication monographs 60*(1), 98 - 105.

Sack, S. A. & Rice, C. E. 1974. Selectivity, resistance to distraction and shifting as three attentional factors. *Psychological Reports 34*(3), 1003 - 1012.

Saussure, F. de 1983 / 1990. *Course in general linguistics*. Edited by Bally, C. & Sechehaye, A. with the collaboration of Riedlinger, A. First published 1983. London: Duckworth.

Schneider, W & Fisk, A. D. 1982. Degree of consistent training: Improvements in search performance and automatic process development. *Perception & Psychophysics 31*(2), 160 - 168.
Schneider, W. & Shiffrin, R. M. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review 84*(1), 1 - 66.
Schneider, W., Dumais, S. T., & Shiffrin, R. M. 1984. Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando (F): Academic Press, 1 - 28.
Schweickert, R. 1992. Information, time, and the structure of mental events: A twenty-five-year review. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artifical intelligence, and cognitive neuroscience*. Cambridge (MA): MIT Press, 535 - 566.
Severin, W. 1967. The effectiveness of relevant pictures in multiple-channel communications. *AV Communication Review 15*(4), 386 - 401.
Shackel, B. 1984. The concept of usability. In J. Bennett, D. Case, J. Sandelin, & M. Smith (Eds.), *Visual display terminals*. Englewood Cliffs (NJ): Prentice Hall, 45 - 87.
Shaffer, L. H. 1975. Multiple attention in continuous verbal tasks. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. London: Academic Press, 157 - 167.
Shannon, C. E. & Weaver, W. 1949 / 1964. *The mathematical theory of communication*. Urbana: The University of Illinois Press.
Shiffrin, R. M. & Schneider, W. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review 84*(2), 127 - 190.
Shneiderman, B. 1987. Designing the user interface: Strategies for effective human-computer interaction. Reading (MA): Addison-Wesley.
Sperling, G. 1984. A unified theory of attention and signal detection. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando (F): Academic Press, 103 - 182.
Staniland, A. C. 1966. *Patterns of redundancy*. Cambridge: University Press.
Strauss, A. & Corbin, J. 1990. Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park (CA): SAGE.
Swets, J. A. 1984. Mathematical models of attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando (F): Academic Press, 183 - 242.
Taylor, J. 1995. We have the technology. *The Thes (The Times Higher Education Supplement) Internet Service*. http://thesis.newsint.co.uk/SPECIAL/taylor.html.
Tesch, R. 1992. Qualitative research: Analysis types and software tools. New York: Falmer Press.
Titchener, E. B. 1908. Lectures on the elementary psychology of feeling and attention. New York: MacMillan.

Tolhurst, D. 1995. Hypertext, hypermedia, multimedia defined. *Educational Technology 35*(2), 21 - 26.

Townsend, J. T. 1990. Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science 1*(1), 46 - 54.

Treisman, A. M. 1960. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology 12*, 242 - 248.

Treisman, A. M. 1964(a). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology 77*(2), 206 - 219.

Treisman, A. M. 1964(b). The effect of irrelevant material on the efficiency of selective listening. *American Journal of Psychology 77*(4), 533 - 546.

Treisman, A. M. & Davies, A. 1973. Divided attention to ear and eye. In S. Kornblum (Ed.), *Attention and performance IV*. London: Academic Press, 101 - 117.

Treisman, A. M. & Howarth, C. I. 1959. Changes in threshold level produced by a signal preceding or following the threshold stimulus. *Quarterly Journal of Experimental Psychology 11*(3), 129 - 142.

Vacherand-Revel, J. & Bessière, C. 1992. 'Playing' graphics in the design and use of multimedia courseware. In A. D. N. Edwards & S. Holland (Ed.), *Multimedia interface design in education*. NATO ASI Series F: Computer and system sciences, 76. Heidelberg: Springer-Verlag, 57 - 82.

van der Heijden, A. H. D. & Stebbins, S. 1990. The information-processing approach. *Psychological Research 52*(2), 197 - 206.

Vertelney, L. & Booker, S. 1990. Designing the whole-product user interface. In B. Laurel (Ed.), *The art of human computer interface design*. Reading (MA): Addison Wesley, 57 - 63.

Web dictionary of cybernetics and systems. 1997. http://pespmc1.vub.ac.be/ASC/REDUNDANCY.html.

Weik, M. H. 1977. *Standard dictionary of computers and information processing*. Rochelle Park, N.J.: Hayden Book. 2nd edition.

Wickens, C. D. 1976. The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance 2*(1), 1 - 13.

Wickens, C. D. 1980. The structure of attentional resources. In R. Nickerson (Ed.), *Attention and perfomance VIII*. Hillsdale (NJ): Lawrence Erlbaum, 239 - 257.

Wickens, C. D. 1984. Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando (F): Academic Press, 63 - 102.

Wickens, C. D. 1992. *Engineering psychology and human performance*. New York: HarperCollins Publishers. 2nd edition.

Wickens, C. D. & Carswell, C. M. 1997. Information processing. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics*. New York: John Willey & Sons. 2nd ed., 89 - 129.

Willows, D. M. & MacKinnon, G. E. 1973. Selective reading: Attention to the "unattended" lines. *Canadian Journal of Psychology 27*(3), 292 - 304.

APPENDIX 1: Form of the pilot study

Classification of unimodal messages and their connections
application nr____ investigators:____&____

element	li	-li	an	-an	ar	-ar	sta	dyn	gra	sou
1.										
2.										
3.										
4.										
5.										
6.										
7.										
8.										
9.										
10.										

Matrix of connections: code of the quality of connection (<u>D</u>istinction, <u>C</u>omplementing, <u>S</u>ubstitution, <u>E</u>mphasizing, <u>R</u>edundancy)

1	1	2	3	4	5	6	7	8	9
2									
3									
4									
5									
6									
7									
8									
9									
10									

Percepted redundant messages (code =e.g. *2-3*, in which 2 and 3 are unimodal messages)

code	redundant message

Remarks:____
____
____
____

APPENDIX 2: Form of the final experiment

Pair code:__________(Please say this number in microphone also)

Product to be investigated:_________________________

Please classify the most important user-interface elements of this node according to their role:

I=Informative; C=Controlling (node/product); De=Decorative; Di=Distinguishing

	User-interface element	I	C	De	Di
1.					
2.					
3.					
4.					
5.					
6.					
7.					
8.					
9.					
10.					
11.					
12.					
13.					
14.					
15.					
16.					
17.					

Please try to identify some meaningful combinations of two elements and classify the relationship between them. Use the numbers of elements from the table above. E. g. if you recognize that the relationship between elements 3 and 6 is *complementing*, write 3-6 in the first column and mark it as "C" like this:

Elements\Relationhip	D	C	S	E	R
3-6		X			

The letters D, C, S, E and R refer to Distinct, Complementing, Substituting, Emphasizing and Redundant.

Elements\Relationship	D	C	S	E	R

If you recognized redundant (or emphasizing) relationships, please describe the redundant message.

Elements	Redundant message