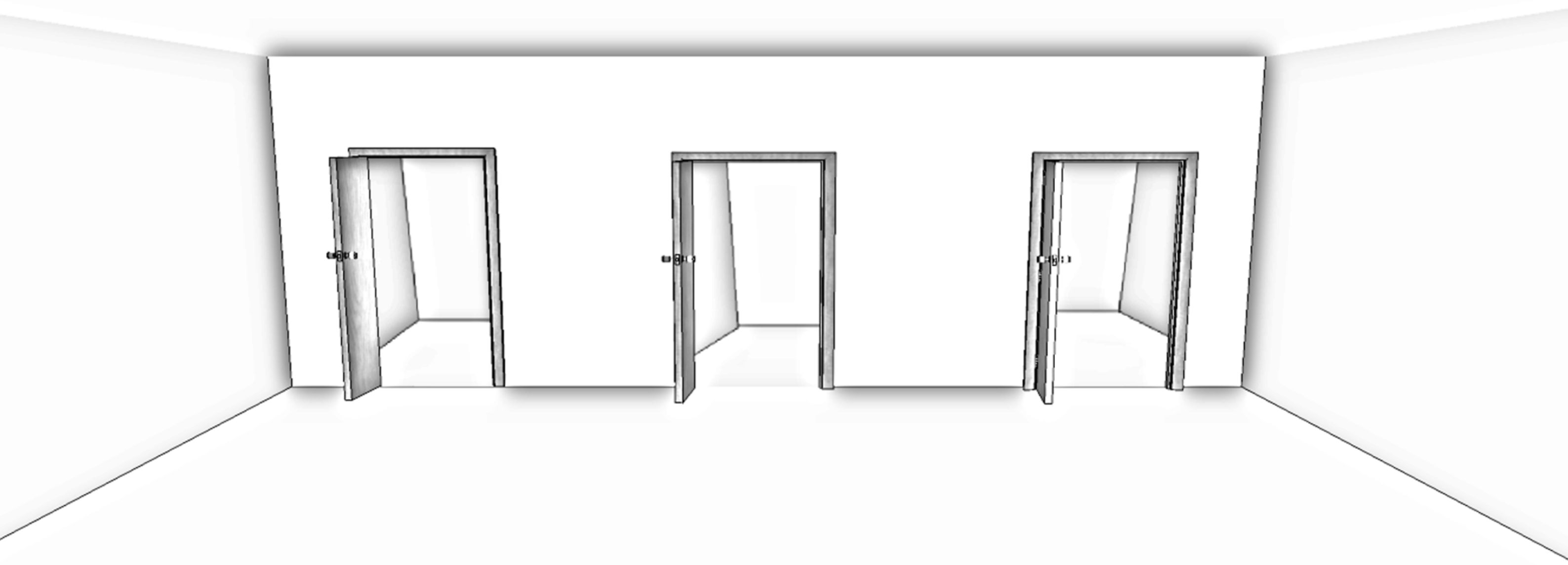
Conceptualizing space

- A study of concepts and the spatial dimension in graphical user interfaces

By Anders K. Møller & Jonas H. Christiansen



Aalborg University
The Faculty of Engineering and Science

Master's Thesis 2012 - GROUP 1073
Engineering Psychology

Aalborg University
Engineering Psychology
Fredrik Bajers Vej 7A
Telephone 96 35 97 31
Fax 98 13 63 93
http://ies.auc.dk

Title:
Conceptualizing space
- A study of concepts and the spatial dimension in graphical user interfaces

Theme:
Master's Thesis
Engineering Psychology

Project period:
February 1st - May 31st 2012

Project group:
1073

Group members:

Anders Kalsgaard Møller

Jonas Høier Christiansen

Supervisor:

Lars Bo Larsen

Copies: 4

Pages: 96

Appendices: 8

Type and number of attachments: DVD, 4

Handed in: May 30th

Abstract:

Current folder systems do not offer enough support for recalling the location of files. Neither do they provide enough information about the context in order for the users to know where the files are located in the system.

The purpose of this study was to investigate if it was possible to create a system with better support for perception of the spatial dimension, and if this enhanced the users' ability to recall objects. It was also investigated if more contextual information about the locations would help the users' awareness of their location in the system.

A prototype system was created which supported perception of the spatial dimension by using a house metaphor. Furthermore, it provided contextual information by using a map. The prototype system was used in an experiment where it was tested if the test subjects were able to find music albums located in the system.

The experiment showed that the test subjects were significantly better at finding music albums in the prototype system compared to a control system. The experiment also showed that the test subjects used the mini-map to get a better overview.

Preface

This study is the result of the Master's thesis written by group 1073 studying Engineering Psychology at the Faculty of Engineering and Science of Aalborg University.

Reading guide

During the paper references occur which refers to the Bibliography at the end of the paper. The Bibliography is set by the Harvard Style of Referencing. Sources are therefore referred to as (author, year). "ibid" is used to refer to the same author or work twice with no other author or work in between. If the page number is different but the author or work remains the same the relevant page number will appear. Sources that are cited consecutively are therefore referred to as (ibid) or (ibid, p. xx). Web pages are referred to as (name of the web page, year) and will likewise refer to the Bibliography at the end of the paper. In the Bibliography books and articles are mentioned by author, title, edition and publisher while web pages will be mentioned by author, year, title, and name of web page.

Figures and charts are numbered in proportion to their chapters. Thereby letting the first figure in chapter 6 be numbered as Figure 6.1. The following figure is numbered as Figure 6.2, etc. This is also applicable for figures in the Appendix, i.e. the figures in Appendix B are numbered Figure B.1, Figure B.2, etc. An explanatory text is written under the relevant figures or charts and these are also available at the end of the report in the List of Figures.

Contents

Chapter 1

Introduction

When the first personal computers started to enter the market you only had very limited storage space and it was primarily used for small programs, games and text files. Besides the use of the storage in computers people were also using tapes and diskettes to store their information.

Today's personal computers have almost unlimited space and we store all kinds of information: Music, video, text files, programs, games etc. Not only do we store different kinds of information on our computers, but the amount of stored information has also increased. This is probably because we do not have to worry about disk space anymore but also due to more and more real life processes have entered the computer.

In order to handle the amount of information people store on their personal computers a categorization system is needed. Such a categorization system existed in MS-Dos. In MS-Dos the files were placed in different virtual locations. The location were only visible to the user by the name of the direction. In order to find your files you typed the direction names. Around the same time the graphical user interface were introduced and with it the desktop metaphor. In the desktop metaphor you were able to sort your files in different folders located on a virtual desktop. In this virtual desktop the files actually have a visual location in one of the folders on the desktop. You find them by clicking on the folder it is located in. The desktop user interface created a user interface where you could actually see and interact with the files. You could also physically move the files around between the different folders.

Since the introduction of the desktop metaphor the direction in categorization of files in graphic user interfaces have not changed much. Of course some changes have occurred like better graphic, more search functions, and prelabelled folders.

Even though some changes have occurred it still has its roots in the desktop metaphor. Studies have shown that people have problems finding their files on computers (Barreau & Nardi 1995). This problem has probably emerged because the amount of data has grown over the limit of our human memory. Even though the desktop metaphor provides some support for users to recall the position of files it does not provide enough. If we remove components like the task bar and simply look at the desktop it is simply a flat surface from which you can enter a folder structure. The problem is not the desktop, it is the folder system. As mentioned above the folder system consists of files categorized in identical folders within folders. The only thing separating the folders from each other is small differences in the look and different labels. This leaves you little information about where you are located. With today's demand on files in the digital world the desktop metaphor with the folder structure may be insufficient.

In an article made by Mark Lansdale (Lansdale 1988) he points out several issues with recalling the position of files. Some of them also exist with the desktop metaphor. The problems occur since the available software tools do not support the nature of the functions behind our semantic memory like our concept creation and the way we store and categorize things in the semantic memory (ibid). For more information about Lansdale's studies see Appendix A.2.

Not only have our amount of data on computers grown. Today most computers have access to the Internet. E-mails, files, pictures, and videos located on the Internet will further accelerate the problem with more data. When searching for files and information on the Internet another problem occurs, people sometimes get lost, and cannot figure out where they are, where they have been, and how to get back (Park & Kim 2000). This is not only a problem when browsing through the Internet, but also a problem when browsing through the files located on a computer. So not only do we have to focus on creating a system where you are able to find you files. The system also needs to provide information that helps you to recall the route in the system so you do not get lost. For more information about a study concerning users' disorientation when using the Internet see Appendix B.

The goal of the study conducted in this paper is not necessarily a break with the desktop metaphor. The study is merely an attempt to create a graphic user interface with an organization structure which allows the user an easier access to files and functions in computer software. In order to create such a system we need to have a better understanding of the processes behind concept creation, categorization, and human memory.

Chapter 2

Problem analysis

By looking further into the psychological processes behind how concepts are formed and how we categorize objects we hope to find ways to improve the existing graphical user interfaces. The improvements could be better support for recall and easier ways to find our files and functions in the digital world. This chapter will therefore present theories and studies that have been used in this work.

The working thesis in this project is:

> *How can the information about our semantic memory and our conceptualization and categorization of objects help us in order to create a graphical user interface that better support recall and help us find our files?*

In this paper the time dimension will be shortly mentioned but will not be addressed further through the paper. Actually we will discuss it with the spatial dimension referring to the spatio-temporal dimension or motion. But we will not discuss time as a dimension of its own, and even though time might be a powerful tool in categorizing and recalling we are not addressing this further.

When reading this paper the terms "concept" and "category" will be used. It can be hard to distinguish them from each other since a lot of concepts will also function as a category, and the other way around. When using the term concept we refer to learning a model or an idea of something. For instance when talking about the concept of a house we do not refer to a specific house but to a house in general. A concept of a house probably has walls and a roof and people live in it. When we use the term category we do so because we focus on its categorical attributes like we can judge whether a car port belong to this house category, or not.

In the following section we will elaborate on findings in different field studies that will explain some of the problems that occur when keeping files in the digital world.

Some of the movements in the area of organizing and categorizing information on a computer have its offspring in observing how people organize and categorize their files and information in real life (Barreau & Nardi 1995, Fertig, Freeman & Gelernter 1996, Malone 1983, Lansdale 1988). For elaborated information about organization and categorization see Appendix A.

Thomas W. Malone (Malone 1983) conducted a study were he observed and interviewed how office workers organize the information in their desks and offices (ibid). He discovered that office workers organized their information in two different ways. Some organized the information in piles, some in files. The piles were simply a bunch of papers lying in a pile on the desktop or somewhere else in the office. Others had their information organized in a system with folders and titles which Malone referred to as files (ibid).

Office workers used the location of the piles to find the information they were looking for. They used the spatial location of the piles as reminders of where they had placed their information (ibid). Barreau and Nardi (Barreau & Nardi 1995) explored that this was not only true when searching for information in piles. In their experiment they observed users who were to find files on their computer.

They concluded that people also use the spatial information of the files when they look for them in folders or on diskettes in a computer system (ibid).

The use of the spatial location is so strong that you do not even have to offer the users a visual space (Fertig, Freeman & Gelernter 1996). The MIT Semantic File System offers a DOS like search were people type the name and a direction to find a specific folder. When users use this system they create their own virtual space when placing their files (ibid). It is possible to create a virtual space in your mind and still use the location of the files. The use of visual elements[1] to support the spatial location seems to help in recalling the position of a given file (Lansdale 1988).

Another dimension (spatial being the first) that people use, when trying to recall where they placed their files, is time (Fertig, Freeman & Gelernter 1996, Lansdale 1988, Rekimoto 1999). People are good at remembering when they worked with a file, therefore a system that supports this could help people find the files they are looking for (Fertig, Freeman & Gelernter 1996, Rekimoto 1999).

Besides the spatial information Malone (Malone 1983) noticed that the piles served as reminders for future tasks. The office workers also wanted information that is frequently used to be easily accessed (ibid). The same goes for elements on your computer such as to-do lists, mails, notepads, newly downloaded articles etc. which you would like to keep on the top level of the folders on the desktop (Barreau & Nardi 1995).

Lansdale (Lansdale 1988) mention another reason why people often has documents lying around on their desk. The reason is either that they do not trust themselves of being able to retrieve the documents from the folders or because they do not know how to categorize them (ibid).

The problem with keeping information in files is that it takes time to create labels, folder the information, and choose the right labels for them (Malone 1983).

Furthermore Barreau and Nardi (Barreau & Nardi 1995) observed that the people, they observed, had problems retrieving the files they had placed into a deeper categorization (ibid). The reason for this could be that by categorizing you also cut off some information. For instance, if you have a document about health insurance you can either store it under health or insurance. If you store it under health you miss the information about insurance, or if you store it under insurance you miss the information about health (Lansdale 1988).

Lansdale (ibid) have looked further into a system were you selected different regions of interest in a hierarchy. In this system people simply made too many mistakes. For instance, selecting the wrong regions of interest because some of the labels were ambiguous. Also some of the ideas or expressions come with different names (ibid).

When you cannot remember which category you have placed a file in it is because you failed to recall the memory of the files' location. The reason why this happens is because we remember the meanings of events not the details (ibid). We interpret the information in the context we learn it when we store information in our memory. When we want to retrieve the information we have to think about things related to the interpretation (ibid). This is why the spatial information, time information, color, form, and other cues can help us recall where we have placed a file (Lansdale 1988, Fertig, Freeman & Gelernter 1996, Barreau & Nardi 1995). We do not always remember all the cues about a file or perhaps the cues were not relevant in the context the file was used. This means that a system based on remembering all three attributes fails. On the other hand a system where the users only have to rely on one of the attributes will help the users (Lansdale 1988).

[1] Form and color.

Over time memories are forgotten or at least weakened (Eysenck & Keane 2005, p. 215). This means that a system which totally relies on recall is prone to error. So in order to create a better system we have to look in other directions. This does not mean that we cannot use the knowledge about recall. When designing a system the aim should be to design a system that provides the best scenario for recall hence provide the right cues to enhance recall. The system should also have a logic categorization, so when people are searching for the location of a file they most likely look the right place. In order to figure out how to do so we have conducted a literature study of the semantic memory and how people create concepts and categorize elements.

So far the field studies have given us an insight about some of the problems and some helpful clues about what helps people when they are trying to find their files, functions, and other stuff in the digital world. In the next section we will try to explain how concepts are formed, and how we categorize items within a given concept. While during so we hope to get a deeper understanding of some of the problems stated in the previous section and why some elements or attributes help us recall better. From this study we hope to get information and tools that can help us create better user interfaces.

2.1 Creation of concepts

First we will explain how we can categorize objects from each other. For this purpose we will introduce Eleanor Rosch's prototype theory (Rosch & Mervis 1975), Jean M. Mandler's theory about conceptual categorization (Mandler 2007), and S. Ratneshwar's work: Goal-Derived Categories (Ratneshwar, W., Pechmann & Moore 2001). We will then explain how concepts are created. The perception part will be explained by James J. Gibson's theory about visual perception (Gibson 1986). This theory will be supplied with the studies conducted by Mel Goodale and David Milner (Goodale & Milner 2006). The cognitive part will be explained by Lawrence W. Barsalou's perceptual symbol theory (Barsalou 1999). We will also explain how context can affect the forming of concept. This is done with the Activity theory by Alexei Nikolaevich Leont'ev (Leont'ev 2002).

One of the most famous theories about categorization is the prototype theory (Rosch & Mervis 1975). According to the prototype theory categories are build upon the family resemblance between items. Family resemblance is visual attributes in common with other items (ibid).

Within these categories some items seems to be more prototypical than others. If we take the category "furniture" you are more likely to think about a sofa than a bookshelf hence a sofa is a more prototypical furniture then a bookshelf (ibid). The more family resemblances an item have with other items within a category the more prototypical the item is. Prototypical items also have a low family resemblance with items in other categories (ibid). For more information about the prototype theory see Appendix E.2.

One of the problems with the original theory is that it only takes the visual attributes into account. The theory does not exclude that other attributes play a role. They just have not examined other attributes - as stated:

> *Family resemblances (even broadly defined) are undoubtedly not the only principle of prototype formation.* (ibid, p. 599)

Jean M. Mandler (Mandler 2007) worked with conceptual categorization - how we separate humans from animals, cars from airplanes etc. Through her study of infants she noticed that when infants start their concept forming and categorizing of objects they do so according to perceptual attributes (ibid, p. 745). The perceptual attributes are attributes like form and color, and can therefore be compared to the family resemblance attributes (ibid).

She further observed that later on infants learn to categorize elements on behalf of their conceptual similarities. This could be similarities like what elements can be used for. A saw and an axe can both be used for cutting down wood. As another example infants can also differentiate between kitchen stuff and bathroom stuff (ibid). The perceptual similarities and concept similarities melt together through our language later on (ibid).

S. Ratneshwar (Ratneshwar, W., Pechmann & Moore 2001) conducted a study where he investigated other attributes than perceptual similarities/family resemblance. The study included 127 undergraduate students. The students were asked to rate the similarity of different foods. The participants judged the similarity of eight pairs of food products. In a test they varied between food categories with high and low surface resemblance. Furthermore they varied the salience of a shared context like "things you can eat in your car". Some of the food categories fitted this context, others did not. They also varied the food categories according to a personal goal of health (ibid).

The food types differed each time in which shared attributes (personal goal, context, surface resemblance) they had in common. And also the salience of the context and the health goal differed (ibid).

The study showed that people are influenced by the context, the personal goal and the surface resemblance when judging on the similarity of different food types (ibid). For more information about the study by Ratneshwar see Appendix E.1.

So not only do we categorize based on perceptual attributes we also categorize according to what Mandler refers to as conceptual similarities (Mandler 2007). But how does this actually work? How do we transform information about color, shape, context etc. into actually concepts?

Mandler (ibid) has studied how our conceptual system evolves in the first place, how we start to learn the attributes of a concept on which behalf we categorize it. She has not found a direct answer but studies have proven that infants pay extra attention to motion and in some cases spatial relations. At the age of two months infants start to notice whether objects move by themselves or by the action of other objects (ibid, p. 745).

When forming our conceptual system we start with creating global concepts[2]. The global concepts then get subdivided according to spatial information. These subdivisions then get subdivided again and so forth until our conceptual system is created (Mandler 2007, p. 747).

Even though Mandler's and Ratneshwar's studies (Mandler 2007, Ratneshwar, W., Pechmann & Moore 2001) show that elements like perceptual attributes, context, motion, spatial relations, conceptual similarities play a role in our understanding of the world they do not fully explain how they work together. Another psychologist named Lawrence W. Barsalou (Barsalou 1999) has come up with a theory that cover all these aspects.

During perception of an event neurons in the sensory-motor system of the brain capture information from the event, and also from the body of the observer. This captured information is known as perceptual symbols. Perceptual symbols could be edges, colors, movements, heat, pain or spatial relations (Barsalou 1999, p. 583).

The perceptual symbols do not affect a whole object but rather a specific schematic aspect of the object for instance a shape or a color. If our selective memory focus on a specific aspect of an object this specific aspect is most likely what we store in our long term memory (ibid, p. 583).

[2]Rosch refers to this as subordinate categories like vehicles, animals etc. (Rosch & Mervis 1975)

When a perceptual symbol is stored in our long term memory it is stored in an associate pattern of neurons, and thereby associated with other perceptual symbols. This gives the perceptual symbols the opportunity to be activated in different patterns according to different contexts. So the perceptual symbol patterns can change over time creating new patterns of categorizing objects, and creating new concepts (ibid, p. 584). For more information about perceptual symbols by Barsalou see Appendix E.4.

The reasons why the perceptual symbols' patterns change over time and why categories can change according to the context is explained with the activity theory. This is because of its underlying basis in how the interaction between a subject and the environment can influence each other and change the subject's activity.

Activity theory by Alexei Nikolaevich Leont'ev (Leont'ev 2002) tries to explain the entire human activity system by implicating the history of the individual, culture, motivations and role of the object. Leont'ev clarifies that the very concept of activity implies an acting subject and an object the activity is directed towards (ibid). The activity therefore becomes the relationship between the subject and object. In Leont'ev's perspective activity is defined when an organism physically seeks objects. For instance, a plant only needs energy and nutrients in a gaseous or liquid form of metabolism, so the plant can merely wait for them to accrue to it. It does not need to search (activity) for these living conditions. When an organism has to search for food material it needs a higher organized activity, which is object- and goal-oriented (ibid).

So activity is the relation between subject and object, but the relation can be considered as an interaction between the subject and its environment, which mutually influence each other (ibid, p. 25). In the process of an activity intentional adjusted mental activities are being executed which brings the organism in contact with objects and conditions which changes the structure of the activity. Thereby it influence the subject's sense organs in ways that give information. The structure of the activity is therefore changes in correlation with the properties of the object and with time experience is accumulated in the organism (ibid).

The terms activity, action, and operation can on a subordinate relational level describe the activity. These three terms are closely linked to motive, goal and condition (ibid). Human activity is initiated by the motive, which is based on a need. The actions, which carries out the activity, is controlled by the goals (ibid, p. 31). So activity and action is linked to what must be done to satisfy the need and the intentional goals. Operations however, are linked to how it is done and are therefore influenced by the attending conditions. The goal of the actions can remain the same, while the conditions, which relates to the operations can change. So, human activity is constituted by a series of actions, which each independently are controlled by the goals, which is adequate to the motives under the given circumstances. The actions "what" is being accomplished by operations "how", which are available under the given conditions.

Activity, whether or not it is interior (e.g.thoughts) or exterior (e.g. physical activity), is mediated and controlled by the mental reflection of the reality (ibid, p. 93). Things in the object-oriented world, which for the subject acts like motive, goals and conditions must in someway be perceived by the subject and imagined, comprehended, located and reproduced by its memory (ibid). This enables the subject to grasp the object-oriented world. The consciousness is in its immediacy a picture of the world, which appears before the subject, and includes the subject itself, its actions and conditions (ibid).

The mental reflection of the reality requires that the subject can collect perceptual information from the object-oriented world (ibid, p. 102). This sensory representation happens through the sensory fabric of the human consciousness (ibid). This means that the sensory fabric forms the experience of reality by allocating the subject a conscious "image" of the world, which exist outside the subject (ibid). Leont'ev presents the term personal meaning to describe the subjective side of the consciousness (ibid). The personal meaning is what reflects and holds its own reality of life and is therefore in close relation to its motive (ibid, p. 41). Personal meaning regulates and guides the human activity (ibid).

So humans consciously generate activities in relation to the context. Context can be internal or external. Therefore the context is not just an outer frame for how people behave, but it involves specific objects, goals, people, settings etc. It becomes part of peoples' own objectives. For more information about the activity theory by Leont'ev see Appendix D.

As mentioned earlier the perceptual symbols are schematic of nature, which means that they are not holistic. The way perceptual symbols are stored in memory makes it possible to separate the colors, shape, and orientation, but still it is also possible to connect the perceptual symbol components to form a specific object (Barsalou 1999, p. 584). Neurons in the perceptual system can code information qualitatively as long as it is unconscious information. Like coding the presence of a line without information about the length, position or orientation. This further explains how the perceptual symbols can be coded into schemes (ibid).

It is important to add that perceptual symbols are multi modal in nature which means that the perceptual symbols also come from other sources such as audio. People acquire perceptual symbols from speech and sounds, from touch, temperature, and texture. Also we acquire perceptual symbols from proprioception - like movements of objects and body positions (ibid, p. 585).

Another modality that can capture perceptual symbols is introspection (ibid). Introspection can be divided into three areas (ibid):

- Representational states include the representation of an entity or event in its absence, as well as construing a perceived entity as belonging to a category
- Cognitive operations include rehearsal, elaboration, search, retrieval, comparison, and transformation
- Emotional states include emotions, moods, and affects

As mentioned earlier perceptual symbols are stored in association with other perceptual symbols. These associations or relations make it possible to simulate an event. If you for instance study a car the information about the body, wheels, windows, and doors are stored in accordance to their spatial location in an object-centered reference frame. More information about the car is added to the reference frame when moving around the car and observing it from different angles or when you look into the trunk etc. Aspects like the sound of the engine, the movement of the car etc. can be included into one of these reference frames or schemes. The information concerning the schema of a car can come from many different experiences with a car. A reference frame like the one in the example just above is named a simulation (ibid, p. 586).

A simulator contains two levels. The first level is the frame that integrate the different perceptual symbols. The second is the potential sets of simulations that can be constructed from the frame (ibid).

The simulator becomes equal with what others might refer to as a concept. During childhood you develop simulators for important types of entities and events. Once you can simulate an event to a culturally acceptable degree you have an adequate understanding of it (ibid). When a clear concept or simulator have been created it is possible to decide if a given entity is a member of a given category.

If the simulator is not able to generate a satisfying simulation the entity is not part of the category and will therefore not be stored in the same association as the given category (ibid).

During the development of our language the perceptual symbols are linked with linguistic symbols and thereby also linked to the simulations. The linguistic symbols develop in the exact same way as perceptual symbols and whenever you hear or read a word of a concept like "a car" the simulation of the concept will start (ibid, p. 592).

Once an entity is categorized you can use the knowledge from this category to provide predictions and thereby further knowledge about the entity. This includes ways of interacting with it. This can be beneficial when working with a new object. Have you once driven a car you already got a lot of information on how to drive a truck (ibid, p. 586).

So far we have explained how spatial location plays a big role in how we create concepts because not alone do we perceive the spatial location of the different elements in the things we see, we also use it in order to construct our concepts. However the focus in Barsalou's theory is how our cognition work and not so much how we perceive the world. Since our focus in this project is on graphic user interface we have chosen only to cover the visual perception and we do so through James J. Gibson's theory about visual perception (Gibson 1986).

Our perception, according to Gibson, is based upon three fundamental elements: medium, substances and surfaces (ibid, p. 16). The medium is what we move through (e.g. water is the medium for fish and air for terrestrial animals), and the medium facilitates different kinds of perceptual systems to evolve. The air for instance makes it possible for different kinds of dispersion of e.g. light particles, chemical particles, and vibrations, and thereby facilitates the visual, auditive and olfactory systems (Trettvik 2001, p. 5). The substances are objects we can touch, look at or eat - it is what we have trouble using as medium (ibid). The surfaces is where the medium and substances meet (ibid).

When light penetrates the medium and hits a given substance the light is reflected and is perceivable by an observer (ibid). Gibson calls the reflected light for ambient light (ibid, p.4). The ambient light is therefore a result of the light which is spread in the medium and illuminates the surfaces in the environment (ibid, p.5). In order for our visual system to make any sense out of light it must be structured. The environment and its surfaces structure the ambient light. This means that when light is reflected it has, because of the surface's layout, a variance in light density, which is specific for the surface the light is reflected from. Gibson calls this ambient optic array which is considered by Gibson to be sufficient for perception (ibid).

In order to perceive the real world it must involve an active observer who is constantly moving his eyes, head and body relative to the environment (ibid). Even though the constant movement of the observer results in a constantly changed image in the observer's retina there is information that remains constant on the retina (ibid, p. 73).

As the observers or the objects move more information becomes available. This is because of the changes in the ambient optic array (Trettvik 2001, p. 6). Every point of observation is surrounded by ambient light that is specific for that exact point of observation, and when the point moves (when the observer moves) the array of the ambient optic changes, or at least some parts of it do (perspective structure) (ibid). The parts that do not change Gibson defines as invariant and these invariants are what makes us perceive the world as rather constant even though we move (ibid). This is because the invariant structure is behind the transformations in the perspective structure and specifies the forms of the rigid surfaces of objects (ibid). For instance when looking at a cup we see its surfaces. If we move our head or eyes just a little bit we see the cup from a slightly different angle. This means that a part of what we are watching has changed, but still we see most of the structure we saw before. This is what Gibson refers to as the invariant parts and from the invariant structure we are able to recognize an object from the rest of the world. Figure 2.1 illustrates an example of the constant information that is available of an object (a cup). For more information about Gibson's theory about visual perception see Appendix E.3.

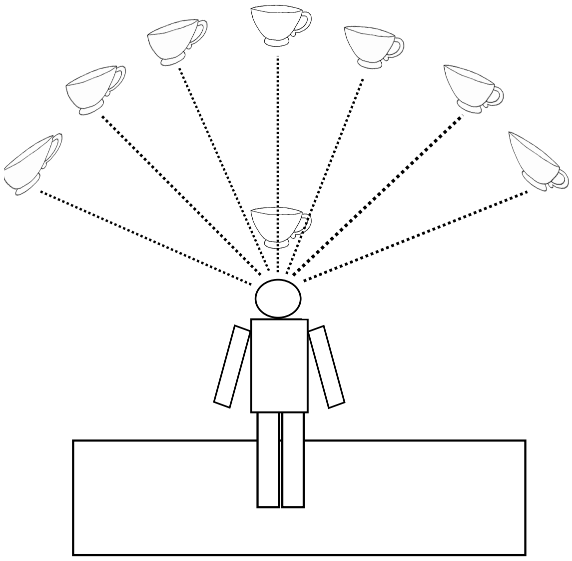


Figure 2.1: A schematic illustration of the invariant parts of a cup (Gibson 1986).

When it comes to visual perception it seems like two systems exist. One can be described as the perception of objects and the other as perception of space (Goodale & Milner 2006). Mel Goodale and David Milner (ibid) showed that the eyes of a rodent does not simply send input to the superior colliculus[3] and the visual system but in fact the eyes send to at least ten different areas in the brain (ibid). Each of these areas appeared to have control of its own separate class of behavior (ibid). Their study showed that the superior colliculus is involved with guiding eye and head movements towards potentially important visual objects and that another subcortical structure plays an important role in guiding animals around their environment (Goodale & Milner 2006).

It is argued that the ventral system, which passes from the primary visual cortex to the inferior temporal lobe is concerned with object identification while the dorsal system, which passes from the visual cortex to the posterior parietal lobe is charged with object localization. They can therefore be considered as two separate processing "streams" (ibid).

For investigation of this view Goodale and Milner made series of neuropsychological studies in the late 1990s with a patient who had severe visual agnosia[4]. The patient had not only trouble with identifying objects but also with basic discrimination of simple shapes (ibid). In their study they examined what the patient was capable of and not. It turned out that even though the patient had trouble identifying or demonstrating the orientation of a slot, the patient could post a card into the same slot without errors (ibid). Furthermore, the patient could manage to tailer her finger-thumb size in order to pick a rectangular block without being able to verbally or manually describe the width of the block (Goodale & Milner 2006, p. 2).

[3]The superior colliculi is two small bumps above the mid brain, which receive optic tract nerve fibers, that are involved in processing spatial aspects of visual information and eye movements in the direction of visual attention (Colman 2009, p. 741)

[4]Visual agnosia impairs the ability to recognize or identify visual images or stimuli (Colman 2009, p. 19).

In their later work Goodale and Milner showed what an important role the ventral stream have in visual memories. Goodale and Milner tested this by presenting a visual object briefly to the same patient as before after which the object was then taken away. The patient was then asked, a few seconds later, to pick up the object as if it was still there (ibid). The experiment showed that the patient completely failed the task. The patient was not capable of tailoring her finger-thumb separation when the patient reached out and tried to pick up the object, that was briefly shown (ibid, p. 3). This outcome was contrary to the normal situation of grasping a visible object where the patient was capable of tailoring her finger-thumb separation perfectly (ibid). This is explained by the patient's lacking possibility of perceiving the dimensions of the object in her consciousness. Therefore the patient had no working memory of the concerned object.

It is therefore safe to claim that the ventral stream plays an important role when acting on our visual memories. As oppose to the dorsal stream, which seems to have no visual memory, the ventral stream allows us to use vision "off-line". It connects the past with the present (ibid). It is however suggested that even though the dorsal stream does not have a visual memory, the visio-motor activities do benefit from experience, and being well-honed by practice the visually guided actions becomes more automatic (ibid).

Even though Goodale and Milner's and others' work does not show how the two streams are working together, their work clearly present the idea of the dissociation of the visual system. For more information about Goodale and Milner's studies see Appendix C.

2.2 Operationalization of theories

To sum up we have first examined field studies about people trying to organize their files, function, and information and how they found their files this either in an office environment or in a graphic user interface (Barreau & Nardi 1995, Fertig, Freeman & Gelernter 1996, Lansdale 1988, Malone 1983).

The field studies have shown us that form and colors along with the spatial location help us remember where we have placed our files. These findings is found to be grounded in our conceptualization and have been explored by Rosch (Rosch & Mervis 1975) and Mandler (Mandler 2007). The field studies also showed that context plays a big role for recall and that when categorizing files it is hard to name the files. The attributes, which are focused on, vary. Thereby giving a file a name that fits all attributes can be hard. The reason for this has been explained through Ratneshwar's studies (Ratneshwar, W., Pechmann & Moore 2001).

So these psychologists have shed light upon why these above mentioned aspects help and give trouble when it comes to recall. They do so through their exploration of how we form concepts. The only problem is that they do so through none-connected theories. This is the reason why we have brought Barsalou's theory about perceptual symbol system into the picture (Barsalou 1999). When connected with the activity theory (Leont'ev 2002) the perceptual symbol system explains how form, color, and spatial location are connected. When forming our concepts the activity theory shows how it is possible to focus on different parts of these aspects in different situations. The activity theory explains why different contexts are priming the conceptual system to focus on different aspects. Finally we added theory about the visual perception to better understand what we perceive when looking on a graphic user interface.

Now that the mechanisms and theories behind conceptualizing and categorizing are explored how do we use this to our benefit in order to create a better structural system?

If we look at the known systems using the desktop metaphor[5] we see elements like folders, text documents, image files etc. located on a flat desktop. Even though the different types of objects looks different the same type of objects look exactly the same. You cannot distinguish one folder from another folder besides by looking at the label. The Windows 7 desktop is shown on Figure 2.2.

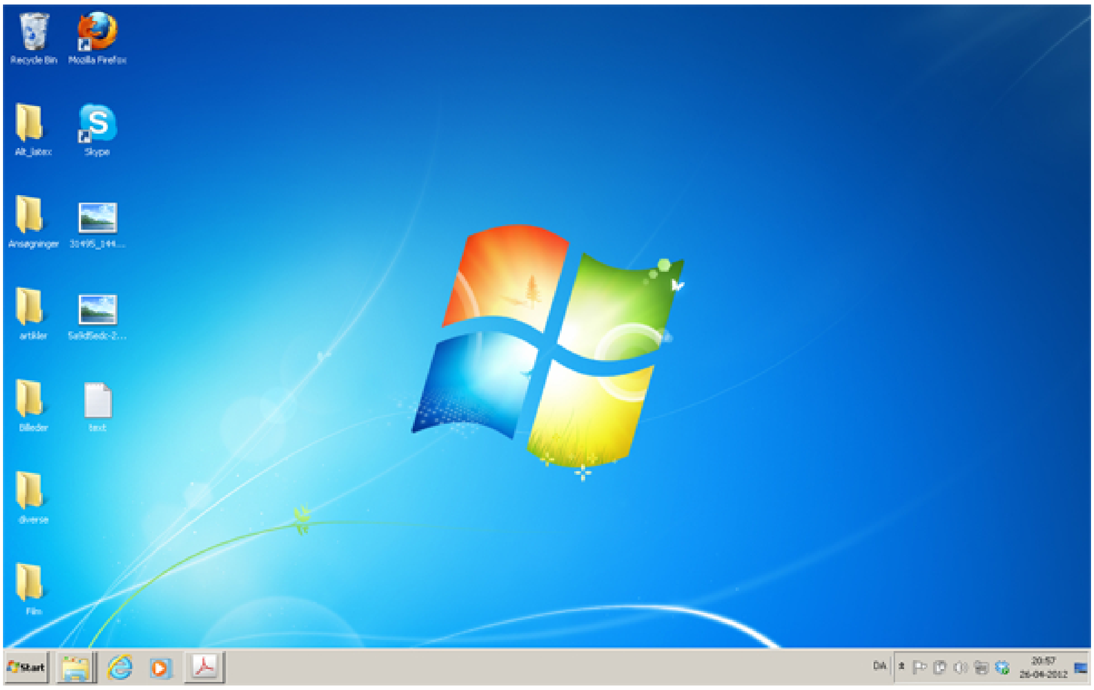


Figure 2.2: Picture of the Windows 7 desktop.

From the desktop you can enter the folder structure. The folder structure is a tree like structure placed on a flat desktop. From the first folder you can enter other folders located in the folder which again have folders located inside themselves. Apparently you can create an infinite number of folders within each other.

In a desktop system it is possible to place objects in different locations. The problem is that the support for the perception of the spatial dimension is not very good. The perception of spatial dimension somehow exists, when placing the file in one of the main folders placed on the desktop but when you enter the folder structure the spatial dimension becomes really abstract. The thought about having infinite folders within folders makes no sense at least in the physical world. Another problem is that we cannot distinguish one folder from another because they look exactly the same. So how do we know which one to pick? The only way we can distinguish the folders from each other is by the spatial location. If for instance you have two subfolders they will be identical only because the folders are at fixed locations. The only difference is the number of folders in the subfolders but even this can be the same. On Figure 2.3 an example of two identical subfolders is shown.

[5] In this example we look at the windows 7 desktop but we believe that most of the claims made about this desktop also goes for almost every user interface that uses the desktop metaphor.

As stated in the perceptual symbol system theory we organize items according to their spatial location. This is further elaborated in the article: "Spatial symbol systems and spatial cognition" (Freksa et al. 1999):

> *Neural representations resulting from perception are often organized in sensoritopic representations, that is, according to spatial structure manifested in the perceived configuration. As entities are perceived in spatial relation to one another* (ibid, p. 615)

The default wallpaper of the desktop is very uniform. The information about the spatial location from the wallpaper is very limited. The information we get about the location of the different folders we get from the interrelationship between the folders. Again we have the problem with the identical folders. Remembering in which folder you placed your file is like remembering a black dot on a white wall with ten black dots surrounding it.

The visual environment in a desktop only support a two-dimensional space. As mentioned in Gibson's work (Gibson 1986) the information for perception is information that remains invariant while an observer moves through the environment (ibid). Since there is no spatio-visual information available when the observer "moves" into the folder the perception of actually moving into the folder does not exist, and is limited to an abstract metaphorical idea not supported by your visual system.

This leads to our first hypothesis:

In a system that supports the objects' spatial locations in a perceived three dimensional space, people will be better at recalling and finding the objects located in the system

When starting at the desktop you can enter one of the folders on the desktop. Afterwards you can choose to either go further down into a folder located in the folder or you can go back. When you have entered the next folder you can again choose either to go deeper into the structure or you can go back to the previous folder. If you want to go to a folder located right beside the folder you are located in, you will first have to go back and then select the side-folder. Figure 2.3 shows how you navigate through folders. If you are located in folder B you have to go back to folder A to enter folder C.

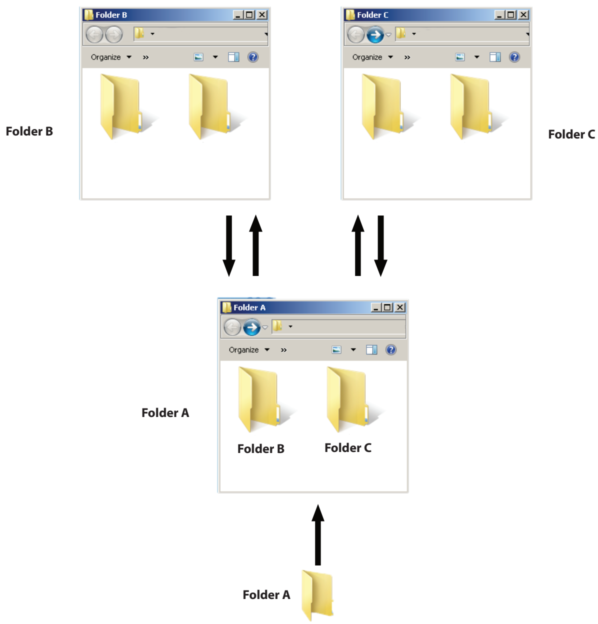


Figure 2.3: The figure shows how the navigation through folders work. Folder A is located on the desktop. From folder A you can enter subfolder B and C.

As stated in the beginning of this chapter a system which only relies on recall is prone to error. So even with the greatest support from the spatial location we still expect that people make mistakes, sometimes people will go to the wrong folders. If people organize their files on their computer the way they categorize items from each other in a hierarchical structure like the one presented by Rosch (Rosch & Mervis 1975) and Mandler (Mandler 2007), then the best place to look next is probably in the side-folder. Therefore a system should allow access to the side-folders.

This leads to the second hypothesis:

A system that allows moving to side-folders will shorten the route to the file they are searching for

If you allow more routes to travel by in a system it will also become easier to get lost, and since people already loose track of their location in the desktop folder system more information about your location in the system is needed. An article by Jonnah Park and Jinwoo Kim (Park & Kim 2000) have shown that if you are offered contextual information about your location in a system it will increase your performance (ibid). By comparison to the real world contextual information is all around us. The visual perception for instance gives us vital information about where we have been, where we can go, and where we are in correlation to our surroundings. Contextual information in a folder system could therefore be information about the surrounding folders.

This leads to the third hypothesis:

A system that gives contextual information about your location in the system will give you a better overview

Chapter 3

Methods

In this chapter we will cover the methods we are using to answer our three hypotheses:

- *In a system that supports the objects' spatial locations in a perceived three dimensional space, people will be better at recalling and finding the objects located in the system*
- *A system that allows moving to side-folders will shorten the route to the file people are searching for*
- *A system that gives contextual information about your location in the system will give you a better overview*

In order to investigate the hypotheses we will first create two prototype systems that the test subjects can interact with. One control system and a system that enhance the spatial perception and support navigation.

We will use an empirical-analytic approach (Kjørup 1997, p. 155-156) where we will see if the manipulation of the system will give an effect. This will be done by using Inferential statistics. Inferential statistics tell us whether the experimental hypothesis is likely to be true. Within most research there is an inherent prediction that the researcher has made. This is called the experimental hypothesis. The experimental hypothesis states that an experimental manipulation has an effect. The null hypothesis states that the experimental manipulation does not have an effect (Field & Hole 2003, p. 141-142).

When testing the hypotheses we work with probabilities. We calculate how high the probability is that our hypothesis is true and the null hypothesis can be rejected. As a threshold for when something is true we have the 95% confidence interval. When we are 95% certain that something is true we accept it as being so (ibid).

When being 95% confident there is still a small chance that we are wrong. We can either believe that there is a difference between two groups when there is not or we can believe that there is not a difference between the two groups when in fact there is (ibid, p. 149-150).
So we cannot be a 100% sure if something is true, but we can say that it is most likely true.

Besides using the quantitative research methods, we will also introduce some qualitative research methods. As stated by Ole Riis (Riis 2005):

> *We need both forms of information. Therefore it is not a question about which method to use but a question about how we can combine them.* (ibid, p. 203)

We might be able to discover differences in the use of the two systems with inferential statistics but we probably will not be able to tell why and how these differences occurred. To that purpose we will use observations of the users' interacting with the systems and furthermore we will conduct an interview to try to make them describe how they experience the system. This is done both in order to answer our hypotheses but also in order to raise the validation of the result from the inferential statistics.

The interview will be a qualitative research interview (Kvale 1997, p. 40). The purpose of the qualitative research interview is to understand the interviewed person's relation to the world and what the person experience within a given topic. The interview should aim at getting uninterpreted descriptions of the topic of the interview. It is important that the interviewer is open to unexpected openings in the interview and is able to explore the new directions the interview can take when it is useful to the topic. Therefore the interview can not be too structured but be able to grasp the different perspectives within the topic (ibid, p. 40-45).

The interview will be a semi-structured interview which means that we will have prepared template questions regarding the present topic. The questions are however not fixed and can be changed during the interview. This gives us the possibility to ask in-dept questions to get more elaborated answers (ibid, p. 133).

Observations in a qualitative study are very flexible. The facilitator can shift focus as he pleases and get details about unforeseen data sources. By using observations the facilitator gets information about the test subjects' way of interacting with a given system. We see what the test subjects do instead of them telling us what they did or are about to do. A problem with this is that the mere present of a facilitator can change the way the test subjects interact with a system. A recording item could have the same effect. There is a risk that instead of focusing on the given task the test subject focuses on the recording device (Leedy & Ormrod 2010, p. 147).

As already mentioned all three hypotheses involves creation of a software system, and in order to compare such a system with the present systems, we had to construct it, or at least be able to simulate such a system. Therefore we created a prototype system and a control system. Details about the construction of the prototype system and the control system can be found in the next chapter.

Chapter 4

Prototype systems

We wanted our prototype system to better enhance the perception of working in a space. In order to do so we needed to implement a metaphor from the real world that actually supports a person moving around in the environment. Another issue is that in the current desktop system you have the opportunity to create infinite folders. The possibility to create infinite storage places for your files gives the user freedom. However, in the folder example you end up with folders located in folders. The problem is that it is hard to support such a function without creating something very abstract with a bad mapping to the real world and thereby create something that does not correspond well with human cognitive skills. Rune Nørager and Johan Trettvik further elaborate this by saying:

> *Human cognitive skills are to different degrees however very rigid and dependent on certain regularities that reflect the physical world. This makes sense since the human brain has evolved to exploit the dynamics of the physical world* (Nørager & Trettvik 2007).

In order to support the involvement of properties that enhance the perception of spatial sensation it can be stated that the brain uses rather low-level cognitive processes when dealing with orientation in an environment (Velichkovsky 1990, p. 4).

A lot of other problems exist when trying to map the real world to the digital. They occur because the computer carries out operations that do not correspond to operations in the real world (Freksa, Barkowsky & A. 1999). Therefore the mapping most be designed in a sensible way. The user must be able to see the relationships between the perceived system and the concept which is mapped from the real world (Norman 2002). This can be achieved by using a metaphor that can be simulated by computer operations in accordance with the mapped concept (ibid).

We decided to use a house as the prototype system's metaphor because it offers perception of a three dimensional space. As it is stated in the article "Intuitive user interfaces":

> *"It is possible to display the affordances of "scenes" as explorable rooms or places in a space, of walls as obstacles, openings in walls as passages to other scenes."* (Bærentsen 2001)

Another reason for choosing the house metaphor is because it allows the construct of infinite rooms in which you can store your files.

The house contained ten rooms. First a lobby that corresponds to the desktop. From the lobby there was access to three other rooms and from each of these three rooms the user could enter additional two rooms. We designed the rooms of the house in Google SketchUp (Google 2012), which is a 3D content creation software[1]. Each and every room consists of a ceiling, floor, three walls and the relevant doors. The colors on the floor and walls vary. All the rooms have different kinds of colors on the walls so it was possible to distinguish the rooms from each other on the basis of the colors. The colors on the floors only vary in correlation to the three levels of depth in the system. This means that the color on the floor in the lobby is one color, the colors of the floors in the three rooms you enter from the lobby are a second color, and the colors of the floors in the six outer rooms are a third color. Figure 4.1 shows how the lobby was designed.

[1] For more information about Google SketchUp visit http://sketchup.google.com

Figure 4.1: The lobby consists of three doors where each door connects the lobby to three separate rooms.

After the rooms were designed in SketchUp, an image of each room was exported from SketchUp to a user interface prototyping tool named ForeUI[2] (EaSynth 2012). In ForeUI we created each image of the rooms into pages which made it possible to create a connection between the pages by adding actions. The actions define the interactive behavior of the prototype. An example of an action could be the switching between pages e.g. from the lobby page to a room page which is executed by a mouse click.

From each of the rooms it was possible to navigate to other rooms by clicking on the doors. We implemented an animation when entering the doors. The purpose was to give the users a visual perception of actually moving into the rooms. We added an arrow pointing backwards. When clicking on it you go back to the previous folder. Furthermore we added doors into "side-rooms". In the structure of the folders in the desktop it is not possible to step to the side folders. The reason why we wanted people to be able to enter the side-folders in our system is that we believe that the objects in a side-folder is just as closely related to the objects in the folder as the root folder. Therefore if people do not find their objects in the folder the next place they will look is most likely the side-folder.

Besides supporting recall we also wanted the system to help the user to a better navigation. As stated earlier Jonnah Park and Jinwoo Kim (Park & Kim 2000) have shown that if you are offered contextual information about your location it will increase your performance (ibid). In the article: "Spatial symbol systems and spatial cognition" (Freksa et al. 1999) it is suggested to use a map as a metaphor in graphical user interface. A map represents a structure in spatial analogy to the world it represents and thereby the things on the map are perceived in spatial relation to each other (ibid). The only problem with a map is that the mapping becomes symbolic instead of direct. You do not see a room as it is, you see a symbol for the room as stated earlier this does not support the basic human cognitive abilities (Nørager & Trettvik 2007). Despite the symbolic nature of a map we still believe it is a good solution to provide contextual information about your spatial location and it is even doing so through the use of spatial relations. For this reason we decided to use a mini-map in order to provide contextual information. The mini-map showed the ten rooms in the house and in which room the user was located.

We also created a control system. We wanted the control system to look like a common folder system. We wanted to build a control system instead of just using one of the existing systems. The reason was that the prototype system would have some limitations compared with the existing folder system like that you cannot interact with the files in the system and there are no labels available. We did not want the system to look directly like an existing system because that might have effected peo-

[2] For more information about ForeUI visit http://www.foreui.com/

ples' opinion about it and their skills in navigating in it.

By avoiding labels in both systems we made sure that it was not too easy to find the relevant albums by reading the labels. Furthermore, when using labels the brain uses rather high-level cognitive processes when dealing with the semantic memory. Velichkovsy (Velichkovsky 1990) put it this way:

> *"[...] semantic memory or learning by restructuring, are from the realm of higher symbolic co-ordinations. "* (ibid, p. 4).

So a system that is based on involving the rather low-level cognitive processes would be preferred.

We build the control system in ForeUI just like the prototype system. The root of the control system was a screen where three identical folders were shown beside each other. It was possible to enter the folders by double clicking on them and inside each folder another two folders were located. We also added an arrow that could be used to go back. In Figure 4.2 the start site of the folder system is shown. Both the systems can be found in Appendix H.

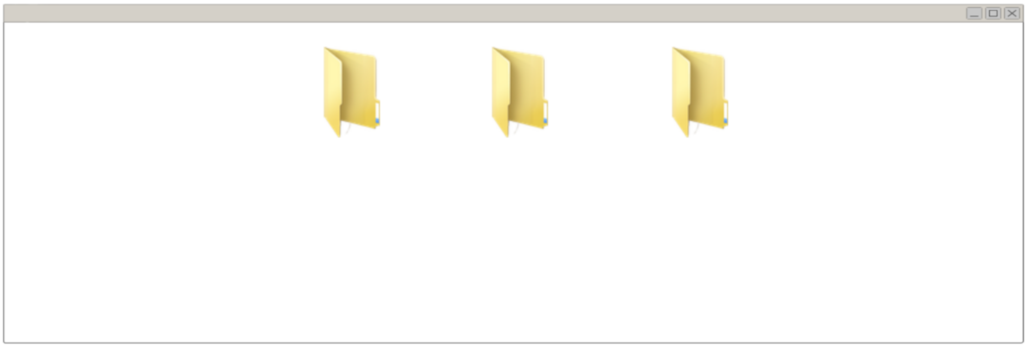


Figure 4.2: From the start site the users can go to one of the three folders.

Both the control system and the prototype system had to contain the same objects. The most obvious choice might have been some kind of text file for this purpose. However, we believe that people sort their text files according to their content or according to their history with the text file like when it was created and when it was used etc. In our systems you do not get the opportunity to interact with the text files and therefore probably would have a hard time remembering and separating them from each other. Instead we choose to use music albums. The reason we choose music albums was that there are a lot of different attributes you can categorize the albums by. This could be genre, year, band/solo etc. Also by choosing albums we were able to show the visual differences by adding a picture of the album.

In order to categorize the albums and in order to decide in which folders/rooms the albums should be located we decided to conduct a study using card sorting. We used card sorting because it is useful when the purpose is to find the categories people use (Rugg & McGeorge 2005). We might just have randomly placed the albums in the different rooms/folders in the systems since we do not think this could directly have an impact on our hypotheses. The reason why we choose to do card sorting anyway was because we believe that people use information about how things are grouped together in the different rooms/folders to remember was their files are placed in a real graphical user interface and we wanted our systems to be as ecological as possible.

Four participants participated in the card sorting study. The participants were not recruited by any particular requirement but they did however consist of two males and two females who were between 27 and 50 years old. We wanted the participants to have an average knowledge about music to make sure the categorization was not an effect of the participants being either expects or novices. Therefore we asked the four participants before the study to rate themselves on a scale from 1-10 - 1 being very limited knowledge about music and 10 being very great knowledge about music. The participants rated themselves between 3-6 which means that none of them could be seen as experts neither novices.

The study was conducted in a room with a white board placed on a table next to 63 cards. There was a picture and a label of an album placed on each card. The cards were spread so all of the cards were visible to the participants. The participants were told to place the albums in three groups. It was up to the participants to decide by which attribute the albums should be categorized. An example was used to make sure they had understood the task. After they had created the three groups they were told to divide each group into two new groups if there were any of the albums that did not fit the two new categories they were allowed to leave them out. In Figur 4.3 a picture of the card sorting session is shown.

Figure 4.3: The picture shows the four participants in the middle of the card sort.

From this card sorting we received a categorization that matched the folder/room system we wanted in the prototype system and the control system. At first level three different folders/rooms are accessible and from each of these folders/rooms another two folders/rooms are accessible. The albums were then placed in folders/rooms according to the categorization made by the participants. The albums, the participants were not able to categorize in the second categorization round, were placed in first level folders/rooms according to the first round of categorization.

In the first round the albums were categorized according to genre, the first group contained "pop" the second "rock" and the third was "others". In the second round "pop" were divided into "bands" and "soloists", "rock" were divided into "soft rock" and "heavy rock", and "others" were divided into "danish" and "foreign".

Chapter 5

Experiment

The purpose of the experiment was to investigate our three hypotheses by giving the test subject an understanding of the interaction of our system compared with a system using the desktop metaphor.

40 test subjects who were all students on Aalborg University participated in the experiment. 22 men and 18 women. The experiment was conducted in a lab. The test subjects were to sit next to a table with a PC and an eye tracker located on it. The experiment was conducted by two facilitators. One facilitator gave the necessary information to the test subjects. The other facilitator started the prototype and eye tracker and made note of the test subjects' statements during the interviews and their performance.

The experiment was a between subject design. The independent variable was the prototype and the control system. Each participant was only interacting with one of the systems. The screen and where the participants were looking, were recorded by the eye tracker. This information was used to figure out how many mistakes they made when trying to recall. It was also used to track their movement through the graphical user interface and to see if they were looking at the mini-map. A dictaphone was used to record the interview so it was possible to analyze the qualitative data. The facilitator used another computer to note the test subjects' statements during the interviews and their performance. The dependent variable was how many mistakes they made, if they used the mini-map or not and if they used the possibility of going to a side-folder when needed. In Figure 5.1 a schematic drawing of the test setup can be found. Figure 5.2 is a picture taken during the experiment.

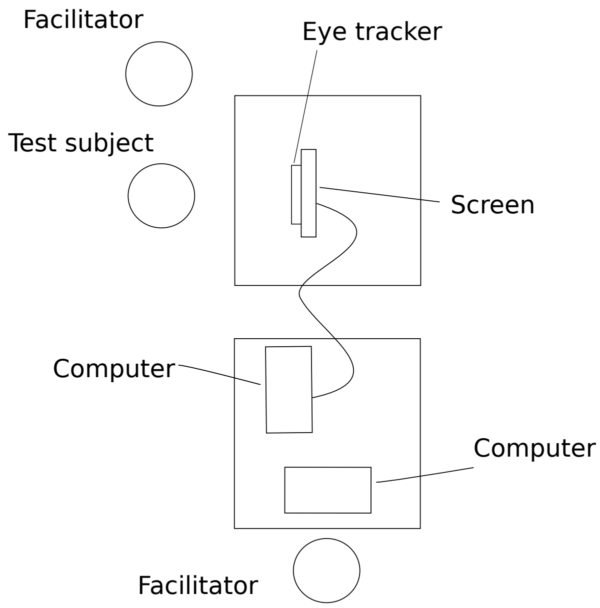


Figure 5.1: A schematic drawing of the test setup used in the experiment.

Figure 5.2: A picture taken during a session with a test subject.

5.1 Procedure

First the test subjects were to sign a declaration of statements[1]. Afterwards the test subjects received instructions about the experiment[2]. Then they were asked to rate their knowledge about music. Afterwards the eye tracker was calibrated. The test subjects were given three min. to interact with one of the systems to learn how to navigate and see where the objects were located in the system. The reason for this was that they needed to see the position of the objects otherwise they would not be able to recall the objects' position. Another reason why the test subjects were given three min. to try the system was that they probably had experience with a system that looked like our control system. We did not want that to effect the results. By giving them this training with both the systems, we hoped to minimize the effect of their past experiences. After the three min. they were asked to go back to the start point of the system. Next they were given the first of ten tasks. The task was to find a specific object in the system. When given the tasks they received a picture of the album with a label. When they had found the object, they were given the next task. When they had finished the ten tasks they were interviewed. After the interview they were done with the experiment. Figure 5.3 shows the pictures of the concerned album covers that were used in the ten tasks.

[1] All the declaration of statements can be seen in Appendix H.
[2] The verbal instructions can be seen in Appendix G.

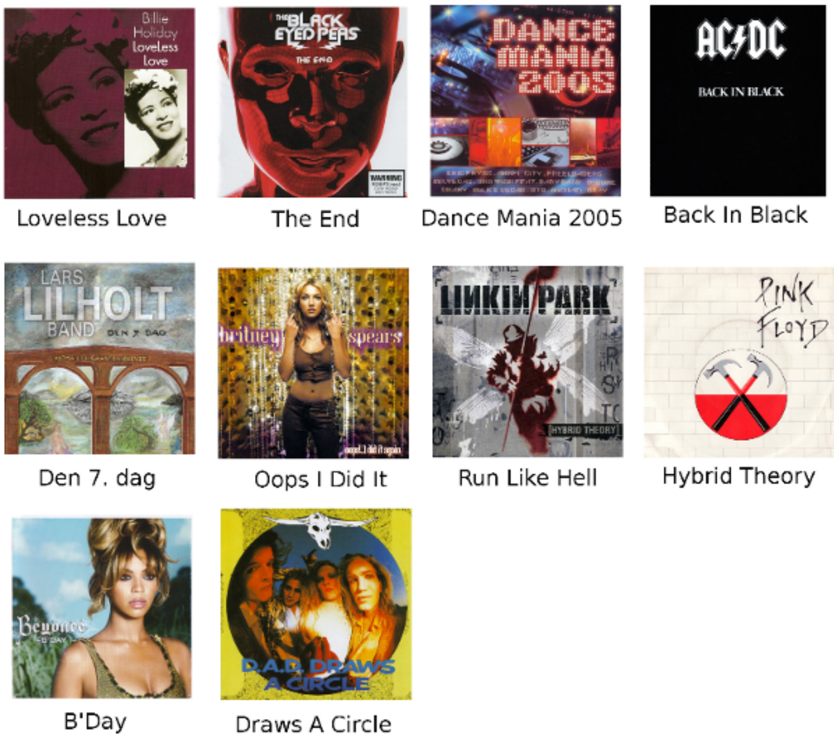


Figure 5.3: The pictures of the album covers were presented in order from left to right as shown in the figure.

5.2 Pilot studies

Before the experiment we performed some pilot studies. The pilot studies were done ad hoc and were not documented. The pilot studies helped us to secure the tasks were not too easy and not too hard. If the tasks were too easy we would not see any errors, and if it was too hard the test subjects might have gotten frustrated and failed to finish the tasks. We regulated the difficulty by adding or removing folders and music albums.

Another reason for running the pilot studies was to figure out how long time the test subjects needed to learn the location of the music albums. With less than three minutes the test subjects felt like they had to stress and with more than three minutes the test subjects would just be waiting for the time to run out.

Furthermore we conducted some pilot studies in order to make sure that the procedures in the experiment were working correct.

5.3 Questionnaire

The test subjects were asked one question before the experiment and five questions after. The test subjects were asked to rate the system they interacted with. In all questions the test subjects were told to rate themselves by making a mark on a line. The line was representing a scale was the end points were labeled. The questions were asked this way in order to produce data on a continuous interval scale. An interval scale has equal units of measurement and its zero point has been established arbitrary (Leedy & Ormrod 2010, p. 26). By using labels (if used correct) it is more likely that the test subjects will think of the intervals of having equal distances among its points. In order to use statistics that uses addition or subtraction the data needs to be measured on an interval scale (Leedy & Ormrod 2010, p. 27).

The first question asked before the experiment was:

How much do you know about music? (do you know many artists, hits, and albums?)
The end labels were: not very knowledgeable about music - very knowledgeable about music

Since we used music albums the knowledge about music might have helped the test subjects to perform better. Therefore we asked this question in order to be able to check for a difference in music knowledge between the two groups.

Below the five questions from the questionnaire are listed:

How easy was it to remember the location of the albums?
Very easy - Very hard

If it, for some reason, was not possible to measure how many mistakes the test subjects made we had the possibility to use their rating instead. This would only measure if they felt like it was easier to remember the location of the albums.

How satisfied were you with the efficiency of the system?
Very unsatisfied - Very satisfied

Even if the test subjects made fewer mistakes with the prototype they might not find it more efficient.

How good were your overview of where you were located in the system?
Really bad overview - Really good overview

This question was used to investigate if the mini-map had improved their overview.

Did it feel like you were psychically moving around in the system?
To a very low degree - To a very high degree

This question was asked to help us determine if the users had perceived the spatial dimension of the system.

How satisfied were you with the system in general?
Very unsatisfied - Very satisfied

This question was asked to see what their general opinions of the systems were.

5.4 Interview

The six questions were followed by a semi-structured interview (Kvale 1997, p. 133). The purpose of this interview was to clarify if the prototype system really did support spatial perception, if they understood how to use the side-folders and if they used the mini-map. The interview was used to cover aspects that could influence the results. This could be if they had misunderstood something or if they had any past experiences with categorizing music.

Before the experiment we prepared an interview guide (ibid). The interview guide consisted of the topics we wanted to cover and potential questions we could use to extract information from the test subjects about the topics. Since this was a semi-structured interview the questions were not fixed and could therefore change during each interview.

Interview guide

First we created a list of topics we wanted to cover with the interview. Afterwards we added potential questions to each topic.

1. Did they perceive the spatial dimension in the system? And did they use it to remember the position of the music albums?
2. Did they use the colors of the rooms and floors to remember the position of the music albums?
3. How much information about the categorization did they use and understand?
4. Did they use the side-doors? Did it help them? Was it intuitive?
5. Did they use the mini-map? And when did they use it? Did it help them from getting lost?
6. Did they have any former experience with categorizing music?

1. Did they perceive the spatial dimension in the system? And did they use it to remember the position of the music albums?
The reason why we wanted to investigate this topic was that we thought it could shed light upon whether they perceived the prototype as more spatial and if they used this spatial dimension to remember the location of the files. This would of course also be covered by the question about if they felt like they were moving around in the system, but we thought more questions about this could further validate our result.

The prepared questions were:

How would you describe the system?
-try to describe the items in the rooms?
-where were the items located?

Which technique did you use to remember the location?

2. Did they use the colors of the rooms and floors to remember the position of the music albums?
This topic was used to cover if they had used things like color of the rooms to remember the location of the music albums.

The prepared questions were:

Did you pay attention to the colors of the walls and floors?
-Did you see any system in the way the walls and floors were colored?

3. How much information about the categorization did they use and understand?
This topic was investigated for two reasons. The first reason was that we wanted to know how much of the categorization the test subjects had guessed and used. The other reason was that we wanted to secure that the differences we might find was not because of a better insight in the categorization compared to the insight in the categorization in the other group.

The prepared questions were:

Did you see any system in the way the music albums were placed?
-How do you think the music albums were categorized?

4. Did they use the side-doors and did it help them and was it intuitive?
This topic was included in order to investigate hypothesis two. To figure out if the test subjects had used the side-doors, if they were easy to use and if they had helped the test subjects.

The prepared questions were:

Did you use the side-doors in the system?
-Did you know where you were in the system when you had used a side-door?
-Did it help you to use the side-doors?

5. Did they use the mini-map? and when did they use it? Did it help them from getting lost?
This topic was included in order to investigate hypothesis three. We wanted to supply our observation data from the eye-tracker with information about what purpose they used the mini map for and if it prevented them from getting lost.

The prepared questions were:

Did you use the mini-map located in the lower right corner?
-Did it help you to navigate through the system?

6. Did they have any former experience with categorizing music? This topic was used to figure out how much experience the test subjects in each group had with categorizing music on their computer. The reason for this was that this might have had an effect on the result if this was not balanced between the two groups.

The prepared question were:

Do you have any experience with categorizing music on your computer?

5.5 Eye tracking

The eye tracker was included in the experiment because we wanted to be able to review the test subjects' interacting with the two systems. This was done for the purpose of being sure the correct number of errors were noted for each test subject. And also to confirm whether or not the test subjects were actually using the side-doors and mini-map. For elaborated information about the eye tracker and its functionality please see to Appendix F.

The eye tracker used in this study was the X120 Eye tracker manufactured by Tobii Technology[3] (Tobii 2012*a*). Figure F.3 shows a picture taken of the Tobii X120 Eye tracker.

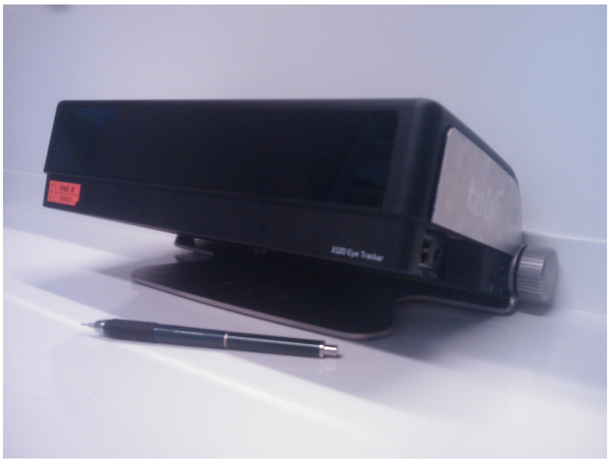


Figure 5.4: The X120 Eye tracker from Tobii.

The X120 Eye tracker is a stand-alone eye tracker unit which was mounted in a fixed position beneath the screen during the experiment. The eye tracker is capable of displaying where a person is looking, how long a person is looking, and in what sequence a person is shifting eye location. To track the eye movements the eye tracker captures the reflection of the user's cornea and center of the pupil with an infrared camera and its associated image software called Tobii Studio. The eye tracker's frame rate was in the experiment set to 120 Hz, i.e. the system will register 120 data points pr. second or every 8.3 ms (ibid). The frame rate could either be set to 60 Hz or 120 Hz. The frame rate of 120 Hz was selected because it would provide twice as many data points and thereby be more precise.

In order to track the eyes' movements most accurately a calibration process is needed. The calibration procedure is done by displaying a dot on a screen which the user must gaze upon for a certain amount of time. The tracking system can record the relationship between the center of the pupil and the corneal reflection as corresponding to a specific x,y coordinate on the concerned screen (Poole & Ball 2005, p. 3).

In order to visualize the fixations of eye movements they are illustrated graphically (Tobii 2012*a*). These graphical fixations are represented by dots, where the larger dots indicate a longer time of fixation. The saccades are indicated by lines between the fixations. In this study the time limit of the minimum length of a fixation was set to 60 ms which is set as default value by Tobii Studio (ibid). Thus fixations that have a duration under 60 ms will not be classified as fixations or graphically represented. Even though several studies have shown that 100 ms is appropriate for most eye tracker studies (Tobii 2012*b*), we estimated that the value of 60 ms was a proper time duration to gather information about the test subjects' fixations. This is because it is possible for viewers to gather information from fixations under 100 ms. In a study by Keith Rayner (Rayner et al. 2009) it is demonstrated that viewers can acquire information to understand the gist of a scene in just 40-100 ms (ibid). Therefore by setting the value to 60 ms, we secured that more information about the test subjects' fixations was collected by the eye tracker. It has however been considered that there might be a risk of information getting lost by neglecting fixation with length under 60 ms, because the test subjects could have fixated upon places as the mini-map e.g. in a duration of 40 ms.

[3] For more information about Tobii Technology visit: www.tobii.com.

In Tobii Studio it is possible to view the eye tracking session with a screen capture of a test subject (ibid). This makes it possible to review the whole session with the possibility of deeper analysis and at the same time be able to see what the test subject looked upon in the session in terms of fixation points (ibid).

Another useful feature in Tobii Studio is the gaze plot. The gaze plot feature displays the eye movement sequence, order, and duration of gaze fixation. The gaze plot makes it possible to see a specific part of a session or the entire session with the fixation dots. The dots are labeled with numbers, so the exact order of fixation points is possible to see. An example of a gaze plot can be seen in Appendix F.

5.6 Evaluation of the experiment

In this section we will evaluate the experiment and the two systems based on our observations during the experiment.

In general the test subjects had no problems with following the verbal instructions. The tasks were not too easy neither too difficult. The amount of time the test subjects had to interact with the systems seemed to be appropriate. However there were one test subject who failed to finish the tasks and the facilitator had to give verbal cues in order to help the test subject. Somehow the test subject must have misunderstood the tasks or how the system worked. Therefore in the further analysis we payed extra attention to this test subject.

As our independent variable we used two different systems - a control system and a prototype system. However, in the prototype system we both introduced a new way of representing a folder system, a mini-map, and side-doors. The problem with introducing all three elements in one system is that they might effect each other. This could have been avoided by making a system for each of the new features we introduced. However, each of these systems would also require a control system of their own. This would require six different systems and if we wanted the same amount of observations for each system it would mean that 120 test subjects were needed.

We implemented side-doors in the prototype system in order to test our second hypothesis. We were considering if this could effect the results of the first hypothesis. However, we came to the conclusion that the side-doors would not have any effect on how many mistakes the test subjects made since the side-doors are not making it easier to recall the location of music albums. It is just making it easier to get to relevant albums.

In order to test our third hypothesis we added a map in the prototype system. This map could actually have affected the number of errors the test subjects made in a positive direction, but a map is a symbolic way of showing elements' location in spatial relation. So it can be seen as a tool that uses the spatial locations to help the users. This does not conflict with the statement of our first hypothesis.

The design of the prototype system with the rooms instead of folders etc. might have been enough to enhance the test subjects' overview. It cannot be completely denied that this could affect the results. Therefore we asked the test subjects about their use of the mini-map. Furthermore the eye tracker made it possible to observe if they used the mini-map or not.

The prepared questions helped us to shed light upon the topics we wanted to explore. However, we did have some problems with clarifying if the test subjects saw the prototype system as more spatial than the control system. When asked to describe the systems they started to evaluate attributes like the usability instead. This made it really hard to evaluate if they had perceived it as spatial or not. Even when rephrasing a few times it was still not all test subjects who answered the question in the intended way. However, we could not have asked in other ways without starting to lead them to the answer.

After the experiment we checked the screen captures from the eye tracker, and we noticed that some of the recording files were corrupted. These recordings could therefore not be played. This was unfortunate but there were not much we could have done different in order to prevent this from happening.

Chapter 6

Results

In this chapter selected results from the experiment will be presented, followed by the statistical calculations of these results and a description of the used statistics. Later the qualitative data which provides information about the topics we have identified in the interview guide will be presented. All the quantitative and qualitative results from the experiment can be found in Appendix H in its original form.

6.1 Presentation of data

Regarding the first hypothesis Figure 6.1 shows a data chart of how well the test subjects performed the tasks in the prototype system. The chart shows how many errors each test subject made in every single task, the total number of errors of each test subject, and the overall number of errors in the prototype system. The errors is defined as when the test subjects went into the wrong room to find the concerned album. However, the counted errors does not implicate the situations where test subjects used the side doors as a shortcut to find the relevant album.

Sub/Task	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	Errors pr. sub.
1					1			1		2	4
3	1	4			1	1					7
5				1						1	2
7									1	1	2
9		1		1		1			1	1	5
11		6	2		1	1			3	1	14
13				1				1			2
15	1		1			1			1		4
17		2				1		1	2	1	7
19										4	4
21	1	4		1		1				1	8
23				1						1	2
25											0
27	5	3			1	1		1	1	1	13
29							1			3	4
31		2		1				5		1	9
33								1			1
35	1							1		2	4
37		2		4	15	6		8	3	4	42
39									1		1
										Errors total	135

Figure 6.1: The data chart shows how many errors the test subjects made in the prototype with the house metaphor when performing the tasks.

Figure 6.2 shows a data chart of how well the subjects performed the tasks in the control system. The chart shows the number of errors each test subject made in the respective tasks, the total number of errors of each test subject, and the overall number of errors in the control system. The errors the subjects made are defined as when the test subjects went into the wrong folder to find the relevant album.

Sub/Task	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	Errors pr. sub.
2	2	5	2							1	10
4	1								1	1	3
6						1			1	4	6
8				1			3	6	1	2	13
10		6			2		2		1	1	12
12		7	1	1	1					1	11
14	1	1	2								4
16			2					1		1	4
18	3	7		1	1			1		1	14
20		4			1			1	2	1	9
22		1			2					1	4
24	4			1	1				1	5	12
26			7	1				4			12
28											0
30	7		4		1					5	17
32	2	3		1	1	1	2	1	7		18
34	1				1	1		1	1		5
36		1									1
38				1				1	1	3	6
40		6	3	1	1		4	6		3	24
										Errors total	185

Figure 6.2: The data chart shows how many errors the test subjects made in the control system when performing the tasks.

The data for our second hypothesis is shown in Figure 6.3, which is a chart over whether or not the test subjects used the side-doors in the prototype system. The test subjects were asked if they used the side-doors. In order to be sure the test subjects actually used them it was necessary to see the screen captures of every single test subject. In the process of viewing the screen captures it was discovered that three screen recordings were corrupted. Therefore it was not possible to clarify whether or not the concerned test subjects actually used the side-doors other than by their verbal confirmation. The specific screen recordings of the test subjects are 11, 23 and 27. All the screen captures can be found in Appendix H.

The chart in Figure 6.3 shows if the test subjects' use of the side doors is confirmed verbally or/and by the screen captures.

Sub/Confirmation	Eye tracker confirmation	Verbal confirmation
Sub 1		
Sub 3	x	x
Sub 5	x	x
Sub 7	x	x
Sub 9	x	x
Sub 11	NA	x
Sub 13	x	x
Sub 15	x	x
Sub 17	x	x
Sub 19	x	x
Sub 21	x	x
Sub 23	NA	x
Sub 25		
Sub 27	NA	x
Sub 29	x	x
Sub 31	x	x
Sub 33	x	x
Sub 35	x	x
Sub 37	x	x
Sub 39	x	x
Number of confirmations	15	18

Figure 6.3: The chart shows how many test subjects confirmed they used the side doors - by verbal confirmation and confirmation by viewing the screen captures made by the eye tracker.

The data for the last hypothesis is shown in Figure 6.4, which is a chart over whether or not the test subjects used the mini-map in the prototype system. As with the side-doors the test subjects were asked if they used the mini-map. In order to be sure that they had in fact used it it was necessary to go through the screen captures of every single test subject. The screen recordings of test subject 11, 23 and 27 was not fit for reviewing due to its corruption.

The chart in Figure 6.4 shows if the test subjects' use of the mini-map are confirmed verbally or/and by the screen captures.

Sub/Confirmation	Eye tracker confirmation	Verbal confirmation
Sub 1	x	x
Sub 3	x	x
Sub 5	x	x
Sub 7	x	x
Sub 9	x	x
Sub 11	NA	x
Sub 13	x	x
Sub 15	x	x
Sub 17	x	x
Sub 19	x	x
Sub 21	x	x
Sub 23	NA	
Sub 25	x	x
Sub 27	NA	x
Sub 29	x	x
Sub 31	x	x
Sub 33	x	x
Sub 35	x	x
Sub 37		
Sub 39	x	x
Number of confirmations	16	18

Figure 6.4: The chart shows how many test subjects confirmed they used the mini-map - by verbal confirmation and confirmation by viewing the screen captures made by the eye tracker.

As mentioned, we used the screen captures to determine whether or not the test subjects used the mini-map. This was done by closely examining the screen recordings and utilizing the gaze plot feature[1]. Figure 6.5 shows an example of a gaze plot. The gaze plot in the following example indicates that the concerned test subject for a duration of time had several fixation points upon the mini-map.

Figure 6.5: The gaze plot indicates that the test subject had several fixation points upon the mini-map.

[1] For more information about the gaze plot feature see Appendix F.

The data from the test subjects' individual ratings from the questionnaire is shown in Figure 6.6. The figure shows a box plot for each of the six ratings from the questionnaire and of how many errors they made during the experiment - from the prototype system and control system.

Additionally, Figure 6.7 shows the means of the six ratings from both the prototype system and control system. The first row concerns the prototype system and the second row concerns the control system.

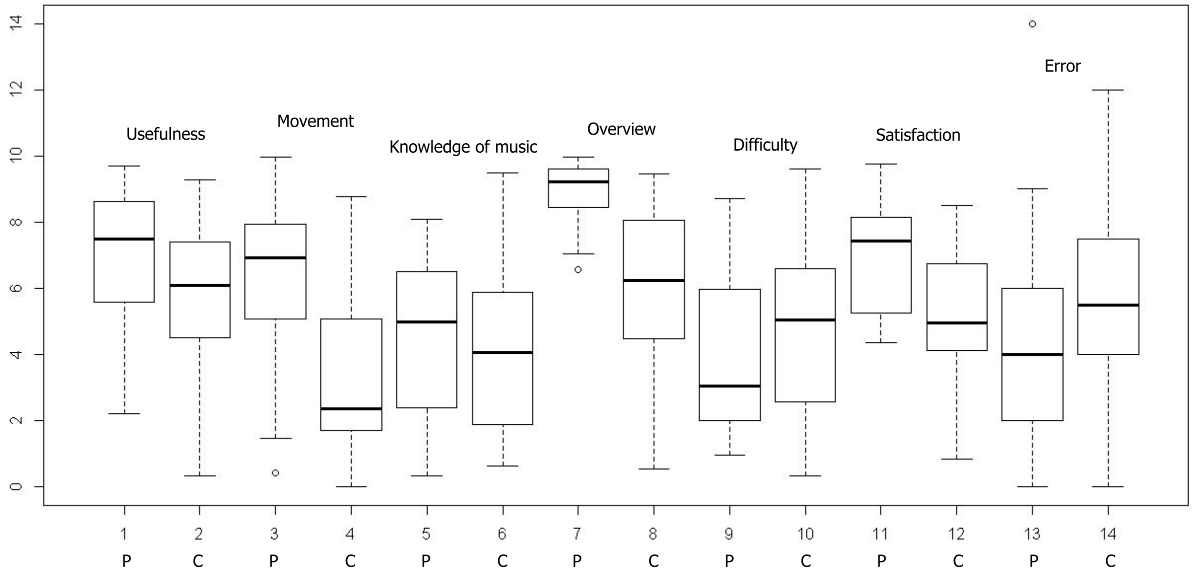


Figure 6.6: Boxplot for the test subjects' rating of the prototype system and control system. The prototype system is labelled by a "P" and the control system is labelled by a "C".

	Knowledge of music	Difficulty	Usefulness	Overview	Satisfaction	Movement	Error
Mean of con. system	4.17	4.87	5.60	6.01	5.10	3.36	5.85
Mean of proto. system	4.48	3.94	6.94	8.90	7.05	6.35	4.40

Figure 6.7: Means for the six ratings of the prototype system and the control system.

6.2 Data analysis

Quantitative data

The data from the test subjects' performance was numeric values of how many errors they made in each task. If one person totally misunderstood the tasks and each of his errors were counted he could have a big impact on our data. Furthermore after a few mistakes it was clear that the test subjects started guessing. It was therefore not relevant for the study of our hypothesis if a test subject made e.g. five or eight errors in a task but just if the test subject made no, one, or several errors in a task. We divided the number of errors into three intervals:

- If the test subject made no errors in a task, the number of errors was classified as "0"[2]
- If the test subject made one error in a task, the number of errors was classified as "1"
- If the test subject made two or more errors in a task, the number of errors was classified as "2"

Figure 6.8 shows the classification for the errors made in the prototype system. The classification for the control system can be seen in Figure 6.9. The original data can be seen in Figure 6.1 and Figure 6.2.

Sub/Task	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	Errors pr. sub.
1					1			1		2	4
3	1	2			1	1					5
5				1						1	2
7									1	1	2
9		1		1		1			1	1	5
11		2	2		1	1			2	1	9
13				1				1			2
15	1		1			1			1		4
17		2				1		1	2	1	7
19										2	2
21	1	2		1		1				1	6
23				1						1	2
25											0
27	2	2			1	1		1	1	1	9
29							1			2	3
31		2		1				2		1	6
33								1			1
35	1							1		2	4
37		2		2	2	2		2	2	2	14
39									1		1
										Errors total	88

Figure 6.8: The data chart shows the classification of the errors of each task the test subjects made in the prototype system.

[2] Due to reader-friendliness the zeros will not appear in the concerned charts but only be presented as an empty space.

Sub/Task	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	Errors pr. sub.
2	2	2	2							1	7
4	1								1	1	3
6						1			1	2	4
8				1			2	2	1	2	8
10		2			2		2		1	1	8
12		2	1	1	1					1	6
14	1	1	2								4
16			2					1		1	4
18	2	2		1	1			1		1	8
20		2			1			1	2	1	7
22		1			2					1	4
24	2			1	1				1	2	7
26			2	1				2			5
28											0
30	2		2		1					2	7
32	2	2		1	1	1	2	1	2		12
34	1				1	1		1	1		5
36		1									1
38				1				1	1	2	5
40		2	2	1	1		2	2		2	12
										Errors total	117

Figure 6.9: The data chart shows the classification of the errors of each task the test subjects made in the control system.

Because the data in the classification is based on counts the data can be seen as a Poisson distribution, and is therefore best treated as non-parametric data. For data that is seen as scores and non-parametric, it is most suitable to use a Wilcoxon test. The Wilcoxon test, also named the Mann-Whitney test, is a non-parametric method that uses the rank of the samples to measure a p-value (Agresti & Finlay 1997). The Wilcoxon test is the equivalent of the independent t-test and is used for testing differences between groups with two levels of an independent variable in a between subject design. The test ranks the data from lowest to highest. The test then calculates the mean of the rank in each group (Field & Hole 2003, p. 237). In our case, the two groups was the total number of errors pr. test subject in the two systems after the classification.

Wilcoxon test
Difference of errors between the two systems
W = 262, p-value = 0.094

With a p-value = 0.094 we cannot reject the null hypothesis. Therefore two groups can be considered equal. However, in the data processing we noticed that test subject 37 had a significant higher number of errors, see Figure 6.1 and 6.8, and is marked as an outlier in Figure 6.6. During the experiment the discussed test subject did not manage to complete the tasks without help from the facilitator. We further examined our recordings of the test subject. By a close examination of the screen capture and sound recording of test subject 37 it can be concluded that the test subject in some way must have misunderstood the tasks. Therefore test subject 37 was excluded for further statistic analysis. Without the data from test subject 37 the the two-sample Wilcoxon test presents a p-value = 0.043.

Wilcoxon test
Difference of errors between the two systems
W = 262, p-value = 0.043

With a p-value = 0.043 we can reject the null hypothesis and instead confirm the alternative hypothesis saying that there is a difference between the two groups. When compared to the means in Figure 6.7 it can be seen that the mean of the test subjects' total errors in the group who interacted with the prototype system is lower than the mean from the group who interacted with the control system group. It can therefore be stated that the test subjects, who used the prototype system were better at recalling and finding the albums than the test subjects who used the control system.

To clarify whether or not there were differences between the performance of male and female test subjects a two-sample Wilcoxon test was conducted, both of the data from the prototype system and control system respectively.

Wilcoxon test
Difference of errors between male and female in the prototype system
W = 57, p-value = 0.34

Wilcoxon test
Difference of errors between male and female in the control system
W = 56, p-value = 0.55

Since none of the p-values are lower then 0.05 we could not detect any differences between males and females in any of the systems. We could not reject the null hypothesis which predicted that there were no differences between the performance of male and female test subjects.

To investigate whether or not the test subjects' knowledge of music were significantly different between the two systems, we had to compare the test subjects' rating of their knowledge of music in the two systems. The ratings were designed as an interval scale, which is a scale where equal intervals represent equal differences in the property being measured (Field & Hole 2003, p. 8). Because the test subjects rated their knowledge on an interval scale, it was possible to calculate a useful mean of the data which is used in parametric tests. However, parametric tests require the data to be normally distributed. In order to determine if the data is normally distributed we used the Shapiro-Wilk test. The Shapiro-Wilk test compares the scores in the sample to normally distributed scores with the mean and standard deviation. If the test gives a p-value higher than 0.05 the data is not significantly different from a normal distribution and can therefore be seen as normal distributed (ibid, p. 159). The Shapiro-Wilk test is performed on both the prototype and control system respectively.

Shapiro-Wilk normality test
Difference from normal distribution for music in the prototype system
W = 0.94, p-value = 0.34

Shapiro-Wilk normality test
Difference from normal distribution for music in the control system
W = 0.95, p-value = 0.43

Since both of the p-values are higher then 0.05 the data is not significantly different from a normal distribution and therefore we could consider our data to come from a normal distribution. Because our data is represented by two groups and is parametric, a t-test is most suitable to use when comparing the two groups. When using a t-test on data from a between subject design the test is called an independent t-test (ibid, p. 163). A t-test calculates a t-value by comparing the differences between the observations in the two groups one by one. The t-value is then compared with a t-distribution in order to calculate the p-value (Dalgaard 2008, p. 100). The t-distribution is a bell-shaped distribution which are symmetric around zero. The spread of the t-distribution depends on the degrees of freedom. As the degrees of freedom increase the spread of the distribution decreases (Agresti & Finlay 1997, p. 181). A Welch t-test is an independent t-test and was used on our data.

Welch t-test

Difference of knowledge of music between the two systems
t = -0.62, df = 37, p-value = 0.54

Since the p-value = 0.54 we could not reject the null hypothesis which predicted that there was difference between the two groups. This means that there was no significant difference in the knowledge of music between the two groups.

To check if there were any difference between the test subjects' ratings of which degree of movement they felt when they navigated in the systems, we had to compare the ratings for both of the systems. Since the ratings were designed as an interval scale we first had to determine whether or not our data came from a normal distribution. We used a Shapiro-Wilk test on the data for both the systems.

Shapiro-Wilk normality test
Difference from normal distribution for movement in the prototype system
W = 0.90, p-value = 0.049

Shapiro-Wilk normality test
Difference from normal distribution for movement in the control system
W = 0.92, p-value = 0.12

The p-value of the data from the prototype system is lower than 0.05 and is therefore significantly different from a normal distribution. Then the data can not be considered as a normal distribution. The p-value from the control system is higher than 0.05 and the data is therefore not significantly different from a normal distribution and can then be considered as a normally distribution. Since the data for the prototype system cannot be seen as normal distributed it is not suitable to use a t-test to compare the two groups. As a result of the Shapiro-Wilk test data from both the groups must be seen as non-parametric data and therefore a Wilcoxon test is more suitable.

Wilcoxon test
Difference in movement between the two systems
W = 75, p-value = 0.0013

With the p-value = 0.0013 the null hypothesis is rejected, which means there were significant differences between the two groups. By this matter of fact and by comparing the means in Figure 6.7, it is safe to say that the test subjects who interacted with the prototype system felt in a higher degree that they psychically moved around in the system than the test subjects who interacted with the control system.

If there were any differences in the test subjects' rating of how well an overview they had of the systems, it was necessary to compare the ratings of overview of the prototype system and the control system. As these ratings also were designed as interval scales it must first be determined whether or not the data for the two systems were normally distributed. Shapiro-Wilk test was used on data from both the systems.

Shapiro-Wilk normality test

Difference from the normal distribution for overview in the prototype system
W = 0.90, p-value = 0.041

Shapiro-Wilk normality test
Difference from the normal distribution for overview in the control system
W = 0.96, p-value = 0.46

The p-value from the data of the prototype system is lower than 0.05 and can therefore be seen as significantly different from a normal distribution. The p-value from the control system is however higher than 0.05, and is not significantly different from a normal distribution. Because the data of the prototype cannot be seen as normal distributed a Wilcoxon test is preferred to use when comparing the two systems in connection with the subjects' rating of how well an overview they had when navigating in the systems.

Wilcoxon test
Difference of overview between the two systems
W = 42, p-value = 3.39e-05

The p-value of the data equals 3.39e-05. This means that we can reject the null hypothesis which predicts that there were no differences between the two groups. This fact can be compared with the means in Figure 6.7. This shows us that the overview in the group who interacted with the prototype system is higher than in the group that interacted with the control system. It can therefore be stated that the test subjects who interacted with the prototype system felt they had a better overview of the system than the test subjects that interacted with the control system.

To investigate if there were any differences between the test subjects' rating of how satisfied they were with the usefulness of the systems, we wanted to compare the test subjects' rating of both the systems. Again the ratings were designed as interval scales and therefore it was first required to determine if the data was from a normal distribution or not.

Shapiro-Wilk normality test
Difference from the normal distribution for usefulness in the prototype system
W = 0.90, p-value = 0.055

Shapiro-Wilk normality test
Difference from the normal distribution for usefulness in the control system
W = 0.94, p-value = 0.20

Since both the p-values are higher than 0.05 data cannot be considered as significantly different from a normal distribution and can therefore be seen as normal distributed. Because of this the Welch t-test is used for comparing the two groups.

Welch Two Sample t-test
difference of usefulness between the two systems
t = -1.60, df = 37, p-value = 0.12

Because the p-value is not lower than 0.05 we cannot reject the null hypothesis which predicts no differences between the groups.

To clarify any differences between the test subjects' ratings of their general satisfaction of both systems we compared the two data sets from the ratings of both systems. An interval scale was also used for this. A Shapiro-Wilk test was used on the data from both systems to see if they came from a normal distribution.

Shapiro-Wilk normality test
Difference from the normal distribution for satisfaction in the prototype system
W = 0.93, p-value = 0.18

Shapiro-Wilk normality test
Difference from the normal distribution for satisfaction in the control system
W = 0.95, p-value = 0.44

Since both of the p-values are higher than 0.05 we could consider both data sets to be normally distributed. Welch t-test is used to compare the data from both systems since they can be seen as normal distributed.

Welch Two Sample t-test
Difference of satisfaction between the two systems
t = -2.95, df = 35, p-value = 0.0055

Because the p-value equaled 0.0055 we could reject the null hypothesis which predicted that there were no difference between the groups. When comparing this fact with the means in Figure 6.7 it can be seen that the mean for satisfaction in the group who interacted with the prototype system is higher than the mean for satisfaction in the group who interacted with the control system. It can therefore be stated that the test subjects who interacted with the prototype system felt higher satisfaction than the test subjects who interacted with the control system.

Qualitative data

The interviews were first transcribed and afterwards we extracted the statements that could give information about the topics we wanted to investigate. From these statements we could analyze what the opinions and thoughts were about each topic. In the following section we will discuss the results from the analysis of the interview and the observations. All the sound recordings and the transcription from the interview can be found in Appendix H.

Did they perceive the spatial dimension in the system? And did they use it to remember the position of the music albums?

Besides confirming or rejecting the hypothesis we wanted to check the validity, if we actually measured what we wanted to measure. To verify the systems we wanted to know if the concerned test subjects actually perceived the prototype system as a three-dimensional space and that the test subjects saw a folder system as the underlying basis in control system. Therefore, the first question in the interview was regarding how they would describe the system. We observed that when test subjects were asked this question, many of them did not immediate answer in the way we originally intended. The objective of the question was that the test subjects would describe the properties of the prototype system that constitutes the spatial dimensions and describe the control system as the well-known folder hierarchy. However, some test subjects interpreted the question such as they should rather answer the question by describing it in terms of how good or bad they thought the system was, e.g. if they felt it was confusing. The following quotes was stated during the interview from a test subject who interacted with the prototype system and from a test subject who interacted with the control system.

> *"It was a bit somewhat confusing. It was not just to remember where you were going and what to look for."* - prototype system

> *"Very simple or plain"* - control system

Others answered the question by describing the categorization and placement of albums:

> *"I do not know. I thought it a bit as different categories. The room in the middle was rock and the left was a little more mainstream, teen music and the right was a little more older generation music like Lars Lilholt and Pyrus. Perhaps slightly more diffuse compared to the other two rooms."* - prototype system
>
> *"I would think that it was divided into some categories. Some rock and some pop and some collections of one form or another. I did not quite acquaint myself with how subfolders worked. But in one of the subcategories, there was something heavy and punk. But otherwise... I do not know."* - control system

In most of the cases when the facilitator evaluated that the test subjects had another interpretation of the question, the facilitator rephrased the question but still there were some test subjects who did not answer in the direction the question immediately intended. Examples of this can be seen in the following quotes from a test subject both from the interaction with the prototype and control system:

> *Test subject: "It was just to remember where you saw it."*
>
> *Facilitator: "If you try to describe this system for someone who has never used it. How would you describe it? How does it look?"*
>
> *Test subject: "Then I will start by saying there are 3 doors to choose from and inside the doors and in the chambers there are albums. At the same time there are doors you can move into."* - prototype system
>
> *Test subject: "It was a bit somewhat confusing. It was not just to remember where you were going and what to look for."*
>
> *Facilitator: "If you had to describe it a little more simple how it looked and what it actually was, how would you describe it?"*
>
> *Test subject: "It looked like a very ordinary Windows Explorer. That is how I would describe it. Windows Explorer which is about 20 years old."* - control system

Even though the interviewer tried to ask different questions about it it was really hard to make the test subjects give a description of the system instead of their opinion about it. This limited the level of details we received from their description. We hoped to see a difference in the degree in which they would describe the two systems with spatial metaphors like: Entering, going into, moving around etc. However in the description of both systems the test subjects used spatial metaphors to describe it, and since the detail information were so limited we were not able to distinguish the description of the two systems.

We could conclude that the majority of the test subjects who interacted with the prototype system described the system as build upon different rooms which contained different albums, and that you could enter the different rooms by moving through the doors. Also, the majority of the test subjects who interacted with the control systems described the control system as the well-known folder hierarchy.

Did they use the colors of the rooms and floors to remember the position of the music albums?

The interview also revealed that most of the test subjects said no when asked if they could see any system in the coloring of the floor and walls and they were also unable to recall what color that was used in the rooms. By judging from the test subjects' statements in the interview there was nothing that indicated that the test subjects used the colors of the rooms to remember the location of the music albums. The following quotes was stated during the interview from test subjects who interacted with the prototype system:

> *Facilitator: " Did you notice the coloring of the floor?"*
>
> *Test subject: "No, I actually did not"*
>
> *Facilitator: "Did you notice if there were any system in the coloring?"*
>
> *Test subject: "There were probably one."*
>
> *Facilitator: " Did you notice the coloring of the floor?"*
>
> *Test subject: "No, not exactly but something was brown and the doors were brown."*

How much information about the categorization did they use and understand?

The majority of the test subject using both the control system and the prototype system had a general clear conception of how the albums were categorized and where the categories were placed. They could not necessarily name the exact albums in all of the rooms but they had a mutual understanding about the first room/folder leading into three rooms/folders from left to right with pop music, rock music and mixed respectively. Some test subjects did however also comment on the further categorization in the six rooms. Overall it did not seem like there were any differences between the two groups' insight in the categorization. The following quotes are from test subjects' statements made during the interview:

> *"It took some time before I found out it was divided into categories, but when I found out I used it, so you knew that there was pop to the left-and rock in the middle and then mixed to the right."* - control system
>
> *"I would describe it as albums that were divided in some room where I could go and find albums in relation to the rooms. [...] I had divided so there was rock in the middle and then there was pop over to the left and then there was a mix on the right."* - prototype system

Did they use the side-doors? And did it help them? And was it intuitive?

The test subjects were asked if they used the side-doors. Several test subjects stated how the side-doors made it faster to reach the relevant room. The use of the mini-map gave an overview of where the side-doors lead to. Furthermore, the majority of the test subjects answered that they were not in doubt about where they were after using the side-doors. These statements were confirmed by our observations that showed that the test subjects used the side rooms without problems without any instructions about them.

The following quote was stated during the interview with test subjects that interacted with the prototype system:

> *Facilitator: "Did you use the side-doors in the system?"*
>
> *Test subject: "Yes I did"*
>
> *Facilitator: "Did you have an overview of where you came from when you used it?"*
>
> *Test subject: "Yes ... and then there was a nice overview down in the bottom corner to just find out where you were. It was fine with those shortcuts"*

Did they use the mini-map? And when did they use it? Did it help them from getting lost?

The majority of the test subjects who interacted with the prototype system stated that the mini-map contributed information to where they were in correlation to the rooms. The majority also stated that not at any point were they in doubt of where they were in the system. The following quote was stated during the interview by test subjects who interacted with the prototype system:

> *Facilitator: "Did you use the mini-map?"*
>
> *Test subject: "Yes, I used it the whole time."*
>
> *Facilitator: "Did it help you?"*
>
> *Test subject: "Yes."*

The facilitators also noted that during the examination of the screen recordings, several test subjects made glances at the mini-map and would navigate to the relevant room. The mini-map were in general used by almost every test subject.

Chapter 7

Discussion

As one of the first steps in the data analysis we removed an outlier from our observations. This is of course always a critical thing to do since it could potentially lead to manipulation of the results. However, in this case the removed test subject were not able to finish the given tasks on her own and needed help from the facilitator to do so. The test subject was either nervous or had misunderstood something in the introduction since the rest of the test subjects finished the tasks with relative ease.

With a p-value that equals 0.043 we found that there was a difference in how many mistakes the users made in the two systems. By looking at the box plot in Figure 6.6 we can confirm that it was the test subjects using the prototype system that made significantly fewer errors than the test subjects in the control system. Furthermore with a p-value = 0.0013 we saw a difference between the two systems in how they rated the perception of physically moving around. Again by looking at the box plot in Figure 6.6 we could confirm that the perception of psychically moving around was significantly better in the prototype system.

From the interview we could conclude that the test subjects in both systems used metaphors that indicated that they were thinking about both systems as an entity with a spatial dimension, where you can move in and out of rooms/folders. However, we could not determine if there was a difference in the degree of their perception of the spatial location.

The results show that we did see the test subjects perform significantly better with the prototype system than the control system. The hard thing is to tell whether or not it was because of the perceived spatial dimension. As we wrote above we did not see a clear differentiation in the perception of the spatial dimension between the two groups. The reason for this was that they used metaphors that indicated that they also thought of the control system and its folders as rooms with actual space and with a location in the system. This is actually obvious since people would create their own virtual space even if they are only provided a direction name (Fertig, Freeman & Gelernter 1996). Even if the spatial dimension is not provided, the users will still think of it as being there. If there is nothing that supports the spatial dimension it could lead to trouble with e.g. recalling files.

While saying this we can not tell for sure if they actually thought of the prototype system as a system with a better spatial dimension than the control system. However, the test subjects did rate the prototype as giving a better perception of moving around in the system. This emphasizes that the test subjects did in fact perceive the prototype system as more spatial than the control system and in connection with the variation of errors between the two systems, this confirms our first hypothesis.

Another concern is that even though they might have perceived the prototype as more spatial the less errors could be due to another variable. Therefore we checked for several other variables that could influence the result.

We checked for a difference between sexes both in the control system and the prototype system. With a p-value = 0.34 for difference between sexes in the prototype system and a p-value = 0.55 for difference between sexes in the control system there was no reason to suspect that sex could influence the results. We asked the test subjects to rate their knowledge about music and with a p-value = 0.43 we could conclude that there was no difference in music knowledge between the groups.

In the interview we asked the test subjects if they had noticed the colors on the walls and floor. Only a few answered that they had. The few who did say they noticed it were not able to recall what color the different rooms had. As a last control question we asked the test subjects about the categorization of the music albums. We found that the idea of the categorization was pretty even between the two groups. So none of the aforementioned factors seemed to effect the results.

We implemented side-doors in the prototype system in order to test our second hypothesis. We were considering if this could effect the results. However, we came to the conclusion that the side-doors would not have any effect on how many mistakes the test subjects made since the side-doors are not making it easier to recall the location of music albums. It is just making it easier to get to relevant albums.

From the observations of the test subjects' interaction with the prototype system we could tell that almost every test subject had used the side-doors. We could also see that in most cases they used the side-doors from the second they started using the system and it seemed like they did so with no doubt about how the side-doors worked. We also observed that they used the side-doors correct to get to a room faster.

Our observations of the use of the side-doors were confirmed through the answers in the interview. Again almost every test subject stated that they had used the side-doors. This made them in full control of where they were in the system after they had used the side-doors. Furthermore the test subjects stated that the side-doors helped them to get to the desired room faster and with fewer mouse clicks.

There was nothing in the statements from the interviews that lead us to believe that anybody was in doubt what the purpose of the side-doors were and how to use them. People that are used to using a regular folder system is not used to the option of going to a side-folder. Test subjects did not receive any instructions about how the side-doors worked, however for the test subjects it seemed very natural to use them. It was easy for the test subjects to adapt to the new functionality even though the only thing that helped them was a suitable mapping hence the house metaphor.

In order to test our third hypothesis we added a map in the prototype system. This map could actually have affected the number of errors the test subjects made in a positive direction but a map is a symbolic way of showing elements location in spatial relation. So it can be seen as a tool that uses the spatial locations to help the user. This does not conflict with the statement in our first hypothesis.

We asked the participants to rate their overview of the system. The Wilcoxon test gave a p-value = 3.39e-05. By comparing this result with the mean of the two groups we could conclude that the test subjects felt that there was a better overview in the prototype system.

From the observations from the eye tracker we could see that almost every test subject had used the mini-map. These observations were confirmed by the interview where the test subjects stated that they used the mini-map to get an overview and they did so without any problems. The mini-map also helped them to see where they were located.

The results showed that the test subjects who had interacted with the prototype system had a significantly better overview than the test subjects who had interacted with the control system. From this fact we cannot tell whether or not the better overview is due to the mini-map. However the observations from the eye tracker and the statements from the interviews show us that the mini-map in fact did help them get a better overview.

The design of the prototype system with the rooms instead of folders etc. might have been enough to enhance the test subjects' overview and it cannot be completely denied that this could affect the results. However with the strong statements from the test subjects and with the observations from the eye tracker it is safe to say that the mini-map has had an effect on the overview of the test subjects.

We also asked the test subjects to rate how satisfied they were with the systems in general. With a p-value = 0.0055 we could conclude that there was a significant difference between the two groups' general satisfaction. By looking at the means in Figure 6.7 we could see that the test subjects preferred the prototype system.

The reason for this difference can probably be found in the synergy between the elements in the prototype system: The spatial dimension which makes the users make less mistakes, the mini-map that helps them with a better overview, and the side-folders that help them with easier access to the files in the system.

Chapter 8

Conclusion

By providing a user interface with more support for the spatial perception we successfully showed a better performance in recalling the location of music albums. Besides a bit unclear picture from the interviews there was nothing that gave us reason to doubt that this result came from the improved support for the perception of the spatial dimension. We can hereby confirm the first hypothesis saying:

In a system that supports the objects' spatial locations in a perceived three dimensional space, people will be better at recalling and finding the objects located in the system

We saw that the test subjects used the side-doors to take a shorter route to the music albums compared to the route they would have had to take if the system had not included side-doors. Furthermore, they did not seem to have any problems using them. These findings were confirmed by the statements in the interview. We can hereby confirm the second hypothesis saying:

A system that allows moving to side-folders will shorten the route to the file people are searching for

By adding a mini-map to the prototype system we added context-information about where the test subjects were in relation to other places in the system. By adding the mini-map the test subjects felt that they had a better overview of where they were and it prevented them from getting lost. Hereby we can confirm our third hypothesis:

A system that gives contextual information about your location in the system will give you a better overview

The working thesis of this project was: "*How can the information about our semantic memory and our conceptualization and categorization of objects help us in order to create a graphical user interface that better support recall and help us find our files?*"

Throughout the literature study we found that one of the most important aspects in how we conceptualize and categorize elements in our semantic memory is by their spatial information. From that point we worked with how we could implement that in a graphical user interface. We ended up building the system we refer to as the prototype system. By confirming our hypotheses we have confirmed that it is possible to create a system that supports our recall of the files in the system better.

To the fact that we confirmed all our hypotheses it is important to add that the way we implemented the spatial dimension, the possibility to move to the side-folders and the contextual-information in form of the mini-map played an important role in our results. We cannot be sure that any other ways of implementing those things will work as well as using a house metaphor, side-doors and a mini-map.

Chapter 9

Further work

Through this study we have shown that it is possible to create a graphical user interface that enhances the support for recall by adding a spatial dimension in the system. As stated in the previous chapter it is important to consider how we implement the spatial dimension.

In this study we have used a house metaphor. We showed that this enhanced the perception of the spatial dimension compared to a common folder system. From this result we cannot tell if it was possible to create a system that even further enhanced the perception of the spatial dimension. In our prototype system the visualization of the music albums was a bit flat. They were all placed in the middle of the room instead of using the foreground and background of the room. We could have done more to support both the visual perception system and the visual action system (Goodale & Milner 2006). The visual action system could be better supported if the users instead of clicking on the doorway were asked to click on the door. The door could give visual animations of opening and the users could then enter the room. However, we cannot just blindly throw in a mapping that supports the spatial metaphor. The way we implement the spatial metaphor is very important. If the users had to click on a door and wait five seconds for the animation to finish, the users would be really impatient. If the users afterwards entered a room where the music albums were located all over the place it would probably lead to frustration. This system might be great for remembering where you have located your files, but the efficiency however would be really poor.

We asked the test subjects to rate how efficient the systems were. We could conclude with a p-value = 0.12 that there was not a significant difference between our prototype system and the control system hence our system was not more efficient than the control system. This shows us that just because you make fewer errors with a system it does not mean that people find it more efficient.

The house metaphor might be sufficient for the test scenario but if it were to be implemented in a fully functional system it would have some flaws.

Music albums located on the floor does not make much sense compared to the real world scenario and users of the system would probably wonder why they were placed there. In our system there were only 38 music albums which is not much compared to the amount of files most users have on their computers. Our prototype system would not be able to contain all these files as it is now.

Another issue with the prototype is that if you want to add a lot of extra rooms the rooms will appear really small on the mini-map or be excluded. And what if you wanted to move the rooms to another location? How will the doors then be connected? What happens if we move the files etc.

These things have to be considered when designing a system. By mapping a system using a metaphor users will create a mental model about the system which is based on the metaphor. Therefore users will act according to the metaphor and if the system does not work that way the users will make mistakes or be confused (Norman 1999).

Which kind of metaphor that is most useful to map a system while supporting the spatial dimension is out of the scope of this project. It is something that would require further work and probably also have to be adapted to the functionality and the concept of the system.

In this study we have only worked with a system that could work as a desktop folder system but our results could also be used in menu-structure as the ones in software programs. The focus in software programs probably has been on how to group the different functions, and buttons instead of thinking about how they are placed in relation to each other. It is also worth considering how we could map the perception of moving to a sub-menu better. The structure of web sites could be improved with focus on the spatial location and perception of the elements on the website. Furthermore the search on the web and any other program with a complex structure could benefit from more information about where you are according to the context.

Through our findings we can conclude that not only should we care about what the elements in a graphical user interface look like, how they work, and what we call them, we should also consider their spatial location and how we map the perception of space.

Appendix

Appendix A

Field studies about categorization

A.1 How do people organize their desks? implications for the design of office information systems

Based on a series of interviews of office workers, Thomas W. Malone (Malone 1983) has studied how to design a computer-based information system (ibid).

Ten office workers participated in the study. The participants were asked to give a tour of their office explaining why and where they were keeping their information. In the end the participants were asked to find selected documents in their office (ibid).

One of the strongest observations in the study were the variation in how precisely organized their offices were. Furthermore people with messy offices seemed to be embarrassing about it. So apparently people perceive higher social value from having a neat office (ibid, p. 105).

He discovered that office workers organized their information in two different ways some organized the information in piles some in files. The piles were simply a bunch of papers lying in a pile on the desktop or somewhere else in the office. Others had their information organized in a system with folders and titles which Malone referred to as files (ibid).

So the participants could not receive information about the documents in the pile from a title and the piles were not necessarily arranged in any particular order. Instead the participants received information from the spatial location of the piles (ibid, p. 106).

The use of piles spatial location is not the only thing people get from arranging their information in piles, the piles also serve as reminders for tasks the participants need to carry out in the future (ibid).

Another reason for using piles is that classifying documents into a file system is time demanding. People had a lot of trouble with selecting the right categories and deciding which category a given document should be filed in (ibid, p. 107).

The conclusion of his findings is that since a computer system can easily classify documents (create folders, labels, and put documents into these folders etc.) this could help the participants' classification problems. Furthermore the computers could also automatically classify your information. Documents could be automatically classified using title, author, and so forth (ibid, p. 108).

A computer system could also give the users the opportunity to create piles on the computer, so they can use the spatial location and do not have to title the information explicitly (ibid).

A.2 The psychology of personal information management

Mark Landsdale (Lansdale 1988), a psychologist that works with categorization, has conducted a study which includes comments on Malone's work (Malone 1983).

As it was mentioned in Malone's study people tend to have their documents lying around on their desk. This is either because they do not trust themselves that they will be able to retrieve the documents or because they do not know how to categorize them (Lansdale 1988, p. 66).

A problem with categorizing files, documents, and objects is that if you have a system where you select different regions of interest in a hierarchy, you make too many mistakes. An error could be to select the wrong regions of interest because some of the labels are ambiguous. Also some of the ideas or expressions come with different names which means that you might search for the wrong expression or name (ibid). Another issue is that when categorizing you also cut off some information. For instance, if you have a document about health insurance you can either store it under health or insurance. If you store it under health you miss the information about insurance if you store it under insurance you miss the information about health (ibid, p. 57).

Another problem with today's user interfaces is that they do not support the way we recall information. We remember the meanings of events, not the details. We interpret information in the context we learn it when we store it in our memory. When we want to retrieve the information we have to think about things related to the interpretation. The key to recall is therefore the context of the word hence we need cues from the context to remember and recall information (ibid, p. 58).

We can also enhance recall in user interfaces by supporting it with better visual cues (ibid). Different studies have concluded that humans are better at remembering pictures than words (ibid). Other studies indicate that if you also include information about the spatial location you are even better at retrieving information. Which means that spatio-visual cues seems to work better than verbals (ibid). It seems like users remember chronological information about their files, like when they created them etc. (ibid, p. 63).

So far it have been described which elements that can support recalling, but one of the most important thing about these observations is that users remember form, color, and location independent (ibid). Hence a system based on remembering all three attributes fails but if you only need to rely on one of the attributes and all of them are present, it helps the user.

In some systems you create your own categories to store information in, and in some systems you use pregenerated categories. The problem with such a system is that if the system generates cues and materials for people, they do not remember them as well as if they had created them themselves (ibid).

A.3 File Organization from the Desktop

The study consist of a comparison between two studies one made by Deborah Barreau and one made by Bonnie Nardi. Barreau studied a mixed group of DOS and Windows users. Nardi studied Macintosh users (Barreau & Nardi 1995, p. 39).

Barreau observed and interviewed seven managers who were given the task of organizing and retrieving information from their electronic work space. The goal of the study were to determine factors affecting individuals' decisions to acquire, organize, maintain, and retrieve information (ibid, p. 39).

Nardi's study included 15 participants in an interview study. Users were asked to give a tour of their systems and a structured set of questions were asked in a conversational style to elicit information about jobs and tasks as well as approaches to organizing and finding files (ibid).

The comparison showed the following similarities (ibid, p. 40):

- A preference for location-based search for finding files
- The critical reminding function of file placement
- The use of three types of information: ephemeral, working and archived
- Lack of importance of archiving files

Barreau and Nardi differentiate between location-based search and logic search. In location-based search the users take a guess at the direction/folder or diskette where they think a file might be located and goes to that location and then browses the list of files until they find the one they are looking for. A logical search is a text search for the file (ibid).

The observations showed that people often separate different types of files in different folder or on different diskettes in order to better remember where they have placed the different files (ibid). Furthermore some users grouped the folders by placing them in different locations on the desktop. For instance users placed files that were soon to be deleted near the trash can. This indicated that users preferred location-based findings (ibid, p. 41).

Another observation were that users liked to keep files, that needed to be taken care of soon, in the top direction of their desktop or directly on the desktop because of it's crucial reminding function (ibid).

They did not use the search function because often they could not remember the name of the file. Also it seemed like users preferred to be in control of the search instead of waiting for the computer to come up with a list of files which may or may not have been relevant (ibid).

Other information that Bearrau and Nardi found in their observations was that the users worked with three types of information: ephemeral information, working information, and archived information. Ephemeral information contains to-do lists, mails, notepads, newly downloaded articles etc. The users liked to keep this kind of information on the top level of the folders on the desktop (ibid).

Working information is documents relevant for the users current work and frequently used often organized and categorized and is easy to find (ibid).

Archived information is infrequently used and exists for month or years, and is often hard to find again. The reason for this were that the users did not want to spend the necessary time to achieve files. This was because they were not sure if they were going to use the files again (ibid, p. 42).

A.4 Finding and reminding reconsidered

A year after Barreau and Nardi published their work (ibid). Another research group made a review of their article and claimed to have discovered some interesting findings (Fertig, Freeman & Gelernter 1996).

They believed that Barreau and Nardi draw the wrong conclusion. The conclusion about separating the users in three groups was considered valid. The rest of the conclusions were explained due to the narrow scope of the study. Since both the Windows desktop and the Macintosh desktop favors the use of location-based search Fertig believed this explained why they found that users used the location-based approach (Fertig, Freeman & Gelernter 1996, p. 2).

Location-based search is based on memory of past events which can be error-prone, also the subjects only had two ways of gaining the information, location-based search, and logical search. Perhaps they would have chosen another way if they had an other solution. Also a better/faster search function might have changed the results (ibid).

Instead they focused on the time dimension and that users often use the information about how old their files are and when they achieved them. This was done in order to recall where they had placed them. Based on this observation they have created a new UI based on a metaphor of a time machine named Lifestream The idea is that you can move forward and backwards through time in order to find files from a specific date (ibid).

Appendix B

Contextual navigation aid for two world wide web systems

A study made by Jonnah Park and Jinwoo Kim (Park & Kim 2000) has shown that contextual information change the users' navigation patterns in World Wide Web (WWW) systems. This increased their subjective convenience of navigation and improving their performance (ibid, p.1).

In their study they clarify that previous studies have shown that users can have problems navigating in WWW systems. The users have trouble identifying where they are or finding the previously visited location (ibid, p.2). The users' uncertainty and lack of awareness in navigation in WWW-systems is seen as disorientation and can be described by four kinds of situations (ibid):

- The user does not know where to go next
- The user does not know how he arrived at the particular place
- The user is not capable of finding the information
- The user does not know how to get to the information

Park and Kim argue that a plausible reason for these problems to occur are the lack of context information. Context information presented to the users can provide the temporal and structural cues of location (ibid).

In their study the contextual information was designed as a structural contextual navigation aid and a temporal contextual navigation aid (ibid, p.5). The aids were designed as two sets of add-on links[1] on top of the basic links (ibid).

The structural context navigation aid provided the users an option to navigate between distant locations (two levels directly upward or downward), and provided structural context information by previewing all the nodes that are two levels upward, two levels downward, and at the same level as the current location. The temporal context navigation aid provided a historical information mechanism (recency-based history which saves the latest URL) (ibid, p.6).

In the study Park and Kim conducted two experiments to investigate the effect of the two types of context information aids in two different WWW-systems: one with extreme well defined structure and one with ill-defined structure. 40 and 46 participants participated in the two experiments respectively. In each experiment the participants was randomly divided into four groups: a control group, a structural context group, a temporal context group, and a group with both structural and temporal. The participants were asked to perform two types of tasks (searching and browsing). In each experiment the participants were asked to fill out a post-questionnaire regarding the ease of use on a 7-point scale (ibid, p.7).

In the experiment with the system, which had extreme well defined structure, the participants were asked to perform four searching tasks and three browsing tasks within 10 min. per task. One example

[1] Add-on links are additionally provided links compared to basic links, which is predetermined by the system's node structure.

of a searching task was that the participants should find a traditional wine. One example of a browsing task was that the participants should order a birthday gift (ibid, p.10).

The results showed that both temporal and structural navigation aid had a significant effect on how many nodes the participants visited in total. There was also a significant effect between the structural and temporal context, and when no structural context was provided there were considerable differences between participants with the temporal context and those without the temporal context. So the temporal context only had influence when the structural context was not provided (ibid, p.13). In general the context aid made the participants visit fewer nodes than the participants with no context aid (ibid). The result of the post-questionnaire showed that the participants found the structural and temporal context to be more convenient for navigation than the participants without structural and temporal context (ibid, p.14). But there was a significant interaction effect between task type and structural context, in that sense that structural context had more influence on the convenience of navigation in the searching tasks as compared to the browsing tasks (ibid).

In the experiment with the system, which had ill-defined structure, the participants where asked to perform two searching tasks and four browsing tasks within the same time interval as in the first experiment. One example of a searching task was that the participants should find information about the differences between two foreign communities. An example of a browsing task was that the participants should select a foreign course for language study (ibid, p.15).

The results also showed that the structural context had a greater effect when participants were performing search tasks with browsing tasks, but there was no interaction effect between the structural context and temporal context. In general the context aid made the participants visit fewer nodes than the participants with no context aid. The result of the post-questionnaire showed that the participants found the structural and temporal context to be more convenient for navigation than the participants without structural and temporal context. Again, the experiment showed that there was a significant interaction between task type and structural context, in that sense that structural context had more influence on the convenience of navigation in the searching tasks compared to the browsing tasks (ibid).

In general the study by Park and Kim made it clear that the participants perceived the WWW-systems with the context information as being more convenient for navigation than those without the context information. Participants with the context information visited fewer nodes repeatedly and fewer pages in total (ibid, p.22). It is therefore possible that the structural context could help participants to reduce navigation problems by previewing information, and the temporal context may contribute to reducing the navigation problems by reviewing information. And because of the diversity in the two different WWW-systems and tasks, it is possible that the context information can provide higher convenient in navigation in other WWW-systems or local-systems (not WWW-stystem) (ibid). It is rather beneficial to investigate the extent of the two structural and temporal context information's effects. The temporal context information helped the participants to the same degree both in terms of searching for specific items and browsing for general ideas. The structural context information helped both the searching and browsing tasks. However it played a more powerful role for the searching task than for the browsing task (ibid, p.23).

Appendix C

One brain - two visual systems

In an article made by Mel Goodale and David Milner (Goodale & Milner 2006) they emphasize that perception of the environment can be seen as two different classes. One can be described as the perception of objects and the other as perception of space.

Until the 1960s and 1970s it was more or less common to think that our visual system best could be described as one class (ibid). But around this time it was suggested that the more ancient subcortical visual system (particular the superior colliculus[1]) enables animals to localize objects whereas the newer cortical visual system allows animals to identify those objects. The formulation "*what vs where*" emerges and was to describe this perspective (ibid).

Goodale and Milner first made some research together in the 1970s to specify the role of the superior colliculus. Their study showed that the eyes of a rodent do not simply send input to the superior colliculus and the visual system but in fact the eyes send to at least ten different areas in the brain (ibid). Each of these areas appeared to have control of their own separate class of behaviour (ibid). Their study showed that the superior colliculus is involved with guiding eye and head movements towards potentially important visual objects and that another subcortical structure plays an important role in guiding animals around their environment (ibid).

In their article it is argued that the ventral system, which passes from the primary visual cortex to the inferior temporal lobe, is concerned with object identification while the dorsal system, which passes from the visual cortex to the posterior parietal lobe, is charged with object localization. Therefore they can be considered as two separate processing "streams" (what and where) (ibid). For investigation of this view Goodale and Milner made a series of neuropsychological studies in the late 1990s with a patient who had severe visual agnosia[2]. The patient did not only have trouble with identifying objects but also with basic discrimination of simple shapes (ibid). In their study they examined what the patient was capable of and not. It turned out that even though the patient had trouble identifying or demonstrating the orientation of a slot, the patient could post a card into the same slot without errors (ibid). The patient could manage to tailer her finger-thumb size in order to pick a rectangular block without being able to verbally or manually describe the width of the block (ibid, p.2):

> *In short, she could guide her movements using visual cues of which she seemed completely unaware.* (ibid)

Goodale and Milner came across studies with monkeys which indicated that a class of neurons were active when the monkeys were reaching towards a target, which fits under the term "where" perfectly. However, the neurons were not concerned at all with where the object was, but was concerned with the size and shape of the objects, the monkeys were grasping (ibid). These neurons was shown to be clumped together at the front end of the dorsal stream. It was concluded that what these neurons had in common with the other neurons in the dorsal stream were their visio-motor properties and not the spatial properties.

[1] The superior colliculi is two small bumps above the mid brain, which receives optic tract nerve fibers, that are involved in processing spatial aspects of visual information and eye movements in the direction of visual attention (Colman 2009, p. 741)

[2] Visual agnosia impairs the ability to recognize or identify visual images or stimuli (Colman 2009, p. 19).

These experiments made Goodale and Milner rephrase the formulation "what versus where" to "what versus *how*" so the division between the ventral and dorsal streams was better explained (ibid, p.1).

Goodale and Milner argue that the ventral stream enables the brain to create the mental "equipment" that allows us to think about the world. It also makes it possible for us to recognize and interpret subsequent visual inputs and plan our actions "off-line" (ibid). In contrast, Goodale and Milner see the dorsal stream acting in real time. It concerns guidance of the programming and execution of actions the second we make them. Goodale and Milner describe their account of the two stream as the distinction between vision for perception (ventral) and vision for action (dorsal) (ibid).

Their work also showed what an important role the ventral stream have in visual memories. Goodale and Milner tested this by presenting a visual object briefly to a subject (the same patient mentioned in the previous sections) after which the object was then taken away. The subject was asked, a few seconds later, to pick up the object as if it was still there (ibid). The experiment showed that the patient completely failed the task. The patient was not capable of tailoring her finger-thumb separation when the patient reached out and tried to pick up the object, that was briefly shown (ibid, p. 3). This outcome was contrary to the normal situation of grasping a visible object, where the patient was capable of tailoring her finger-thumb separation perfectly (ibid). This is explained by the patient's lacking possibility of perceiving the dimensions of the object in her consciousness. Therefor the patient had no working memory of the concerned object.

> *Evidently she had no working memory of the object - not because her memory was not working properly, but because she has not consciously perceived the dimensions of the object in the first place.* (ibid)

As Goodale and Milner foresaw, the patients with optic ataxic[3] had the opposite problem (ibid). These patients were not capable of tailoring their grip when objects were present (real time), but when these same patients performed the delayed pantomime task they were able to tailor their grip perfectly (ibid). The study actually showed that the patients with optic ataxic were capable of tailoring their grip perfectly even if the object was still present in the delayed pantomime task (ibid). This means that if a healthy ventral stream is involved it tends to dominate the patient's actions even if the patient's is presented with a visible object (ibid). Goodale and Milner examined this assumption by switching between different sized object, without the patients knowing, during the delay on some trials. It showed that the patients tailored their grip according to the previewed object - i.e. the object the patient remembered was shown to them in the beginning (ibid).

It is therefore positive to claim that the ventral stream plays an important role when acting on our visual memories. As oppose to the dorsal stream, which seems to have no visual memory, the ventral stream allows us to use vision "off-line", it connects the past with the present (ibid). It is however suggested that even though dorsal stream does not have a visual memory, the visuomotor activities does benefit from experience, and being well-honed by practice the visually guided actions becomes more automatic (ibid).

Goodale and Milner assumptions were additional confirmed by the use of MRI technology, which proved the existence of the ventral and dorsal streams (ibid). Goodale and Milner made MRI scans on the same patient who suffered from visual agnosia and other healthy volunteers to show which visual areas in the patients' brain worked and not. To find the object-recognition area in the brain they contrasted the pattern of brain activity that would occur when a subject would look at line-

[3] Optic axatic describes patients with bilateral damage to the parietal lobe. These patients have difficulties with pointing or grasping objects in front of them visually, but can still report what the objects looks like and where they are located relative themselves (Goodale & Milner 2006, p. 2).

drawings of real object with the brain activity that would occur when they looked at scrambled versions of those same line-drawings (ibid). The healthy volunteers' brain activity were strong in the object-recognition area, but in the brain of the visual agnosia afflicted patient there was no activity regardless if it was the line-drawings or the scrambled versions (ibid). It was concluded from Goodale and Milner's recent and earlier studies that because of the patient's accident which caused her visual agnosia she had lost her shape-recognition system. The patient can perceive lines and edges at early stages of the visual system but cannot put these elements in order to form perceived "wholes" (ibid). The patient's condition had however not an impact on her dorsal stream, as normal brain activity occurred when the patient were asked to reach out and grasp objects in different positions (ibid).

One of the main findings that Goodale and Milner can best be described as:

> *[...] what we are consciously seeing is not what is in direct control of our visually guided actions.* (ibid, p. 4)

The patient with visual agnosia could guide her movements using visual cues but was not actually aware of the visual clues. Further, the patients with optic axatic were not capable of tailoring their grip when objects were present in real time, but only when the objects were taken away and they should pretend to pick up an object, were they capable of tailoring their grip perfectly. These findings suggest that what we see is not necessarily what is in charge of our actions (ibid). One of the recent studies of this separation between perception and action is shown by the "hollow face" illusion (ibid). Volunteers were asked to flick off a small bug-like target which was stuck on the inside of a mask, which for the volunteers impelled them to see as a normal protruding face (ibid). In Figure C.1 a picture of the hollow face is shown.

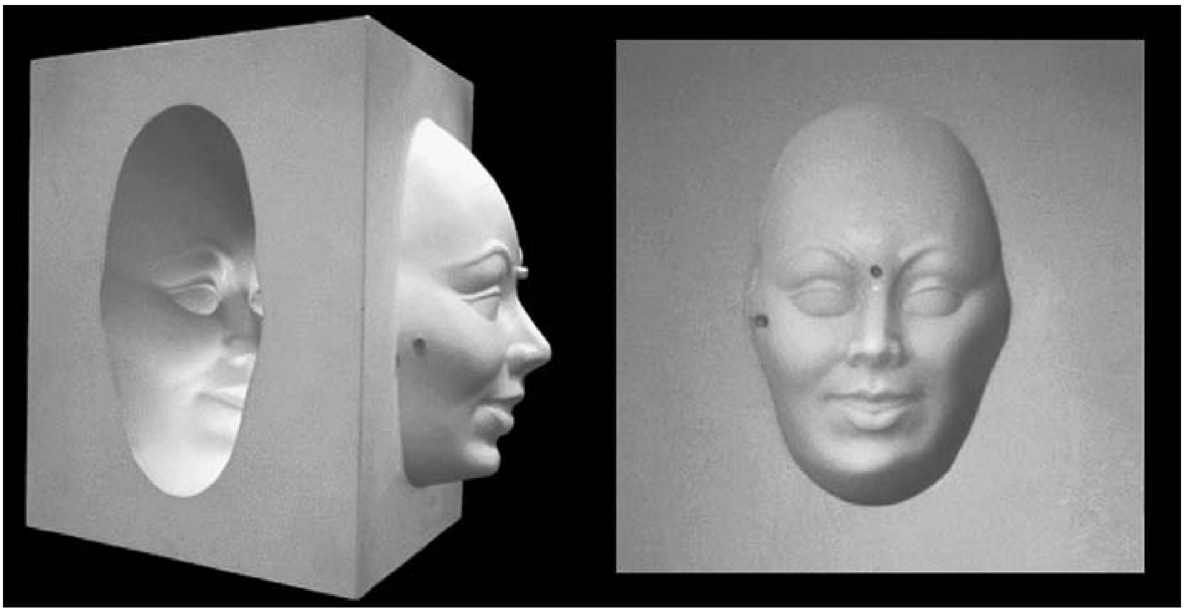


Figure C.1: Figure C.1 shows an example of the "hollow face" (Kroliczak et al. 2006).

The experiment showed that the volunteers were capable of flicking the target as they reached out to the correct point in space (ibid).

Even though Goodale and Milner's and others' work does not show how the two streams are working together, their work clearly present the idea of the dissociation of the visual system - one visual stream for perception and one visual stream for action. This is shown with various experiments with patients who suffered from damage to the one or the other stream and with healthy volunteers.

Appendix D

Activity theory

The activity theory should not be considered a theory in the strict interpretation of the term "theory". It should rather be seen as a descriptive framework (Bannon 1997), that consists of principles. These principles tries to explain the entire activity system beyond the basis of just one being (or actor/individual). This is done by implicating the history of the individual, culture, motivations, and role of the object. The very concept of activity implies that not only does it involve an acting individual, but also something that the activity is directed towards. So the activity is what connects the subject with the object-oriented world. The activity is therefore the relation between subject and object (Leont'ev 2002).

Alexei Nikolaevich Leont'ev was a student of Lev Vygotsky, the founder of cultural-historical psychology (ibid). Leont'ev developed the conceptual framework called Activity Theory (ibid). The activity theory by Leont'ev, has origins from Vygotsky perception of a human being's triple anchoring as a social being, consciousness being, and a member of the animal species (ibid, p. 17). Based on Vygotsky's triple perception of the human individual, Leont'ev tries to give a general psychological theory and therefore an important part was to clarify the mind's biological anchoring (ibid). Leont'ev does this by implicating comparative psychology which compares psychological phenomenons among different life forms including the human species (ibid). Leont'ev attempts to create a connection between the studies of other animal species' behavior and the human psychology. Here Leont'ev focuses on the term: activity. Leont'ev does not only look back at the development of the species before man, but all the way back to when comparative psychology was not useful because no behavior existed (ibid). When Leont'ev's talks about life, it is sufficient that the organism has a metabolism, a substance, and energy decomposition, between the organism and its environment. No activity is required. The plant kingdom is categorized by metabolism, but is stagnant species. Leont'ev sees this as a substance and energy decomposition without activity or without any behavior as a behaviorist would have put it (ibid).

So in Leont'ev's theory the life of an organism can be defined by metabolism, but an organism do not only have a metabolism but may also have an activity. As long as the organism only needs energy and nutrients in an gaseous or liquid form it does not need to search for these living conditions, but can merely wait for these to accrue to it. It does not need higher organized form for activity. In the moment the organism physically seeks objects, a new phase is present. The original metabolism, which concerns the organism secondary decomposition of what it consumes, is still maintained. When searching for food material a new and higher organized type activity is required: an activity that is object- and goal-oriented. This activity is what have named Leont'ev's theory: Activity theory (ibid).

So the activity theory begins with the notion of activity. Activity can be seen as a system of human "doing" whereof a subject interact with an object. In other words a relation between subject and object, whereas the relation can be considered as an interaction between the subject and its environment, which mutual influence each other (ibid, p. 25). In the process of an activity intentionally adjusted mental activities is being executed which brings the organism in contact with objects and conditions which changes the structure of the activity and thereby influence the subject's sense organs in ways which gives information. The structure of the activity is therefore changes in correlation with the properties of the object and with time experience is accumulated in the organism (ibid). In this way activity always happens in proportion to the environment, social, and objects and cannot be seen as isolated from the social relations or as an activity without an object (ibid).

So, Leont'ev sees the human activity as an object-related, but how is the object defined? The object can be treated in two different ways. In the most wide sense an object is a thing, which is in somewhat connection to other things, and therefor as a thing which has an existence (ibid). In a more narrow sense it can also be something that an action is directed against, that is, something the organism relates itself to as object for its activity (ibid). This means as contrary when the activity is directed towards an object as in a tangible thing like a computer, it can also be an intangible thing like a status in a group (an immaterial matter) (ibid), which the activity is directed towards. The notable and somewhat problematic with an intangible thing is that the object is given in and with the activity itself (ibid). An example of this could be if the object for a monkey is status in a group, the tangible "status" is defined by the monkey's (and other monkeys of the group) activity rather than activity is defined by the object "status".

The terms activity, action, and operation can on a subordinate relational level describe the activity. These three terms is closely linked to motive, goal, and condition (ibid). In the animal kingdom the activity is motivated by a certain need, for which an object exists. This object has a meaning to the animal, so it immediate should satisfy the animal's need (ibid, p. 29). This object's meaning is in a way the activity's motive, and therefore the motive and object becomes consistent. In other words, the object(s) in an activity is the activity's real motive (ibid). This means that the objects always correspond to a need and that the activity always links to a motive (ibid). As said before, the animals' activity is initiated and controlled by the object (motive), which can satisfy their needs. However, with humans it is different. Human activity is initiated by the motive, which is based on a need. The actions, which carries out the activity, is controlled by the goals (ibid, p. 31). The goals is defined as "intermediate results", which the actions produce as a "step" towards the object, which is the motive for the activity. This makes the actions a process which is subordinated to the notion of the result, that is supposed to be achieved. In other words, what we do (actions) to achieve an intentional goal (ibid, p. 32). So, activity and action is linked to what must be done to satisfy the needs and the intentional goals. Operations however, is linked to how it is done and is therefore influenced by the attending conditions. It can be seen as the object-oriented activity is directly translated into operations, which is directed towards the conditions (or obstacles), so the activity can achieve it's object (ibid, p. 29). The goal of the actions can remain the same, while the conditions, which relates to the operations can change. The difference can be hard to clarify, since the operations are how the actions are being accomplish in practice and are determined by the circumstances these operations are accomplish in (ibid). In short sense this can be explained by the following: The human activity is constituted by a series of actions, which each independently are controlled by the goals, which are adequate to the motives under the given circumstances. The actions "what" are being accomplish by operations "how", which are available under the given conditions.

The subject's activity, whether or not it is interior (e.g.thoughts) or exterior (e.g. physical activity), is mediated and controlled by the mental reflection of the reality (ibid, p. 93). Things in the object-oriented world, which for the subject, acts like motive, goals, and conditions must in someway be perceived by the subject and imagined, comprehended, located, and reproduced by its memory (ibid). This enables the subject to grasp the object-oriented world. The mental reality, which appears immediately before us, is the consciousness subjective world (ibid). In other words, the consciousness is in its immediacy a picture of the world, which appears before the subject, and includes the subject itself, its actions, and conditions (ibid).

The mental reflection of the reality requires that the subject can collect perceptual information from the object-oriented world (ibid, p. 102). This sensory representation happens through the sensory fabric of the human consciousness (ibid). This fabric forms the sensory structure of the present representation of the reality in the moment it is perceived, or when it appears in the human memory (ibid). This means that the sensory fabric forms the experience of reality by allocating the subject a conscious "image" of the world, which exist outside the subject (ibid). The conscious can externalize itself and this externalization can be seen as the objectified consciousness and holds mankind's social experiences and social practice (ibid, p. 39). It becomes the consciousness' generalized abstractions of social experience and practice, and should be seen as independent from the individual's subjective relationship to the reflected. Leont'ev present the term personal meaning to describe the subjectively side of the consciousness (ibid). This means that unlike values, the personal meaning is what reflects and holds its own reality of life and is therefore in close relation to its motive (ibid, p. 41). Personal meaning is what creates the human consciousness' engagement. It regulates and guides the human activity (ibid).

Appendix E

Conceptualization

E.1 Goal-Derived Categories: The Role of personal and situational goals in category representations

Ratneshwar and co. conducted a study with 127 undergraduate students. The students were asked to rate the similarity of different foods. The participants judged the similarity of eight pairs of food products. In a test the food categories varied between high and low surface resemblance. Furthermore they varied the salience of a shared context like "things you can eat in your car". Some of the food categories fitted this context others did not. They also varied the food categories according to a personal goal of health. The test subjects were presented with two kinds of food. Example given, a plain granola bar and a candy bar. The students then had to rate the similarity of the to food types on a scale from 0-11.

The food types differed each time in which shared attributes (personal goal, context, surface resemblance) they had in common. And also the salience of the context and the health goal differed.

The study showed that people are influenced by the context, the personal goal, and the surface resemblance when judging on the similarity of different food types.

Ratneswar concluded that not only is our mental category representations based on a bottom-up process from the perceptual elements, but also based on a contextual schema top-down process.

E.2 Family Resemblances: Studies in the Internal Structure of Categories

This article documents a study in prototypes. Prototypes are the clearest example of a category. If we look at different items within a given category the thing that binds them together is that:

> "*Each item has at least one and probably several, elements in common with one or more other items*" (Rosch & Mervis 1975, p. 575)

The elements in common is referred to as family resemblance (ibid).

The work includes three different types of categories: Superordinate semantic categories, basic level semantic categories, and artificial categories. The superordinate semantic categories are ordinate categories such as furniture and vehicle they are not physical objects. Basic level semantic categories can be physical objects such as a chair and car. The artificial categories are formed from random digits and letter strings (ibid).

In the first part of their article they try to confirm these hypotheses considering superordinate semantic categories:

- Members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which overlap those of other members of the category

- Items viewed as most prototypical of one category will be those with least family resemblance or membership in other categories

In order to test their hypotheses they conducted two experiments.

In the first experiment six categories were used, each were listed with 20 items, a category could be fruit and the items listed within the category could be banana, able etc. (ibid).

400 students participated in the first experiments where their first task were to rate how prototypical six items were to the categories they were presented in. Each category were presented with one item per test subject. In the second part of the experiment the test subjects had to list as many attributes as possible for the six items, one by one. Afterwards all items were credited with the attributes chosen by the test subjects. Two judges checked the list in order to exclude false attributes. The attributes were then rated on a scale from 1-20 according to how many times the attribute were listed in the category (ibid).

The result showed that the number of attributes an item had in common with other items within the categories corresponded to how high the item were rated as a prototype (ibid).

In the second experiment they wanted to prove the opposite hence items viewed as most prototypical of one category will be those with least family resemblance or membership in other categories. Another 400 students participated (none from experiment 1 participated). This time the experiment only included five categories, again listed with 20 items. The test subjects were told to list three categories for which the items could belong. From the list of these three categories the facilitator could calculate the category dominance for each item. The result showed that there was a significant correlation between prototypicality and dominance of membership in the category for which prototypicality had been measured, and thereby have least membership in other categories (ibid).

Another two experiments were conducted in order to confirm that the same two hypotheses were correct for basic level categories. The procedure of the experiment were almost the same as the previous two experiments. One of the major differences was the use of basic level categories like cars and chairs (ibid).

Again the hypotheses were confirmed which means that for both superordinate categories and basic level categories, the rated prototypicality is corresponding to how many attributes the item have in common with the other items within it's category. Also it confirmed that there was a significant correlation between prototypicality and dominance of membership in the category for which prototypicality had been measured, and thereby have least membership in other categories (ibid).

The advantage of using words of natural languages is that they occur in actual human usage and are therefore closer to the real world than constructed words. Unfortunately this also means that there are many uncontrolled variables which are unanalyzable (ibid, p. 581). Therefore they created another experiment using only artificial categories. The hypotheses of this experiment were that people are faster at learning an item's membership of a category if the item is more prototypical to the category (ibid).

The artificial categories consisted of six items each. The items were strings of a mix of five letters or digits. As letters they only used vowels so none of the words could be pronounced and function as a factor. The artificial categories were not given any names so the only thing saying anything about the category were the letters and digits in common within the category (ibid).

The strings prototypicality within each category were measured by counting in how many of the other words the letters were used within the category. The experiment included a control group were all strings had the same family resemblance score. Besides the control group there were two other categories with six strings in each (ibid).

The test subjects were 30 students. The test subjects were first shown the two categories. Afterwards they were shown a random string and were to guess to which category this string belonged. The test subjects were then told if they had guessed correctly or not. They were then presented with another string and had to guess again and they continued to show the strings to the test subject until they had completed two runs errorless. When they had completed the task they were given the list of the items in each category and were told to judge the prototypicality of each string (ibid).

The number of errors for each category were noted, so was the reaction time[1], and the rankings of prototypicality. When comparing the three variables it showed that strings rated with higher prototypicality gave lesser mistakes and shorter reaction time (ibid).

Based on the experiments they conclude that for both superordinate, basic level, and artificial categories. Items judged as most prototypical have the highest resemblance between items within the category, furthermore other studies has concluded that this also goes for other types of categories like dot patterns (ibid, p. 599). This study have focused on family resemblance, but there are probably other principles or attributes which can connect different items in one category like social factors, memory, or goals (ibid).

E.3 Visual perception

In James J. Gibson's approach to the study of perception it is described how an active observer gathers information from the environment (Gibson 1986, p. 147). Gibson explains perception as an active exploration of the environment. The substantial in Gibson's approach to visual perception is that visual space is defined by the information contained on environmental surfaces. The information for perception is information that remains invariant while an observer moves through the environment (ibid, p. 73).

In a more detailed explanation, our perception, according to Gibson, is based upon three fundamental elements: medium, substances, and surfaces (ibid, p. 16). The medium is what we move through (e.g. water is the medium for fish and air for terrestrial animals), and the medium facilitates different kinds of perceptual systems to evolve. The air for instance makes it possible for different kinds of dispersion of e.g. light particles, chemical particles and vibrations and thereby facilitates the visual, auditive, and olfactory systems (Trettvik 2001, p. 5). The substances is objects we can touch, look at, or eat - it is what we have trouble using as medium (ibid). The surfaces is where the medium and substances meet (ibid).

So when light penetrates the medium and hits a given substance the light is reflected and is perceivable by an observer (ibid). Gibson calls the reflected light for ambient light (ibid, p.4). The ambient light is therefore a result of the light which is spread in the medium and illuminates the surfaces in the environment (ibid, p.5). In order for our visual system to make any sense out of light it must be structured. The environment and its surfaces structures the ambient light. This means that when light is reflected it has, because of the surface's layout, a variance in light density, which is specific for the surface the light is reflected from. Gibson calls this ambient optic array which is considered by Gibson to be sufficient for perception (ibid).

Gibson emphasize in his work that perception of real life cannot be done by a stationary observer

[1] How long it took before they guessed.

but must involve an active observer who is constantly moving his eyes, head, and body in relation to the environment (ibid). Even though the constant movement of the observer results in a constant changing image in the observer's retina there is information that remains constant on the retina (Gibson 1986, 73). The more movement an observer or the object have the more information becomes available because of the changes in the ambient optic array (Trettvik 2001, p. 6). This is because that every point of observation is surrounded by ambient light that is specific for that exact point of observation. When the point moves (when the observer moves) the array of the ambient optic changes, or at least some parts of it do. This is due to the perspective structure (ibid). The parts that do not change Gibson defines as invariant and these invariants is what makes us perceive the world as rather constant even though we move (ibid). This is because the invariant structure is behind the transformations in the perspective structure and specifies the forms of the rigid surfaces of objects e.g. (ibid). Fig E.1 illustrates an example of the constant information that is available of an object (a cup).

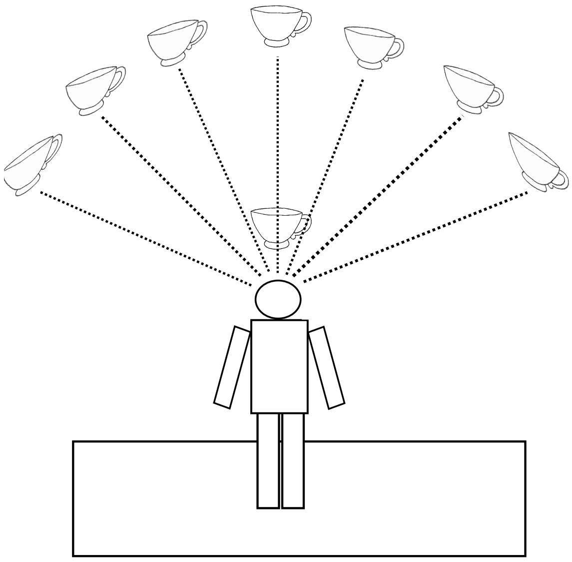


Figure E.1: A schematic illustration of the invariant parts of a cup (Gibson 1986).

But when Gibson talks about perception of objects not only do he talk about visual characters of the objects but also the perception of what objects affords (Gibson 1986, p. 127). These aforementioned invariants is what specifies the subject's options with the object and Gibson called these options for affordance. Affordance is in general what the object e.g. offers the subject (ibid). In other words, it is what the objects are appropriate for or serve or allow (Trettvik 2001, p. 7). An example of this could be how a stable horizontal surface (like a chair or a ledge) offers sitting. So, affordance refers to the meaning objects have for the observers and remain invariant in most cases (Goldstein 2009, p. 3).

The visual perception of affordance can be described as the subjects perception of a possible activity with an object (Trettvik 2001, p. 7). Therefore the subject perceives the relations of activities and these relations of activity are specified by the invariants in the ambient optic array. This is because it is the invariants (as mentioned before) which is behind the transformation in the perceptive structure which e.g. is created by locomotion through the environment (ibid). Perceiving the affordance does not only include what to do with the object, but also includes how one should avoid the object, how to catch it, how to throw it etc.(ibid, p. 8). So, affordance has meaning and value for us, we can perceive affordances which are useful for us but also at the same time be a treat to us.

A cliff affords a high edge to hang-glide from and also afford that we could fall of it by accident and hurt ourselves. Therefore affordance can be seen as information and depends on the observer and the activity (ibid, p. 9).

E.4 Perceptual symbol systems

In his article "Perceptual symbol systems" Barsalou (Barsalou 1999) explains his thoughts about the human conceptual system and perceptual symbols (ibid).

During perception of an object, an event etc. neurons in the sensory-motor system of the brain capture information from the object or event, and also from the body of the observer. These captured information are known as perceptual symbols. Perceptual symbols could be edges, colors, movements, heat,pain, or spatial relations (ibid, p. 583).

The perceptual symbols do not affect a whole state but rather a specific schematic aspect of the state for instance a shape or a color. If our selective memory focus on a specific aspect of the object this specific aspect is most likely what we store in our long term memory (ibid).

When a perceptual symbol is stored in our long term memory it is stored in an associate pattern of neurons, and thereby associated with other perceptual symbols this give the perceptual symbols the opportunity to activate in different patterns according to different contexts. So the perceptual symbol patterns can change over time, creating new patterns of categorizing objects and creating new concepts (ibid, p .584).

As mentioned earlier the perceptual symbols are schematic of nature, which means that they are not holistic. The way perceptual symbols are stored in memory makes it possible to separate the colors, shapes, and orientation, but still it is possible to connect the perceptual symbol components to form a specific object (ibid). Neurons in the perceptual system can code information qualitatively as long as it is unconscious information. Like coding the presence of a line without information about the length, position, or orientation. This further explains how the perceptual symbols can be coded into schemes (ibid).

So far the focus has primarily been on perceptual symbols gained from visual stimulus. However it is important to add that perceptual symbols are multimodal in nature which means that the perceptual symbols can also come from other sources such as audio. People acquire perceptual symbols from speech and sounds, from touch, temperature, and texture. Also, we acquire perceptual symbols from propriception - like movements of objects and body positions (ibid, p. 585).

Another modal that can capture perceptual symbols is introspection (ibid).

Introspection can be divided into three areas (ibid):

- Representational states include the representation of an entity or event in its absence, as well as construing a perceived entity as belonging to a category
- Cognitive operations include rehearsal, elaboration, search, retrieval, comparison, and transformation
- Emotional states include emotions, moods, and affects

As mentioned earlier perceptual symbols are stored in association with other perceptual symbols. These associations or relations makes it possible to simulate an event. If you for instance study a car the information about the body, wheels, windows, and doors are stored in accordance to their spatial location - an object-centered reference frame. More information about the car is added to the reference frame when moving around the car and observing it from different angles or when you look into the trunk etc. Aspects like the sound of the engine, the movement of the car etc. can be included into one of these reference frames or schemes. The information concerning the schema of a car can come from many different experiences with a car (ibid, p. 586).

A simulator contains two levels. The first level is the frame that integrates the different perceptual symbols. The second is the potential sets of simulations that can be constructed from the frame (ibid).

The simulator becomes equal with what others might refer to as a concept. During childhood we develop simulators for important types of entities and events. Once you can simulate an event to a culturally acceptable degree they have an adequate understanding of it (ibid). When a clear concept or simulator has been created it is possible to decide if a given entity is a member of a given category. If the simulator is not able to generate a satisfying simulation the entity is not a part of the category and will therefore not be stored in the same association as the given category (ibid).

When our language develop the perceptual symbols are linked with linguistic symbols and thereby also linked to the simulations. The linguistic symbols develop in the exact same way as perceptual symbols and whenever you hear or read a word of a concept like "car" the simulation of the concept will start (ibid, p. 592).

Once an entity is categorized you can use the knowledge from this category to provide predictions and thereby further knowledge about the entity. This includes ways of interacting with it. This can can be beneficial when working with a new object. Have you once driven a car, you already got a lot of information on how to drive a truck (ibid, p. 586).

Figure E.2 gives an overview of the human conceptual system.

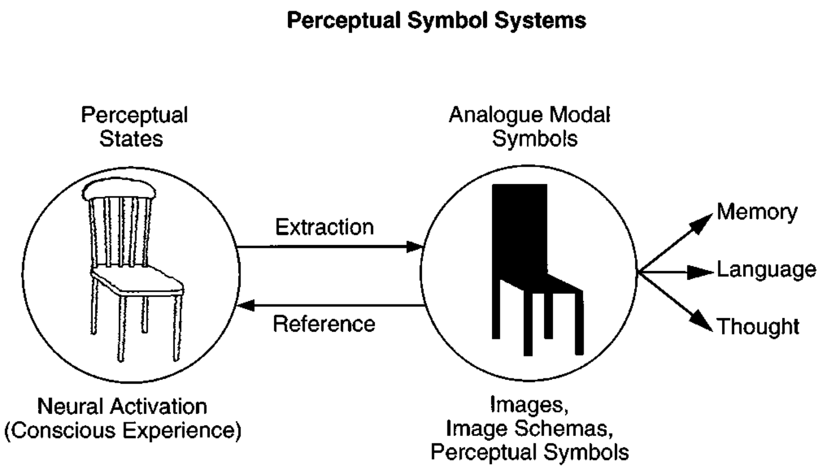


Figure E.2: When the chair is perceived it turns the entities of the chair into perceptual symbols which are stored in a schemata association. The associations can be used in memory, language, and thought through the use of simulators (ibid, p. 2).

E.5 Spatial symbol systems and spatial cognition: A computer science perspective on perception-based symbol processing

This article explains some thoughts about spatial structures in computer systems (Freksa, Barkowsky & A. 1999, p. 615).

> *"Neural representations resulting from perception are often organized in sensoritopic representations, that is, according to spatial structure manifested in the perceived configuration. As entities are perceived in spatial relation to one another."* (ibid, p. 615)

When you perceive entities in a scene you also structure them according to their spatial location hence the perception processes make use of spatial organization. So generating a representation that preserves spatial relations is an easy process. This means that it is easy to recall the spatial structure of a scene. Also the spatial structure works as an inter-modality - as a combiner for different modalities (ibid).

The benefits from the perception of spatial structure can be used in computer systems. The computer is good at implementing arbitrary concepts and processes. This has been achieved by generalizing principles in specific domains to apply to all domains. These principles does not always apply to the domains in the real world which have alienated the computer systems from the real world. In the real world rules and structures are naturally applied to every domain. Structures and principles in computer systems are not given intrinsically but are instead simulated through expensive computational processes. Which means that if the computer should obey every rule of a specific domain the computational cost would simply be to high (ibid).

As already mentioned it is very easy for humans to interpret entities in spatial relation to other entities. One of the things that uses these principles are maps (ibid).

> "*The pictorial map space is organized in spatial analogy to the world it represents. Hence, the map becomes an interface between spatio-perceptual and symbolic concepts that are integrated in a single representational structure*". (ibid)

So maps order the information of the scene into symbols but keep the information about the spatial location, and thereby support the use of spatial information in a natural way.

A computer system could support a map structure and thereby become closer to the real world (ibid).

E.6 On the Origins of the Conceptual System

Mandler (Mandler 2007) worked with conceptual categorization - how we separate humans from animals, cars from airplanes etc. (ibid).

When discovering a new item we must decide whether it can be categorized as the same as a known thing. This can be done by comparing how things look, but this can be misleading. Therefore our categorization is more frequently based on concepts. Since babies have not created concepts yet they can only categorize on perceptual similarities (ibid, p. 743).

Through her study on infants she noticed that infants categorize objects according to perceptual attributes (ibid, p. 745). This is perceptual attributes such as form and color. Later on infants learn to categorize elements on behalf of their conceptual similarities such as what things can be used for, e.g. a saw and an axe can both be used for cutting down wood. Infants can also differentiate between kitchen stuff and bathroom stuff.

Later on perceptual similarities and concept similarities melt together trough our language (ibid).

Mandler has also studied how our conceptual system evolves in the first place, how we start to learn the attributes of a concept on which behalf we categorize it. She have not found a direct answer but studies have proven that infants pay extra attention to motion and in some cases spatial relations. At the age of two months infants start to notice whether objects move by themselves or by the action of other objects (ibid).

When forming our conceptual system we start with creating global concepts[2] . The global concepts then get subdivided according to spatial information. These subdivisions than get subdivided again and so forth until our conceptual system is created (Mandler 2007, p. 747).

[2] Rosch refers to this as supordinate categories like vehicles, animals etc. (Rosch & Mervis 1975).

Appendix F

Eye tracking

An eye tracker is capable of measuring an individual's eye movements. This makes it a useful research tool, because it is possible to know where a person is looking, how long a person is looking, and in what sequence a person is shifting eye location. Utilizing this kind of technology gives insight in the field of human computer interaction to understand visual and display-based information processing (Poole & Ball 2005, p. 1). Additionally it can clarify what factors may have an impact upon the usability of a graphical user interface (ibid). Another use of the eye tracking technique is to capture eye movements and use them as control inputs to a system. This enables people to interact with a graphical user interface by using only the eyes. This is rather beneficial for individuals with a certain incapability.

There are several ways of tracking eye movements, but capturing the reflection of the user's cornea and center of the pupil, using a video-based tracker with an infrared camera and its related image processing software, is a common used method for interactive systems (Duchowski 2007, p. 54). The eye tracker, is often, mounted in a fixed position beneath or next to a screen, and is pointing an infrared light, embedded in the infrared camera, directly into the eyes of the viewer to create a strong reflection of the eye's features (Poole & Ball 2005, p. 2). The infrared light will make the pupils appear bright and also generates the corneal reflection which will appear as a small sharp glint (ibid). In Figure F.1 the bright pupil and the corneal reflection is illustrated.

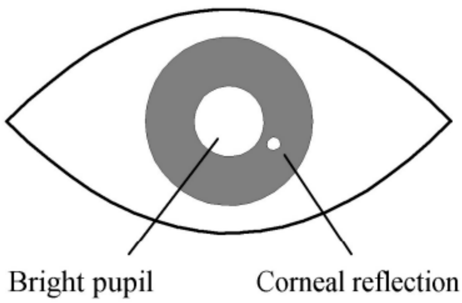


Figure F.1: Illustration of the bright pupil and corneal reflection made by infrared light (ibid, p. 2).

The image software can then identify the center of the pupil and the location of the corneal reflection. The vector between them is then measured, and by using trigonometric calculations it is possible to determine where the user is looking ("point of regard") (ibid). So by using this technique to track eye movements, it is not required to be in physical contact with the user's eyes (Namahn 2001, p. 2).

In order to track the eyes' movements most accurately a calibration process is needed. In this process the eye tracker system is fine-tuned to each of the individuals' physiological properties of their eyes. The calibration procedure is done by displaying a dot on a screen which the user must gaze upon for a certain amount of time, so the tracking system can record the relationship between the center of the pupil and the corneal reflection as corresponding to a specific x,y coordinate on the concerned screen (Poole & Ball 2005, p. 3). The dot will appear several times across the screen to ensure an accurate calibration all over the screen. After the calibration procedure the system should have an accurate direction of the user's gaze. An example of a calibration can be seen in Figure F.2.

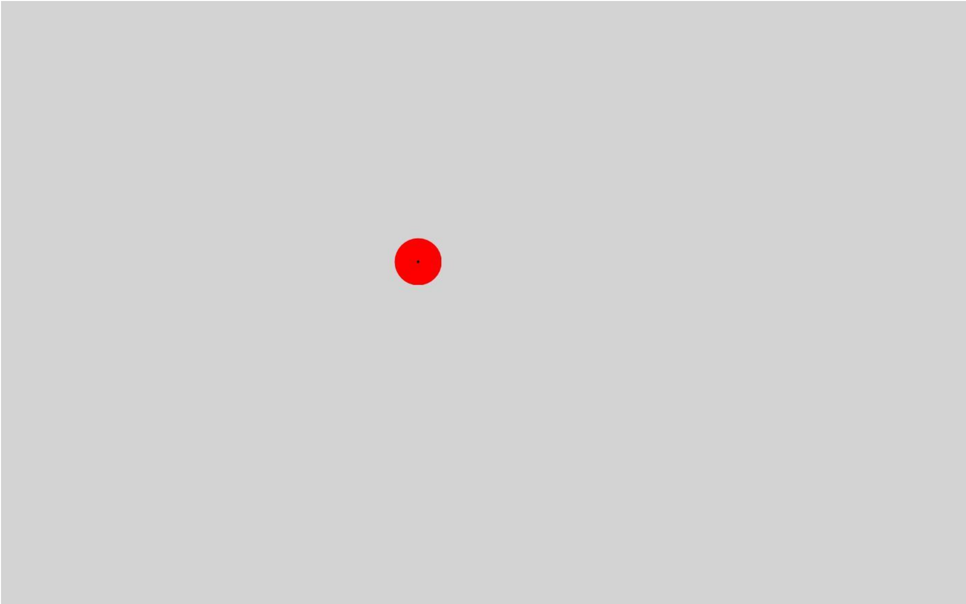


Figure F.2: The calibration of the eye tracker is done by displaying a dot which appears several times across the screen.

Tobii X120 Eye tracker

The eye tracker used in this study was the X120 Eye tracker manufactured by Tobii Technology[1] (Tobii 2012*a*). Figure F.3 shows a picture taken of the Tobii X120 Eye tracker.

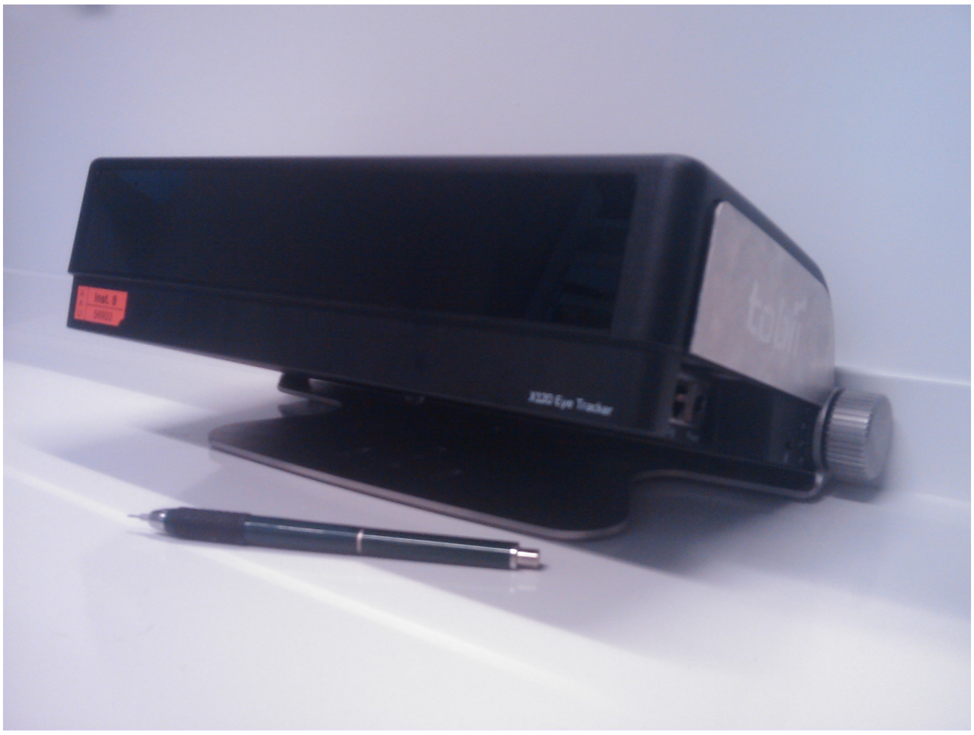


Figure F.3: The X120 Eye tracker from Tobii.

The X120 eye tracker is a stand-alone eye tracker unit which makes it capable of tracking eye movements of any surface (ibid). Therefore the X120 Eye tracker is not prohibited to only measure eye movements on technical surfaces like a screen but also physical surface from objects like on a wall.

[1] For more information about Tobii Technology visit: www.tobii.com.

The eye tracker can have a frame rate of 60 or 120 Hz, which means that the system will either register 60 or 120 data points pr. second or every 16.6 or 8.3 ms (ibid).

Tobii Studio™Eye Tracking Software is a software program that, with the eye tracker's functionality, offers to record, observe, analyze, and visualize data in heat maps and gaze plots[2] (ibid). In order to visualize the data of eye movements the data is processed into graphical fixations (ibid). These graphical fixations are represented by dots, where the larger dots indicate a longer time of fixation. The saccades are indicated by lines between the fixations. It is possible to set a time limit of the minimum length of a fixation. In this study it was set to 60 ms which is set as default value by Tobii Studio (ibid). Thus, fixations that have a duration under 60 ms, is not graphically represented. Even though several studies have shown that 100 ms is appropriate for most eye tracker studies (Tobii 2012*b*), we estimated that the value 60 ms was a proper time duration to gather information about the test subjects' fixations. This is because it is possible for viewers to gather information from fixations under 100 ms. In a study by Keith Rayner (Rayner et al. 2009) it is demonstrated that viewers can acquire information to understand the gist of a scene in just 40-100 ms (ibid). Therefore by setting the value to 60 ms, we secured that more information about the test subjects' fixations was collected by the eye tracker. It has however been considered that there might be a risk of information getting lost by neglecting fixation with length under 60 ms, because the test subjects could have fixated upon places as the mini-map e.g. in a duration of 40 ms.

In Tobii Studio it possible to view the eye tracking session with a screen capture of a test subject (ibid). This makes it possible to review the whole session with the possibility of deeper analysis and at the same time be able to see what the test subject looked upon in the session in terms of fixation points (ibid).

Another useful feature in Tobii Studio is the gaze plot. The gaze plot feature displays the eye movement sequence, order, and duration of a gaze fixation. The gaze plot makes it possible to see a specific part of a session or the entire session with the fixation dots. The dots are labeled with numbers, so the exact order of fixation points is possible to see. An example of a gaze plot can be seen in Figure F.4.

Figure F.4: The gaze plot shows where and in what sequence the user has look upon.

[2]Features in Tobii Studio may vary according to what version is used. In this study Tobii Studio 3.0 was used.

Appendix G

Verbal instructions

Mundtlige instruktioner

Hej velkommen.

Som det første vil jeg bede dig rate din musik kendskab.

Dette forsøg omhandler at huske placeringen af nogle musik albums i et system.

Det første vi skal gøre er at kalibrerer eyetrackeren. Eyetrackeren er den her maskine, som vil følge med i, hvor du kigger hen på skærmen. Når vi har kalibreret eyetrackeren vil du få 3 min til at lære placeringen af nogle musik albums at kende. Herefter vil du få nogle opgaver, som går ud på, at du skal prøve at finde nogle albums. Efterfølgende har vi nogle spørgsmål, vi gerne vil stille dig. Har du på nuværende tidspunkt nogle spørgsmål? Ellers kan du bare spørge, når vi gennemgår tingene. Vi vil gøre dig opmærksom på, at vi optager lyden fra forsøg og tager nogle billeder.

Det er vigtigt at du under forsøget ikke bevæger dig alt for meget frem og tilbage, da det kan forvirre eyetrackeren du må selvfølgelig gerne dreje hovedet (viser med kropsprog, hvad der menes).

Kalibrering: Du skal nu følge kuglen på skærmen for at kalibrere eyetrackeren.

Systemet (Der gives en kort introduktion til systemet hvor alt funktionalitet bliver gennemgået)

Du får nu 3 min til at lære alle albums at kende efterfølgende vil du få et billede ligesom dette fremvist (billede af album med label vises) og du skal så prøve at finde det i systemet. Er der nogle spørgsmål inden du går i gang?

Godt du har tre min fra nu af.

Appendix H

Enclosed DVD

The enclosed DVD contains the following:

- The two interactive prototype systems (with instructions of how to run the executable files)
- Quantitative data in form of the test subjects' rating and performance
- Qualitative data in form of the transcribed interviews
- Screen captures from the eye tracker
- Sound recordings of the test subjects
- Declaration of statements
- Google Sketchup install file
- ForeUI install file
- A copy of this paper

List of Figures

Bibliography

Agresti, A. & Finlay, B. (1997), *Statistical methods for the social sciences*, Prentice Hall.

Bærentsen, K. B. (2001), 'Intuitive user interfaces', *Scand. J. Inf. Syst.* .

Bannon, L. (1997), 'Activity theory'.
http://www.irit.fr/ACTIVITES/GRIC/cotcos/pjs/TheoreticalApproaches/Actvity/ActivitypaperBannon.htm.

Barreau, D. & Nardi, B. A. (1995), 'Finding and reminding: file organization from the desktop', *SIGCHI Bull.* .

Barsalou, L. W. (1999), 'Perceptual symbol systems.', *the Behavioral and Brain Sciences* p. 577.

Colman, A. (2009), *A Dictionary of Psychology*, Oxford Paperback Reference, Oxford University Press.

Dalgaard, P. (2008), *Introductory Statistics with R*, Statistics and Computing, Springer.

Duchowski, A. T. (2007), *Eye Tracking Methodology: Theory and Practice*, Springer-Verlag New York, Inc., Secaucus, NJ, USA.

EaSynth, S. (2012), 'Foreui'. http://www.foreui.com/.

Eysenck, M. W. & Keane, M. T. (2005), *Cognitive psychology*, Psychology Press.

Fertig, S., Freeman, E. & Gelernter, D. (1996), 'Finding and reminding reconsidered'.

Field, A. & Hole, G. (2003), *How to Design and Report Experiments*, Sage publications Limited.

Freksa, C., Barkowsky, T. & A., K. (1999), 'Spatial symbol systems and spatial cognition: A computer science perspective on perception-based symbol processing', *the Behavioral and Brain Sciences* .

Gibson, J. (1986), *The Ecological Approach to Visual Perception*, Resources for ecological psychology, Lawrence Erlbaum Associates.

Goldstein, E. B. (2009), 'The ecology of perception gibson's'.

Goodale, M. & Milner, D. (2006), 'One brain - two visual systems'.

Google (2012), 'Sketchup'. http://sketchup.google.com/.

Kjørup, S. (1997), *Forskning og samfund : en grundbog i videnskabsteori*, Gyldendal.

Kroliczak, G., Heard, P., Goodale, M. A. & Gregory, R. L. (2006), 'Dissociation of perception and action unmasked by the hollow-face illusion'.

Kvale, S. (1997), *InterView: En introduktion til det kvalitative forskningsinterview*, Hans Reitzels Forlag.

Lansdale, M. (1988), 'The psychology of personal information management', *Applied Ergonomics* .

Leedy, P. & Ormrod, J. (2010), *Practical Research: Planning and Design: International Edition*, Pearson Education, Limited.

Leont'ev, A. (2002), *Virksomhed, Bevidstehed, Personlig*, Hans Reitzel.

Malone, T. W. (1983), 'How do people organize their desks? implications for the design of office information systems', *ACM Transactions on Office Information Systems* .

Mandler, J. (2007), 'On the origins of the conceptual system'.

Namahn (2001), 'Using eye tracking for usability testing'. a research note by Namahn.

Nørager, R. & Trettvik, J. (2007), 'There is no "where" only "what" in user interfaces: Lack of spatiotemporal dynamics undermines user interaction'.

Norman, D. (1999), *The invisible computer: why good products can fail, the personal computer is so complex, and information appliances are the solution*, MIT Press.

Norman, D. (2002), *The Design of Everyday Things*, Basic Books.

Park, J. & Kim, J. (2000), 'Contextual navigation aids for two world wide web systems', *International Journal of Human-Computer Interaction* pp. 193–217.

Poole, A. & Ball, L. J. (2005), Eye tracking in human-computer interaction and usability research: Current status and future, *in* 'Prospects, Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc'.

Ratneshwar, S., W., B. L., Pechmann, C. & Moore, M. (2001), 'Goal-derived categories: The role of personal and situational goals in category representations'.

Rayner, K., Smith, T. J., Malcolm, G. L. & Henderson, J. M. (2009), 'Eye movements and visual encoding during scene perception'.

Rekimoto, J. (1999), 'Time-machine computing: A time-centric approach for the information environment'.

Riis, O. (2005), *Samfundsvidenskab i praksis: introduktion til anvendt metode*, Hans Reitzel.

Rosch, E. & Mervis, C. B. (1975), 'Family resemblances: Studies in the internal structure of categories', *Cognitive Psychology* pp. 573–605.

Rugg, G. & McGeorge, P. (2005), 'The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts'.

Tobii, T. (2012*a*), 'Flexible eye tracking - tobii x60 & x120'. http://www.tobii.com/en/eye-tracking-research/global/products/hardware/tobii-x60x120-eye-tracker/.

Tobii, T. (2012*b*), 'Tobii studio analysis software release'. http://www.tobii.com/Global.

Trettvik, J. (2001), 'En økologisk tilgang til perception og aktivitet'.

Velichkovsky, B. M. (1990), 'The vertical dimension of mental functioning. special issue: Domains of mental functioning: Attempts at a synthesis'.