Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Social Sciences 142

On the Relative Utility of 3D Interfaces

BY

MONICA TAVANTI

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2004

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2, Ekonomikum, Uppsala, Friday, December 3, 2004 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Tavanti, M. 2004. On the Relative Utility of 3D Interfaces. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 142. 74 pp. Uppsala. ISBN 91-554-6102-6

Three-dimensional (3D) interfaces are made with the ambition to reinvent traditional two-dimensional (2D) displays into intuitive 3D environments that allow natural interactions and guarantee users' satisfaction. However, reality is far from the expectations and 3D interfaces remain experimental. The results of empirical studies comparing 2D and 3D interfaces are incoherent and do not indisputably endorse the development of 3D applications. The reason behind this incoherency is that empirical comparisons often discard several factors that go beyond the specific style of the interface itself although they play a major role in human performance. Specifically: perceptual factors, factors related to semantics, contextual factors.

The identification of these factors was carried out through a set of empirical studies tackling two applications domains: information management (namely, the retrieval of data and on spatial memory tasks) and Air Traffic Control. Concerning the first domain, the results suggest that 3D interfaces can support spatial memory if the 3D interface is provided with a spatial structure that has also a semantic function. Also, the specific content of the objects disseminated in the 3D structure seem to affect performance. It is argued that when there is a strategic coupling of the semantics of the spatial structure with the meanings of the objects, 3D interfaces could enhance spatial memory. Concerning the second domain, the results indicate that 3D interfaces could support controllers' tasks only for a limited set of activities. These do not include standard monitoring tasks where the presence of 3D would actually cause detriment to the performance due to perceptual biases. Finally, it is claimed that the idea of familiarity with an interface style can affect the way people interact with it and, despite the fuzziness of the concept, familiarity may represent a real challenge in users' acceptance of 3D interfaces.

Keywords: human-computer interaction, information visualization, 2D and 3D interfaces

Monica Tavanti, Department of Information Science, Box 513, Uppsala University, SE-75120 Uppsala, Sweden

ISSN 0282-7492
ISBN 91-554-6102-6
urn:nbn:se:uu:diva-4681 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4681)

A Nerino Neri

List of Papers

1. Tavanti, M., Lind M. (2001). 2D vs. 3D, Implications on Spatial Memory. Proceedings of the IEEE Symposium on Information Visualization 2001, San Diego, California, October, 2001.
2. Tavanti, M. (2004). Deconstructing Space: Spatial Memory for symbols and pictures in 2D, 3D and in Deconstructed Displays. Research Report 2004:2, Department of Information Science, Uppsala Universitet.
3. Dang-Nguyen, T., Tavanti, M., Le-Hong, H., Duong, V. (2004). 3D stereo Displays for Air Traffic Control: a multidisciplinary framework and some results. Research Report 2004:3, Department of Information Science, Uppsala Universitet. An earlier version of this paper appeared as: Dang Nguyen T., Le-Hong H., Tavanti M., (2003). Analysis of the Applicability of 3D Stereoscopic in Air Traffic Control. Proceedings of the IEEE 6th International Conference On Intelligent Transportation Systems Shanghai, China, October 12-15, 2003.
4. Tavanti M., Dang Nguyen T., Le-Hong H. (2004). Usability Inspection of a 3D Interaction Metaphor. Proceedings of the 2nd RIVF '04 Research Informatics Vietnam-Francophony, Hanoi, Vietnam, February, 2004.
5. Tavanti, M., Le-Hong H., Dang T. 3D for ATC displays: ask controllers, they know better. Submitted to: AIAA Journal of Aerospace Computing, Information, and Communication (JACIC). An earlier version of this paper appeared as: Tavanti, M., Le-Hong H., Dang T. (2003). Three-dimensional Stereoscopic Visualization for Air Traffic Control Interfaces: a Preliminary Study. Proceedings of AIAA/IEEE 22nd Digital Avionics Systems Conference, Indianapolis, Indiana, October, 2003.
6. Tavanti, M., (2002). Using 3D to Visualise Medical Data. Proceedings of ACM Advanced Visual Interfaces AVI 02, Trento, Italy, May 2002.
7. Tavanti, M. (2004). Representing medical information: a subjective evaluation. Submitted to Behaviour & Information Technology.

Permission to reprint the papers in this thesis has been obtained from the publishers

Contents

1. Introduction

De Kerckhove (1993) uses the term "psycho-technology" to define a technological tool that:

> Emulates, extends, or amplifies perceptual, cognitive and interaction functions […]. The telephone, the radio, the television, the computers […] contribute to the creation of intermediate spaces for the information elaboration[1] (p.22).

According to this definition, the radio and the telephone can be envisioned as tools that extend human speaking abilities while the television is a tool that expands people's collective imagery; but what about computer interfaces? In this view, such interfaces are cognitive artefacts (Norman, 1993) since they are tools that amplify cognitive abilities.

Interfaces can be characterized by usability. In the definition of the International Standards Organization (ISO/IEC 9241-11, 1998) usability is defined as: "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". Usability is here defined across three key-elements: effectiveness, efficiency and satisfaction. According to Dix, Finlay, Abowd & Beale, (1998) the effectiveness is the accuracy and completeness with which specified users can achieve specified goals in particular environments; the efficiency refers to the resources expended in relation to the accuracy and completeness of goals achieved; the satisfaction is the comfort and acceptability of the work system to its users and other people affected by its use.

As they are defined, the concepts of effectiveness and efficiency refer to the objective dimensions of the notion of usability, leaving out more subjective notions as the users' personal satisfaction and well being in relation to a system. These objective dimensions can be thought of as defining utility, that is how well the functionality of a system supports the execution of some task.

---

[1] Translated by the present author

2. Focus of the thesis: utility of three-dimensional interfaces

The thesis focuses on the concept of *utility* of particular interfaces, namely three-dimensional (3D) interfaces, especially in comparison to more common two-dimensional (2D) interfaces. Two application domains are considered by this research: information management (with its main focus on the retrieval of data and on spatial memory tasks) and Air Traffic Control (ATC). For the two domains tackled by the thesis, an in-depth analysis of the empirical studies at the state of the art was carried out. The results of the analysis indicated that, besides the type of interface, (2D or 3D) there are several factors that, if not properly considered, could affect users' performance with the 3D interfaces in a number of ways. The original set of empirical and survey studies encompassed by this thesis provide support for this hypothesis. The utility of 3D against 2D interfaces cannot be decided *a priori,* just relying upon the common-sense assumption that 3D is more similar to real world settings than 2D and therefore easier to interact with. The utility of 3D interfaces is relative and dependent on *perceptual factors* (which are related to human perceptual abilities (and/or limits), *factors related to semantics* (for instance the semantic organization of the space, of the content of objects distributed within it and of the relation among the objects and the space) and on *contextual factors* (which are related to pre-existing knowledge of the interface and therefore to the time used to get acquainted with some interface's style and its functionalities). This idea does not herald a non-scientific approach to 3D interfaces design. On the contrary it is argued that these factors are essential elements to be taken into account in the design of effective 3D interfaces and that these elements can be scientifically investigated.

3. Thesis structure

As stated above, the idea of a *relative* utility of 3D interfaces will be investigated through empirical data. The argument is quite vast and the thesis is based on specific contexts, thus a first part of the empirical data is derived from the literature, while a second one is provided by the original contributions of the thesis.

The first part of the thesis comprises the sections from 4 to 7.3.

Section 4 and 4.1 provide a description of the dimensions across which the notion of '3D' is defined, including the types of 3D interfaces that the thesis is dealing with, namely *2½D* and *3D stereoscopic* interfaces. Section 4.2 provides a description of the systematic biases that 3D perception entails.

Section 5 provides some theories of external representation and section 5.1 introduces to the notion of Information Visualization. Section 5.2 presents some theoretical approaches supporting the view that the utility of graphical interfaces cannot be claimed *per se*.

Sections from 6 to 6.2 present some concepts in interface design relevant for this thesis, including the notions of spatial metaphors and spatial cognition.

Sections from 7 to 7.3 present a summary of the empirical results found in literature studying 2D compared to 3D interfaces. The studies provide inconsistent and sometimes contrasting results. An analysis of the possible reasons behind these inconsistencies is also provided. This analysis is articulated into three sections.

Section 7.1 presents some results dealing with information retrieval, with attention to studies comparing 3D and 2D interfaces across spatial memory tasks.

Section 7.2 deals with a special case of information management and it refers to medical data. In this case study, logic connectives are used to describe the development of diseases. In the section, some details about what is known of the way people perform logical reasoning is given, as well as a description of graphical representations that may allow people to effectively tackle such problems.

Section 7.3 present the summary of the main empirical results comparing 3D and 2D interfaces for ATC applications.

In one empirical study presented in the thesis is used a special device for interacting with 3D volumes. In order to provide a better understanding of this new interaction technology, section 8 provides a brief introduction to the notion of *3D interaction technique*.

Section 9 gives to an introduction to the methodologies used during the empirical studies conducted for the thesis work.

The second part of the thesis comprises the sections from 10 to 13.

Section 10 gives a summary of the original contributions, organized according to the domains of information management and ATC.

Section 11 provides the description of the papers upon which the thesis is based, including some survey and empirical results.

Section 12 and 13 present the conclusions and a discussion of future work.

4. Three-dimensional Interfaces

Shneiderman (1996) suggests a taxonomy of graphical interfaces based on the type of data they support. There are seven data types in this taxonomy: temporal data, multidimensional data, tree data, network data, 1-dimensional, 2-dimensional, and 3-dimensional data. Computer generated 3D views can be implemented in several ways. A common one, that is often referred to as *2½D*, make use of monocular and pictorial cues (see below) to render a 3D effect and can be displayed on any graphic interface. A more elaborate one, often called 3D stereoscopic, adds the use of binocular cues to depth (specifically, the stereopsis) and usually requires special equipment. The next section briefly describes the differences between these two types of 3D interfaces because both of them were used during the empirical evaluations of this research.

4.1 3D displays: monocular and binocular cues to depth

As previously stated, *2½D* make use of monocular and pictorial cues to render an effect of depth and three-dimensionality. These cues are called *monocular* since a single eye is sufficient to make distance and depth judgments based on them and *pictorial* because they are widely used in paintings and pictures to provide the viewer with a sense of depth and three-dimensionality on flat canvases. Even if those interfaces can be displayed on a two-dimensional surface, the content of the graphical representation *looks like* 3D. Important pictorial cues are:

- *Interposition (or occlusion*) that is, the overlap of objects within a space; the relative size of objects in relation to the others (bigger objects are perceived as closer than smaller ones).
- *Linear perspective*, according to which parallel lines appear as converging in the distance; shading, which is the body of shadows and lights within a scene.
- *Height cues,* that is objects set closer to the line of the horizon seem farther away from the observer.
- *Atmospheric perspective*, for which blur and undefined objects within a scene are set far away from the observer.
- *Texture* gradients according to which the texture of surfaces become thinner and denser as the distance from the observer increases. It is rather easy to deliver the depth effect with the support of such cues on standard computer screens, using more or less sophisticated graphics 3D modelers.

There are also binocular cues to distance perception, specifically convergence and stereopsis. The convergence consists of the movement of the eyes to fixate objects; the two lines of sight converge at a certain point and depending on the distance of the fixated object, the angle formed at their intersection will be narrower (far away object) or wider (closer objects). Stereopsis is based on the slight difference of the images projected onto the retina of each eye. Such differences between the two images are called *binocular disparity* and provide information about distal 3D structure.

The state of the art offers several techniques to produce stereopsis to display 3D scenes and objects, but their review goes beyond the scope of this thesis. Only the technique used in this thesis is briefly described herein. It is based on special eyeglasses. The lenses of the glasses contain shutters, letting the light pass, or blocking it, alternately to the right and left eye. The images emanate from a CRT screen and an infrared signal tells whether the left or the right image is written on the screen. The trick is that neither the left nor the right eye can discover this flickering since the refresh rate is very high (usually a 120 Hz refresh rate is used). Therefore the two images are perceived as simultaneous and not as sequential.

Below the term *3D* is used to indicate monocular -*2½D*- interfaces, and *3D stereoscopic* (or simply, to *stereo*) for such interfaces.

4.2 Distortions in depth estimations

Recent research in space perception (Lind, Bingham & Forsell, 2003) shows that the perception of space is not metrically accurate and that it is not based on precise Euclidean estimations of distances. More precisely, in 3D spaces distances along the line of sight are perceived with a large amount of error. Reliable judgments of the positions of objects falling along this line are still possible by a perceiver given that they are qualitative judgments, they are basically referring to the ordinal relationships of the objects, not the their actual distances.

The perception of 3D distances in the depth dimension over longer distances is, in addition, distorted and compressed (that is, perceived as shorter than it actually is). This discovery will be used to explain some of the empirical results reported in section 7 and, in more general terms, it has a heavy impact on the research concerning the design and the utility of 3D interfaces. As a matter of fact, this also suggests that 3D can adequately convey to the users only the qualitative characteristics of 3D space, but that these can be easily grasped and understood. In summary, any use of 3D that entails fine-grained judgments based on metric structures has to be rejected because the users will only perceive uncertain and distorted estimations.

5. External representations: images and symbols

According to some approaches, the effectiveness of tools (not necessarily graphic interfaces) depends on a number of constraints. Tools can be defined as external artifacts (Norman, 1993) when they allow a direct and tangible access to their whole contents. The impact of the artifact on its users is not neutral. Zhang (one of Norman's students) demonstrated this idea, modifying the size and the shape of some objects (cups of coffee) to be used in a puzzle game. He showed that the physical structures constrained people's behavior during a problem-solving task. This finding is close to Gibson's concept of affordance, which is "the fit between an animal capabilities and the environmental supports and opportunities" (Gibson, 2000, p.15). There are some intrinsic properties of objects that determine their use and this functional relationship cannot be determined *a priori* since some objects afford some opportunities of action while others do not.

In this context, we want to talk about *cognitive affordances* in reference to external representations (defined below). It is necessary to make a distinction between this idea of *cognitive affordance* and the original meaning defined by Gibson. For instance, Gibson denied the existence of internal representations; he asserted that the environment provides the affordance that are external to the observer and that the observer does not have to elaborate information, but simply capture it as it is. This original concept of affordance was then extended and re-shaped by Norman (1990) into the concept of *cognitive* (or *perceived*) *affordance.* The idea of *cognitive affordance* does not deny the existence of internal representations and does imply a mental elaboration of the objects that are perceived and used by the observer. *Cognitive affordance* refers to the extent to which designed external representations 'invite' users to intended (by the designer) interpretations and actions.

Another difference between the two notions of affordance is that, according to Gibson, affordance mainly refers to physical objects and to physical environments while Norman's *cognitive affordance* refers to the general concept of *external representations*.

Gleitman (1995, p.301) defines external representations as representations that "stand for something else, such as maps, blue-prints, menus, price lists" and therefore also graphical interfaces. Representations can be divided into two main groups, symbolic and analogical: "Analogical representations capture some of the actual characteristics of (and they are analogous to) which they represent. In contrast symbolic representations bear no such relationship to the item for which they stand" (Gleitman 1995, p.302). A simple example (Schwamb, 1990) is given by two objects, one of which is longer than the other; the symbolic representation will be **{longer (a, b)}**; while the analogical will be **{________ and ___}**. Whilst in the first representation the meaning is derived by the presence of the term 'longer', in the second, the relationship between the two items is implicitly given through the physical isomorphism existing between what is represented and its representation. Analogical representations, also when produced by computer graphics, may have the potential of better supporting people in performing some tasks compared to symbolic ones. Examples in this sense are provided by Zhang (1997), which shows that people playing with a graphical version of a tic-tac-toe game, played better when compared to others using a numeric version of the same game; others (Stenning & Oberlander, 1997) suggest that graphical representations limit the arbitrariness of abstraction and therefore support processibility.

Bruner (1966) re-elaborate Piaget's theory of stages of development identifying three different types of mentality: enactive (manipulation), iconic (recognition, comparison) and symbolic (abstraction). Based on his experimental results Kay (1990) claims that: "Doing with Images makes Symbols", that is, working with images may support abstract reasoning. However, graphical representations are not to be considered *a priori* better than the symbolic ones.

What might make graphical representations better is the extent to which they enhance information processing (Stenning & Oberlander, 1997; Larkin & Simon, 1987) their ability to *afford* tasks like reasoning and understanding. One of the main goals of 'Information Visualization' as a scientific discipline is defining ways to build graphical interfaces able to support users tasks and enhance their performance.

5.1 Information Visualization defined

In a recent publication, Colin Ware (1999) defines the expression 'Information Visualization' as follows:

> Until recently the term visualization meant constructing a visual image in the mind (Shorter Oxford English Dictionary). But now it has come to mean something like a graphical representation of data or concepts. Thus, from being an internal construct of the mind, visualization has become an external artifact supporting decision-making. (p.1)

Gershon, Eick & Card (1998) state:

> Visualization is the process of transforming data, information, and knowledge into visual form making use of humans' natural visual capabilities.

These definitions are compatible with the ideas exposed in the previous section claiming that visual interfaces can act as a communication layer between data and the user. However, these definitions remain still partial, as they do point towards to the function of information visualization but do not allude to its precise content.

So far, analogical and graphical representations have been discussed. But information visualization does not necessarily exclude the representation of symbolic data like words or numbers; graphical information may either substitute or integrate symbolic information in a functional manner. The key issue is defining the elements that allow the design of effective graphical interfaces. Such elements are presented in the next section.

5.2 Three factors in information visualization: culture, perception and tasks

There is an ongoing debate regarding how to design effective graphical visualizations. Theory suggests that at least three issues have to be considered in the design of graphical interfaces: culture, perception, and tasks.

For some semiotics approaches, there is an arbitrary relation between a symbol and its meaning (de Saussure, 1959). The understanding of the meaning is determined by a cultural link, accepted as shared a convention. If we extend this approach, we can claim that also graphical representations are constrained by cultural rules that preside over their construction and that visualization is also dependent on the knowledge of those rules. However, it is reasonable to assume that culture is only one of the factors that have to be considered. For instance, Eco (1992) states that: "iconic signs do not have the same physical properties as the object, but they set a perceptive structure which is similar to the one triggered by the object".[2] This last statement introduces the notion of perception and of perceptual factors in information visualization design.

Ware (1999, p.16) writes that the "sensory aspects" of visualizations derive their expressive power from being well designed to stimulate the visual sensory system. Take Venn Diagrams as an example: the effectiveness with which they are understood, can be found in their compliance with the Gestalt laws (Kanizsa, 1980) about perceptual organization (Ware, 1999, p.210). Even more striking examples of the importance of sensory aspects can be found in the phenomenon called *simultaneous color contrast* (Gleitman,

---

[2] Translated by the present author

1995; De Valois & De Valois, 1975) where the appearance of one color is altered by the simultaneous presence of another color. Another example is the *successive color contrast phenomenon*, where the appearance of a color changes if another color is presented before. Such types of perceptual phenomena are general and they take place independently of the users' culture; therefore they represent a strong constraint in the design of graphical interfaces.

There is also a third element to be considered in information visualization, and that is the task. Tasks are at the core of visualization since they are what visualization should support. An example of this is given by Norman (1993) who criticizes the use of vertical bars to display qualitative data. This design choice is likely to induce errors, as bars better support tasks like the extraction of quantitative dimensions. Some (Card, Mackinlay & Robertson, 1991) claim that the interface should provide the exact amount of information necessary to carry out the task, nothing more and nothing less. This last advice seems reasonable but is very difficult to implement. For instance, it is not always possible to define 'the exact amount' of information required to perform a task. However, it is of course advisable to have as much knowledge of the task as possible. Tasks remain at the center of interface design

To summarize, the theories sketched above on external representations and information visualization claim that graphical representations can efficiently integrate or surrogate symbolic data to support users' activities. However, the utility of graphical interfaces, being either 3D or 2D, cannot be argued in absolute terms, but is relative and dependent on a variety of factors like cultural rules, perceptual issues, the tasks the visualization are intended to support, and the semantics of the displayed content. The aim of the present thesis is to apply this reasoning to the issue of 2D versus 3D interfaces and to provide some pieces of empirical evidence to further enhance our understanding of these issues.

6. Interface design

The interface is a “discrete and tangible thing” (Laurel, 1990) that visually represents what is going on in the hidden and intangible parts of computers. With the introduction of menu-oriented interfaces, and later graphical ones, computing technologies became more widely spread and accessible to a greater number of users (not only to computing scientists or programmers). This also made Human Computer Interaction (HCI) a discipline of a more general concern. According to definition provided by Hewitt, Baecker, Card, Gasen, Mantei, Perlman, Strong & Verplank (2004):

> Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.

As a matter of fact, the term interaction itself implies a mutual and reciprocal sense in the action, so that not only the human acts but also the systems and the interfaces do act towards the users. Again, whether they do behave in a usable manner is a matter of how the interface, and of course the underlying system, is designed.

A common strategy used to design effective interfaces, is metaphors. The metaphor expresses the similarity in topological or metrical structure between different *dimensions* (Gärdenfors, 1996), so one word or token that represents a certain quality in one dimension can be used to express the same quality in another dimension. Metaphors, which are widely used in everyday language, are adopted in interface design. The Star system developed at the Xerox Parc was the first commercial computer that used a graphical user interface and that exploited the so-called ‘desktop metaphor’; the interface graphically displayed a space with folder, files, filing places, etc. in a congruent analogy with the way information is managed in physical office settings. The goal of this metaphor was to make the system understandable and easy to learn. A major problem is how to create appropriate metaphors. Thomas D. Erickson (1990) suggests a three-steps process:

- First the identification of the systems’ functionalities.
- Second the identification of the users’ problems in the interaction with a suggested system.

- Third the generation of a metaphor that adequately comprises and explains the system functionalities and put appropriate constraints so as to prevent users' errors.

In this sense the metaphor does not have a simple normative value but, more pragmatically, it should help the user in developing the actual understanding of a domain (Carrol, Mack & Kellogg, 1988).

The effectiveness of the use of metaphor in interface design is explained in terms of the human tendency to analogical reasoning (Carrol et al., 1988). Learning and understanding are strongly affected by the analogy existing between a system, with its whole set of functions, and its corresponding metaphor, which is concretized in the interface. However, it may happen that users assume the existence of analogies that do not really exist or, on the opposite, discard analogies embedded in the interface, but belonging to a different metaphorical domain. When a single metaphor does not hold, designers often make use of "composite metaphors" for a single system, providing different (and possibly puzzling) points of view for the system interface. This may be surprising but it perfectly reflects the limits of metaphorical language. In everyday language metaphors create partial and imperfect mapping between qualities (Lakoff & Johnson, 1980). Similarly, it is very unlikely that a single metaphor is complete enough to embrace a whole system. In some cases the mismatches are paradoxical: "a hideous failure of consistency is the garbage can on the Macintosh, which means either destroy this, or eject it for safekeeping" (Nelson, 1990). On the other hand, what it is interesting about metaphors is that they offer a rough structural skeleton that helps users in making sense of the system's functional organization. In that, both physical and cultural analogies are useful in offering the users a direction to develop a mental model about the system.

In the case of 3D interfaces the notion of metaphor has two implications. For instance many claims supporting the utility of 3D are based on the idea that 3D has a higher degree of similarity with reality (Robertson, Czerwinski, Larson, Robbins, Thiel & Van Dantzich, 1998). Therefore metaphors can be considered redundant since 3D aims to create interfaces that look like real things. Yet, the simple addition of the third dimension to an interface does not seem sufficient to support the users in the full understanding of the functionalities of a system since analogical language is used to bridge the system functions and the users' understanding of the system even in 3D interfaces (cf. section below).

6.1 3D interfaces and spatial metaphors

Examples of '3D metaphors' can be found in (Miller, Hetzler, Nakamura & Whitney, 1997) and (Miller, Wong, Brewster & Foote, 1998) in which geometrical and river-likes shapes are used to display large data sets. In

(Andrews, 1995) collections of documents are portrayed by three-dimensional colored blocks, while (Bray, 1996; Sparacino, Davenport & Pentland, 2000; Costabile, Malerba, Hemmje & Paradiso, 1998) uses landscapes, buildings and cities as 3D metaphors to display information. In (Robertson et. al., 1998) the metaphor of a mountain is used on which documents are visible on a slope, while in (Robertson, Van Dantzich, Robbins, Czerwinski, Hinckley, Risden, Thiel, & Gorokhovsky, 2000) 3D graphics are used to organize tasks, appearing as pieces of information hung on walls of a gallery. A more traditional approach is provided by Card, Robertson & Mackinlay (1991) in which a set of 3D rooms is used for data storage and retrieval; and in (Card, Robertson & York, 1996) where web pages are visualized in a 3D book embedded in an office, with a desktop and shelves used to store and retrieve data.

What emerges from these examples is that their creators seem to have a strong sense of the importance of a spatial continuity between 3D representations and their real correspondents. Representing information in a 3D fashion provides a strong analogy with real settings and the basic idea seems to be that this may let/make people perceive and act as they would if interacting with the real world which, in turn may enhance their spatial cognition abilities.

6.2 Spatial cognition: several factors at stake

Spatial cognition is defined as “the knowledge and internal or cognitive representation of the structure, entities, and relations of space; in other words, the internalized reflection and reconstruction of space in thought” (Hart & Moore, 1973). Spatial cognition is very important for the design of graphical interfaces. In information related tasks, people tend to make an extensive use of space to store and retrieve data. For example, the results of interviews conducted in two different studies (Barreau & Nardi, 1995) show that there is a tendency to prefer location-based strategies to logical strategies when people are demanded to search documents, and that place was a critical reminding function. Remembering locations seems to be relatively easy (Tversky, 2000) however the effectiveness with which places of items are recalled is sometimes questioned in the HCI literature. Location alone is not a sufficient strategy to organize information, especially when the amount of information increases. Jones & Dumais (1986) indicate that when location is isolated from meaning and task contexts, it is a poor retrieval cue, but the combination of spatial and symbolic information seems to provide some benefits.

As a matter of fact, there are several factors that impact on spatial cognition. Examples include, geographical relations among elements like similarity and proximity; the saliency of landmarks; the organizational structures

(Tversky, 2000). These factors affect the encoding, the storage and retrieval of spatial information (Winn, 1995).

Spatial cognition is also affected by logical and semantic distance (or proximity) that can impact on spatial tasks. For instance, in (Hirtle & Jonides, 1995) it was found that people tended to group buildings according to their semantic content (commercial, institutional, etc.) despite the fact that the buildings were spatially interspersed. Objects in space are related according to some functional dependency principles that depend on the content of the objects, on how they are related and also on the prior knowledge of the users (Gapp, 1995).

To summarize, mental representations of geographic spaces are only one of the dimensions of a more general semantic space, where objects are related semantically in memory (Hunt & Waller, 1999), according to different principles like hierarchical structures or familiarity (Kelley & Jacoby, 2000). This leads to the main idea of this thesis.

The effectiveness of space in spatial memory tasks seems dependent on many factors such as the content of objects to be memorized, their spatial distribution, their logical relation to other objects and the semantic and spatial proximity of the items. In addition more personal factors seem important, for instance the prior knowledge of a user and his or her familiarity with the system. Therefore, the incoherency of the empirical results comparing 3D interfaces to their 2D counterparts has to be viewed in this perspective. An analysis of the results considering only the spatial dimensions as the main rationale would be incomplete and superficial.

7. 2D vs. 3D: evaluations and empirical results

7.1 Spatial memory tasks

This section presents four empirical results from four studies that entailed the comparison of 2D and 3D interfaces across memory tasks.

The first work was an evaluation of Data Mountain (DM) (Robertson et al., 1998) a 3D interface designed to control the storage and retrieval of favorite Web pages. The interface shows an inclined surface, on which thumbnails of web pages are vertically displayed. When a simulated page is selected, it moves to a preferred position, central and closer to the viewer. The interface (there were actually two versions of it) implements audio feedback, visual landmarks on the surface, animation, and (in its newer version called DM2) a minimum distance separation that is kept among close pages (so as to reinforce the avoidance of complete occlusion of interposed thumbnails). A visualization mechanism that relates the selected page with its pop-up title (both of them becomes highlighted) is also available. The two versions of the 3D interface (DM1 and DM2) were compared to the Internet Explorer (IE) mechanism, across a retrieval task and involving 4 cueing conditions, the thumbnail, the title, the summary of the page and all the three cues. In the study the time (deployed to retrieve pages) and the accuracy (the number of failed trials and of incorrect retrieval) were analyzed. The results showed that the DM was more efficient in supporting the retrieval task than the IE interface. Moreover, the DM2 better supported the retrieval terms of time since the performance was as fast or faster than DM1 and than the IE group. However it is not clear whether the performance differences were caused by the interfaces' style (2D or 3D). In fact, in the study there were four factors which might have contributed to the results and that need to be discussed.

1. The first one is time. The authors state at the end of their work that the storage time was reduced using the DM but, as it appears from the study, no clear indications are available regarding the time given to users during the storage tasks (i.e. when users were demanded to store and organize their pages, before the retrieval phase). This is crucial since the time used to organize the information influences performances (Atkinson & Schriffin, 1968) when people are asked to retrieve that information. Moreover, the IE mechanism might have required scrolling the window to retrieve data (the authors stated that with the IE, the default amount of screen was indeed restricted). It can be assumed that in the IE condition users had to use part of

their time to scroll the window, instead of rehearsing the positions of the web pages, dividing their attention between the main task (memorizing web pages' position) and this secondary task (scrolling the IE window). Again time might have influenced the subsequent task.

2. The second factor is represented by the additional visual and acoustic elements of DM like the sophisticated animations, the landmarks, and the enhanced spatial sounds that allowed multi-sensory encoding of information. The influence of these variables in the performance was not analyzed separately. Therefore it is difficult to properly estimate their contribution in the spatial memory task.

3. The third one concerns the type of cue used during the retrieval, which had an impact both on the time and on the accuracy performances. For instance, the performance with IE was very slow when thumbnails were used as retrieval cue, while the performance with DM1 was negatively affected by the use of title as cues. This 'cue-type' effect had a similar detrimental effect on accuracy performances. That is, the number of failed and incorrect trials was higher for the IE group when the cues were the thumbnails or summaries compared to when title were used. On the other hand, performance with DM was adversely affected by the use of textual cues. This result can be explained in terms of the Encoding Specificity Principle (ESP) (Tulving, 1983), according to which memory is influenced by the correspondence between the information presented at the encoding phase with the information presented as cue during the retrieval. The DM2 was less influenced by the use of textual cue types for the retrieval especially when the cue was the title. Probably this was due to the implementation of a coherent mechanism that provided a 'visual link' between the encoding and the retrieval phase (a highlight effect was created to relate the selected page with its title). However, besides its 3D style the DM interface was relying on the use of pictorial data rather than textual information. Therefore, the role of the thumbnails had to be further studied to determine the contribution of the pictorial images in the spatial location of the web pages.

Another study (Czerwinski, Van Dantzich, Robertson, Hoffman, 1999) that consisted of a long-term follow up of (Robertson et al., 1998) was carried out six months after. For half of the trials the subjects were presented with a version of the DM that had the thumbnails 'turned-off' (white placeholders were replacing the images of the web pages). The same cues (summary, thumbnails, title and all the three cues) were used for the retrieval phase. A comparison between the results obtained in (Robertson et al., 1998) and the results of this new investigation was made. Unsurprisingly, the results showed that when the cues were thumbnails and when they were 'turned-off' in the scenes, the time responses were slower. 'Turned-off' thumbnails were also detrimental to accuracy performances. These results seem to confirm the ESP hypothesis in that they show that the contribution of the pictorial images was very important in the spatial task if compared to the first study (Robertson et al., 1998); memory for items can be improved if the encoding conditions are matching the retrieval conditions (Tulving & Osler, 1968).

4. Another factor that emerges from the detailed analysis of (Czerwinski et al., 1999) concerns the ‘familiarity’ with DM. The ‘cue-type’ effect impacted on the time performance but this effect was mostly present within the first blocks of trials and tended to disappear over time. An explanation of this effect could be the ‘familiarity’ that can be acquired spending time with a particular spatial DM organization. As a matter of fact, the subjects involved in this experiment took part in the first DM investigation, then they took part in another session 6 weeks after the first study; and finally they returned after 4 months to perform the final study (Czerwinski et al., 1999). The retention interval was therefore interrupted at different time stages and subjects had to carry out the task always using the same spatial DM organizations that they had built-up during their first exposure to the interface (Robertson et al., 1998). The subjects went through a process of re-learning of the pages’ spatial position, and this might explain the relatively good performance over time, even when the thumbnails were ‘turned-off’.

To conclude, (Robertson et al., 1998) and (Czerwinski et al., 1999) tried to provide some empirical evidence that memory tasks are better supported by a 3D-styled interface; nevertheless many uncontrolled elements could have influenced the results. If we assume that the interface style (2D vs. 3D) can improve memory tasks, then the interface dimensions should be studied separately, without the interference of confounding factors.

An attempt to more stringently investigate the 3D visual properties of DM was done in (Cockburn & McKenzie, 2001). This study varied only the number of dimensions of the interface and compared a 3D DM to a 2D DM. The task consisted of the storage and retrieval of web pages thumbnails, across three different information densities. In this study the cues did not vary and both the thumbnails and the pages’ titles were –at the same time- used as retrieval cues in the two conditions. The results, which accounted only for time responses, did not show any significant difference between the two interfaces for both for the storage and the retrieval phases. There was a significant time difference for information density in the retrieval, but, again, no interaction with the interface type. These results seem to contradict the first DM study (Robertson et al., 1998) since it was showed that no benefits can be gained from the use of the 3D DM for memory tasks, at least, when the information content and the cues are equivalent for all conditions and when only the spatial dimensions of the interface style vary. However the results of Cockburn & McKenzie (2001) can still be questioned if we look more in detail at the manner the 3D and 2D DM were designed. Basically, two main issues can be identified:

1. The first one is that –at least according to the photos provided in the article - the 2D DM implemented a misleading use of occlusion. The thumb-

nails are vertically arranged one behind the other, they are partially occluded, and the general effect is of a series of pages vertically arranged on a large and slightly tilted surface (whose edges are not visible). With such a representation very little is left to visually detect the difference between the 2D and 3D interfaces. It is possible that the lack of a clear difference between the two interfaces might have affected subjects' performance.

2. The second is the absence of a mechanism allowing for minimal separation among pages (in contrast with (Robertson et al., 1998). As it was also acknowledged by the authors, the high occlusion could have influenced the performance simply because when occlusion is high, the identification of objects is more difficult and retrieval time increases. This explanation was further supported in a follow up study (Cockburn & McKenzie, 2002) that implemented the partial occlusion algorithm to avoid full overlapping of pages if compared to the results in (Cockburn & McKenzie, 2001), reaction time was reduced by almost half.

This last study (Cockburn & McKenzie, 2002) also implemented several instances of the DM but it was more complicated than the previous since it involved three types of interfaces 2D, *2½D* and 3D. For every type of display there was a 'virtual' and a 'physical' implementation. The virtual interfaces were made of electronic representations visualized on a monitor; while the physical ones, were physical containers with real paper pages. The virtual 2D flat interface made possible a flat arrangement of pages (along the *x* and *y axes*). In the virtual *2½D* (so defined by the authors) interface the pages' arrangement was constrained upon a surface (vertical movement was not allowed) and in the virtual 3D interface, the pages could be placed inside a wire-framed space, in any direction. The 2D virtual and the *2½D* interfaces implemented a partial occlusion algorithm that constrained the pages overlapping, so that "pages with a lesser value on the y-axis are placed in front of the pages with higher y-axis". The physical interfaces were made of physical structures; the 2D interface was 900*710 mm; the *2½D* was the same as the 2D, but tilted of 25°; the 3D interface was a rods-made box slightly larger than the others and provided with 750 mm depth; the physical pages (or thumbnails) were paper pages of 90*90 mm. In coherence with the previous investigation, there was a main effect due to the information density, that is, denser amount of information required more time for the retrieval independently from the interface style. The other results did not provide any empirical evidence in support of the *2½D* or of *3D* virtual DM and actually, the performances with the virtual interfaces were rather poor; however, several factors involved in the experiments were not considered in the analysis, specifically:

1. As stated above, the partial occlusion (or minimal separation) algorithm implemented on the virtual interfaces, efficiently supported the retrieval tasks since time responses were faster both with the 2D and the *2½D*

by almost half. This means that avoiding the full overlapping of pages contributed to the results.

2. The time performance improved evenly across the different interfaces but no difference was found between the 2D and the *2½D* interfaces. Again, this may be explained by the misleading use of occlusion (c.f. point 1 above). As a mater of fact, the same algorithm was used for both interfaces therefore the graphical difference between the two interfaces was, once again, minimal and probably difficult to detect.

3. The physical interfaces supported a faster performance than the virtual interfaces. However, the difference between the physical and the virtual condition consisted of two things: the type of support (paper versus monitor) and the size of the supports (the physical surface and the paper web pages were bigger than their virtual co-respective). For instance, deriving from the data available in the study, the amount of pages that could be placed within the surface of the physical interfaces without any occlusion, was 78; while for the virtual conditions, the authors state that this amount goes from 63 to 71 pages. Moreover, also the size of the pages used in the test was quite different, ranging from a minimum of 40 pixels to a maximum of 130 pixels for the virtual conditions (aggregated) and 90*90 mm for the physical pages. With those sizes, it is very likely that the amount of things that the viewer could learn about the page and the number of graphical details available during the retrieval was (at least intuitively) higher in the physical than in the virtual condition. In the virtual condition, the pages were smaller and therefore less clearly visible. This might have resulted in interferences with previously or later stored items in memory (Roediger & McDermott, 2000) during the retrieval as the small size probably affected the ability to correctly perceive details in the target thumbnails.

3. Time responses were faster with the physical 2D than with 3D physical interfaces. The authors stated that performance decreased because of the increased freedom used to locate objects in the third dimension. But there could be also other explanations. In the study, subjects were standing 50 cm from the front-edge of the interfaces. However, the in 2D and in the *2½D* conditions (900*710mm) the pages could be seen clearly at that distance; while, in 3D physical interface this distance was more than doubled (as the depth of the interface was 750 mm). Perhaps the decreased performance was also due by the lower visibility of the pages that this interface offered.

To summarize, the objective data of the studies presented above seem to provide incoherent results since no clear evidence is given in support of a specific interface style (either 2D or 3D). The only common result among the studies outlined (Czerwinski et al., 1999; Robertson et al., 1998; Cockburn & McKenzie, 2001 and Cockburn & McKenzie, 2002) concerns the subjective observations of subjects' behavior. For instance, all the subjects tended to spatially arrange the pages creating a sort of grouping and coupling the place with categorization (semantic clusters). It is also briefly reported that the subjects identified spatial ‘qualities’ within the DM; i.e. the dimension closer (to the subject)/farther (from the subject) which was coupled to meanings like ‘more interesting’ (closer)/ ‘less interesting’ (farther). If people appeared to intensively use spatial arrangements while interacting with all the different versions of DM, how can the difference be explained? One reason is that the many confounding factors involved in the experiments had a decisive influence on the results of the 2D vs. 3D comparison.

Therefore, the following conclusion can be tentatively drawn from these studies: spatial cognition can probably be exploited by 3D interfaces however, their utility with respect to spatial memory cannot be determined *per se* and the following factors must be considered:

- Perceptual factors:
    - The size of objects is important in memory tasks, both with regards to the objects and to the space in which objects are placed. Larger objects containing detailed visual structure embedded in larger spaces allow the users to have a higher visibility on them, to learn more about the details of the objects to discriminate them more clearly therefore, making memory tasks easier.
    - Monocular 3D interfaces make use of pictorial cues (like occlusion, relative size, perspective, etc.). These pictorial cues are essential in the definition of 3D interfaces and have to be properly designed so as to guarantee a clear and visible distinction between 3D and 2D interfaces.
- Contextual factors:
    - Time is an important variable in learning and in memory tasks. In order to lower contextual differences in the comparison of 2D and 3D interfaces, the time variable should be kept equivalent and therefore comparable in all conditions.
    - Another way in which time affects performance is related to familiarity. People using graphical interfaces are well acquainted with 2D information structures; they normally perform storage and retrieval of data, using two-dimensional views of information. It is stated in (Cockburn & McKenzie, 2001) that 2D against 3D comparisons can be biased by the fact that people are not used to manipulate objects in 3D interfaces. Moreover, the relatively good performances obtained in (Czerwinski et al., 1999) can be easily explained by training and familiarity with the 3D DM.
- Factors related to semantics:
    - The semantic organization of the space. People tend to relate space with a meaning and to spatially cluster information accordingly; 3D can better support spatial related tasks when there is concurrence between the meaning of the space and the objects placed within the space.

- Objects placed within a space are characterized by semantic content and by a ‘generation code’ (for example, textual or pictorial or both). As indicated by the studies found in literature (Tulving, 1983) the continuity between the type of information that is stored and the type of information that is retrieved has a positive impact on memory tasks.

The theories in section 5.1, 5.2 and 6 as well as this summary of results, indicate that the utility of 3D interfaces in the domain of information management, especially concerning memory related tasks, is dependent on the above factors. One part of this thesis tries to investigate the difference between 2D and 3D interfaces in spatial memory tasks when the contribution of these factors is being controlled (as will be described in section 10).

7.2 Using three dimensions to explain logical notations

Another topic tackled by the thesis concerns the displaying of information in the specific domain of medical data. In one of the papers presented in this thesis, a new approach is described in the design of the interface to a medical database. A part of the database makes use of logical notations to explain the temporal development of diseases (the *temporal simulator*). Formal logic is an effective way to describe the development of diseases, but people are usually not very good in understanding formal logic (especially if they have only little or no specific background).

There are some studies that try to explain the difficulties that people face when they are requested to tackle problems involving formal logical reasoning (Johnson-Laird, 1993). In deductive reasoning people decide whether certain conclusions can be drawn starting from a set of initial premises. Examples of such types of reasoning are *syllogisms*, in which statements are organized into two premises and one conclusion, as the following example: *all men are mortal; Socrates is a man; Socrates is mortal.* This syllogism is valid and people can easily infer from the premises that Socrates is a mortal. But, while Socrates’ syllogism is easy to understand, some others are not so straightforward. Gleitman, (1995) reports another example: *all heavenly angels are accomplished harp players; some accomplished harp players are members of the American musicians’ union; some heavenly angels are member of the America musicians’ union.* ‘Logically speaking’, this syllogism is not valid, the conclusion is false; Gleitman claims that people often make errors when requested to judge the validity of such syllogisms because they interpret an implication as an equivalence. For instance, the first line of the syllogism above claims that all A (the angels) are B (good harp players) but it does not claim the opposite (that is, all good harp players are angels). It is this misinterpreted symmetry that is the basis for the mistake. There is a clear distinction between the semantic of an expression (that pertains its content), and the syntax that is composed of the set of rules that prescribe the construction of an expression. People tend to confound them and to rely on the *content* of the proposition rather than on its *logical validity*, (thus, we could imagine that some angels could be indeed members of the American Union). This bias is also central in the understanding of logical reasoning: a statement can be reasonable, but logically false; by way of contrast, a statement that is logically true, may not be reasonable at all.

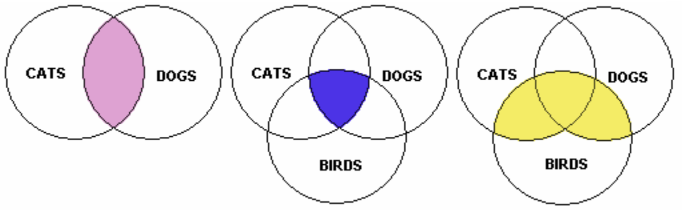
Propositional logic does not deal with the actual meaning of sentences, but only with the logical validity of the sentences (that is, weather a sentence is true or false). An example is given by the way the properties common to many (or all) propositions are described with logic. Instead of stating them for each single proposition, propositional variables can be used. Variables represent an arbitrary proposition and make statements about the properties in terms of those variables. The proposition $(P \vee Q)$ contains the variables P and Q; the value of the proposition depends on the value of the constituent variables (thus, the value of P and the value of Q). A table called the *truth table* represents those values. The truth table charts the value (either true or false) of a proposition for all the possible values of the constituent variables. In the example, P and Q are linked together by a special symbol, which is called *connective*. There are five connectives that describe the relationship between the truth-values (true or false) of a proposition and the values of its component (P and Q). In the specific case, one may refer to the truth table of the '$\vee$' (OR). But there is also a truth table for '$\neg$' (NOT), '$\wedge$' (AND), for the '$\Rightarrow$' (implication) and for '$\Leftrightarrow$' (equivalence). Beside the difficulties implied in the learning this, for most users, *new type of language* (e.g. the symbols, their meaning, the values described by the truth tables, etc…) there is also a difficulty caused by the limited capacity of human memory. For instance, the above proposition is very simple and constituted of only two variables (P and Q). With little bit of learning, it is easy to judge the validity of the proposition at glance (if the values of P and Q are known!). If we consider the following composite proposition: $(P \wedge Q) \wedge (\neg Q \vee P) \vee (P \wedge \neg Q) \Rightarrow (Q \vee \neg P) \vee (\neg P \wedge Q)$, it is, however, very unlikely that a judgment may occur at glance. The clusters of the proposition have to be analyzed sequentially, one after the other. But as soon as the second cluster is tackled, the memory of the result of the first one may decay or be displaced by the new information (Gleitman, 1995) and be lost.

To summarize, formal logic can be very hard to deal with for three reasons: a) implications are mistaken for equivalences; b) people tend to confound logical validity with how reasonable a proposition is; c) when logical propositions present a considerable amount of information, the ability of finding easy and 'at a glance' solutions is hindered by our limited short-term memory capacity.

One way to support people in understanding formal logic is displaying information graphically. In section 5.2 it was noted that graphical representations as the Venn diagrams can support people in understanding logical relationships between objects (or concepts) linked by connectives (specifically, AND, OR and the NOT).

In Fig. 1, three expressions that make use of connectives are graphically represented (from left to right, Cats AND Dogs; Cats AND Dogs AND Birds; [Birds AND (Cats OR Dogs)]). Venn diagrams display complex relational concepts and translate them into a graphical form, which is easier to understand (Michard, 1982). However, Venn Diagrams have a limited expressive power (the information that they can graphically describe cannot be too complex and the numbers of the terms is usually small); moreover, they specify relationships only across one dimension. Yet, the logical strings describing the temporal aspects of disease, had to be defined across three different dimensions: the symptoms, their frequency, and their development over time; thus choosing a three-dimensional representation of the temporal development of the diseases was an intuitive solution. However, every dimension specified in the temporal simulator is also multidimensional (e.g. there are several symptoms; there are different frequency levels; time is described along several days and weeks). Therefore a simple 3D representation was not sufficient to display the complexity of the temporal architecture and other features like color and shape codes were added to the representation.

Figure 1. Boolean expressions



Three-dimensional displays are not new in the medical domain. For instance three-dimensional interfaces are used for educational purposes like the learning of surgery (Molendi & Patriarca, 1992; Stone, 1999) or even tele-surgery (Satava, 1993). But these tools are highly experimental and they are not largely exploited. Moreover, they pertain to specific medical disciplines and are limited to very few medical tasks. Thus, it is not very surprising that the characteristics of the 3D temporal simulator makes the general appearance of the interface quite unusual for a domain where textual information and -at most- photos are used as supports to gather and understand data. The subjective evaluation of the 3D temporal simulator suggests that the utility of such an interface is debatable. The physicians who took part to the study clearly stated that the 3D style was too unusual and bizarre and that they preferred to rely on traditional textual information, independently from the complexity of the text. In the medical domain, especially in the case of experienced general practitioners, the choice of particular information tools is a function of the familiarity and the acquaintance with the tools. When familiarity is lacking the tools become suspicious, distrusted and their utility can be questioned.

7.3 Air Traffic Control

Recent research within the domain of in Air Traffic Control attempts to evaluate the utility of three-dimensional perspective interfaces in comparison to traditional 2D top-views (or simply 2D). However, also in this domain it is evident that no clear conclusions regarding the 2D vs. 3D issue can be drawn with regards to the utility of 3D interfaces for air traffic control. There are a number of elements that may affect the incoherency of the empirical results found in literature. Most importantly:

- The variety of tasks across which the 2D vs. 3D issue has been investigated.
- The types of interfaces taken into consideration (3D perspective, 3D stereoscopic, color coded interfaces and simple black and white interfaces, etc.).
- The different types of subjects involved in the studies, like controllers, pilots and subjects with uneven operational experience.

In order to focus on the domain of the Air Traffic Control tackled by the thesis, only the empirical results involving air traffic controllers will be reported. Some results derived from the evaluation of cockpit applications (therefore involving pilots as subjects of the study) will be also presented; however, these last results provide data regarding perceptual abilities that can be generalized, that is, they are not constrained by the specific activity of the subjects, either pilots or controllers. The need to discriminate between the two groups of subjects (controllers and pilots) and between two different activities (monitoring/piloting aircraft) can be explained as follows.

If we accept the perspective of this thesis and look at the problem in terms of contextual factors a clear distinction between the pilots and controllers tasks emerges. From a spatial perspective we can say that pilots are already embedded in a three-dimensional space, they have an egocentric view given by a ego-centered perspective of the space, and their activities require information extracted both from this egocentric view of the space and from the cockpit instruments (as well as from the indications spoken by ground control agencies). In contrast, controllers' position in relation to the space they manage is exocentric and external; and (at least for the en-route and approach facilities) their activities are grounded on *ad hoc* training to derive both 2D and 3D dimensional information from 2D planar interfaces. Thus, the problems at stake for these two groups of users are inherently different and difficult to aggregate coherently. This point of view is shared by Tham

& Wickens (1993) who questions the results provided by Ellis & McGreevy (1987) showing that the pilots made more avoidance maneuvers in the vertical dimension when given a perspective interface. Also, the same authors question the explanation provided for this effect, that 3D perspective interfaces, in large and general terms, present information in a more homogeneous way (Ellis & McGreevy, 1983; Haskell & Wickens, 1993). The critical point is the inherent difference between two different activities, the ones of the pilot and the ones of the controllers. Generalizations are difficult because different effects on performance can be observed on controllers and on pilots across the same tasks and same interface.

There are some results concerning biases when using 3D interfaces that is possible to generalize. Those biases concern perceptual aspects rather than background or activities. In general there are costs induced by 3D interfaces for tasks requiring lateral estimation (Tham & Wickens, 1993; Wickens, Campbell, Liang & Merwin, 1995) (Wickens & May, 1994) and for tasks impacted by the lack of height constancy, which is disrupted in 3D interfaces because of incorrectly estimated distances along the line of sight. More precisely, specific problems pointed out by previous studies are the following. The foreshortening (Perrone, 1982), which refers to fact that displacement in depth (relatively to the screen surface) is perceived as smaller than the amount of vertical or lateral displacement leading over-estimation in judging altitude relative to distance (Banks & Wickens, 1997). The loss of display resolution, which refers to fact that in 3D interfaces, a change in position of depth, will be perceived as a much smaller change in visual angle than a change in position of lateral or vertical separation. This has an impact on judgments related to accurate positioning and will degrade the perception of objects that are at the far end of a perspective display from the user's viewpoint (Boyer & Wickens, 1994).

Most other studies, however, present conflicting data concerning the utility of 3D interfaces in ATC; these studies usually differ in type of users involved and the type of tasks used. Here is a summary of these results.

1. The comparison between 2D and 3D perspective interfaces does not provide unequivocal results about a critical and important ATC task like conflict resolution. A study carried out by Burnett and Barfield (1991) showed that conflict resolution was quicker for the 3D interface only at lower traffic density (only 7 aircraft in each scene), but there are no benefits when the amount of traffic is increased (up to 17 aircraft). Another study carried out by Wickens et al. (1995) compared 2D, 3D perspective and 3D stereo interfaces across a number of tasks, among which the conflict resolution. The results did not show any difference in performance and this is contrast with a previous study (Tham and Wickens, 1993) that showed detriment in performance with the 3D interface. This last study was, however, criticized by Wickens himself because the 3D stereo interface did not produce actual fused images, preventing any benefit from stereo from being

realized. Nevertheless, a deeper analysis of (Burnett & Barfield, 1991) and (Wickens et al., 1995) reveal some differences between the two studies. First, the interface used in (Burnett & Barfield, 1991) implemented some color codes for the traffic; Wickens does not mention in his work any use of color code for the aircraft. Second, in Wickens' study, subjects were requested to make a prediction about a possible conflict, that is, judging if there was a loss of safe separation between two aircraft. In (Burnett & Barfield, 1991) the subjects were informed before the task if the conflict was about to take place and they were simply requested to choose the most appropriate solution to resolve the conflict (a list of 3 possible solution was presented by the experimenter). Thus, this last task entailed the judgment of the appropriateness of the conflict solution type, rather than a prediction of an event.

2. Empirical results show that the subjects' previous knowledge and their familiarity with a specific interface style may influence the performance, determining a more or less efficient use of 3D interfaces. This 'familiarity effect' was revealed in two different studies. In the first one (Wickens, Miller & Tham, 1994) 23 subjects were used, six were ATC specialists while the remaining were pilots only trained to ATC skills. The results show that while pilots were equally fast in extracting information from the 2D and from the 3D interfaces, controllers were quicker than pilots in extracting the information from 2D interfaces.

Moreover, the study simulated pilots' requests using three types of communications means:

- The verbal mode (radio).
- The print mode (textual messages in the screen).
- The spatial vector mode (vectors attached to the aircraft were graphically displaying the pilots request, e.g. a vector pointing to the left was equivalent to a 'turn west' request).

The authors discovered that combined requests (like change request involving both altitude and heading) together with the "vector request mode" condition were time-costly only for the 3D interface and that "this pattern of effect was not significant for pilots […] but it was for controllers and was not reflected in the accuracy data". The authors explain that such situation is particularly ambiguous since controllers cannot exactly determine the relative contribution of the vertical and the lateral components when both dimensions have to be derived form a single vector. But then, how to explain the fact that pilots were not affected by this problem? One explanation could be that controllers usually perform tasks with two-dimensional interfaces and that their conservative attitude in terms of safety might have imposed a time-consuming evaluation of the traffic picture. Another hint in this direction (given in the same study) is that controllers performed with more accuracy when discriminating whether the pilots' requests were safe or unsafe. Moreover, controllers were erring "on the side of caution" as it happened that they were rejecting a safe request even when it was safe. Similar behaviors were observed in (Wickens & May, 1994) where seven controllers and nine pilots were involved in terrain scenarios. Some advantages towards the 2D interfaces were discovered for controllers but not advantages for pilots (e.g. in the case of conflict detection between traffic and terrain). Again, this may suggest that there is a sort of negative transfer on controllers due to the past experience on 2D planar interfaces.

3. 3D interfaces produce detrimental effects for particular types of tasks. It was already stated in section 4.2 that relating distances along the line of sight to distances in other planes produce errors; and that exact judgments on these distances can be done only in a qualitative manner (Lind et al., 2003). This observable fact explains also to ATC tasks in which controllers are asked to make judgments on the lateral axes using 3D perspective interfaces. For instance, in (Wickens & May, 1994) tasks requiring the determining of the lateral distance between the aircraft and the potential threat of the terrain showed favor for the 2D and detriment for 3D interface. In (Wickens et al., 1995) subjects were asked to assign a single vector to avoid weather formation, the response time was equivalent for the two interfaces, but subjects were less conservative with the 2D interface, defining vectors closer to the weather (lateral distance was correctly estimated, thus the vectors definition was less conservative). In another task, having as a target a location point, subjects had to assign a new set of vectors to safely avoid weather (requiring changes in altitude or heading) and directing the aircraft to the requested point. For the 3D interface subjects preferred using a smaller amount of heading vectors while with the 2D interface the number of headings was higher and again, defining less conservative trajectories (that is smaller distance margins from the weather formation).

4. 3D interfaces imply another distortion that can impact on judgments performed on a same plane, for example the problem of height constancy (Tham & Wickens, 1993). In (Wickens et al., 1995), when subjects were asked to judge if an aircraft was about to penetrate a weather formation, there was no significant difference in the performance with respects to the interfaces styles for accuracy, but the 2D interface seemed to support more rapid discrimination. By way of contrast in (Wickens & May, 1994) for the task requiring estimation of terrain penetration, when the flight was not level but climbing or descending, there was a small time benefit using the 3D interface. It is unclear which specific characteristics the aircraft involved in the weather study had; therefore a straight comparison between the two cases is unfeasible. However some hints of explanations can be found in the special characteristics of the aircraft entering (or not) in collision with terrain formation. The authors claim that the time benefit with descending/climbing traffic was determined by “a more direct spatial extrapolation of changes [that] can be made with the perspective display”, while in the 2D “a fairly complex extrapolation of changes in digital data tag reading must be per-

formed”. Yet, they also state that time benefits are not obtained when flights are level, because of the perspective viewing disrupted altitude estimation. In other words, the ‘ecological’ displaying implies that altitude cannot be kept constant: as the aircraft proceeds farther away along the line of sight altitude will ‘appear smaller’. This affects judgments that require exact altitude estimation, and perhaps it might have been the case for the traffic entering weather formation. Contrasting evidence is given by (Burnett & Barfield, 1991) in which an altitude extraction task provided slightly faster performance for the perspective interface. However, the efficiency with which such task is carried out strictly depends on the geographical extent of the 3D traffic scenes. It is feasible to assume that the problem of height constancy can be reduced if the depth area is also reduced.

5. The last results concerns the subjective evaluation that the subjects involved in the experience with the 3D interfaces provided. Burnett & Barfield (1991) suggest that the controllers preferred the perspective interface in terms of “extracting immediate spatial situational and directional information”. Subjects of (Wickens et al., 1995) revealed that in general they were not satisfied with the 3D interface due to the visual clutter. Nevertheless, spontaneous comments of one of the subjects showed interest towards the use of 3D interfaces as training tool to understand the three-dimensionality of the airspace. Interestingly this last suggestion found empirical evidence. In (Wickens et al., 1995) when subjects performed estimation of terrain collision with the 3D interface first, they improved in their performance with the subsequent 2D trials, but the opposite effect was not discovered.

To summarize some general conclusions from these results may be drawn. There are several factors that impact on performance with 3D interfaces for ATC.

- Perceptual factors
  - When precise metric judgments are required, 2D interfaces provide a better support, as correctly projected distances are more clearly perceived on planar views. In 3D interfaces there is a systematic bias and large errors in the depth direction. Therefore, people can only provide qualitative judgments with high reliability.
  - A similar problem is shown for tasks requiring altitude judgments along extended depth; as depth increases aircraft altitude appears reduced; relying only on the ‘ecological’ appearance of levels may induce errors in judgments.

- Contextual factors
  - It seems that past knowledge and acquaintance with the interface plays quite a role in performance. As a matter of fact, from the previous studies emerge that the best performances obtained with 2D planar interfaces is mostly provided by controllers, but sub-

    jects with no operational background can perform equally both with 2D and 3D interfaces, showing no particular bias.
- Factors related to semantics
    - In the domain of ATC, space has a specific, concrete meaning and its understanding is critical to a safe ATC activity. Three-dimensional interfaces could, however, support a better understanding of the traffic and some of its three-dimensional spatial relations during training and this, in turn, could also positively impact later performance with 2D interfaces.

To conclude it is not possible to generally support a general utility of 3D interfaces in the domain of ATC. If there is a utility, it is only relative and dependent on the factors provided above. An initial analysis of the contribution of some of these factors is provided by the present thesis and it is described in section 10.

8. 3D interaction

2D interaction is based on the selection and manipulation of objects placed within a two-dimensional surface, while 3D interaction concerns actions taken in 3D spaces. A 3D space[3], by definition, entails three dimensions and six interaction variables that have to be controlled. For instance, in a 2D surface, objects are subjected to simple translation along the *y* and the *x* axes and, possibly, rotation around one axis. But in a 3D space, objects can be translated and also rotated along 3 different dimensions. It is understandable that new interaction modalities are required to adapt to this new and complex situation.

As a matter of fact several attempts has been done in order to support an intuitive and efficient interaction with 3D volumes. Novel input devices have been produced (e.g. the Teletact glove (Stone, 1991). These devices are different from the standard mouse commonly used for 2D interaction and allow more flexible and complex actions. Also new interaction techniques have been created. The interaction techniques attempt to map the users' inputs on the interaction device and to visually display the result of those inputs. Just to give a rough example, a 3D interaction technique can be compared to the pointer of the mouse which displays on computers' screens the output of a drag-and-drop action as well the movements followed to perform the action.

The domain of 3D interaction is quite vast, and a detailed description of both the hardware and interaction techniques of the state of the art goes far beyond the main theme of the thesis. Therefore, only a brief description of the interaction device (the Wand) and of the technique (the Ray Casting) that was used for the Papers III, IV and V is presented here. The Wand (Intersense) is a hand-held input device that contains many tracking sensors (used to detect the position and orientation of the user's hands), a single joystick device, and four control buttons (Dang, Hong & Tavanti, 2003).

The Ray Casting technique is displayed as a simple ray emanating from the Wand and, more in general, is based on the pointer metaphor. As a matter of fact, through the ray, users can firstly point, and then select 3D ob-

---

[3] In this section the definition "3D" refers to three-dimensional interfaces that make use of binocular cues, and not to *2½D* interfaces, which can be displayed on two-dimensional supports like ordinary computer monitors and for which a simple mouse usually would suffice.

jects. This technique is easy to use and is very close to the mouse pointer idea (thus, it could be easily approached by non expert users).

Some empirical comparisons of several interaction techniques (Poupyrev, Weghorst, Billinghurst & Ichikawa, 1998; Bowman, Kruijff, LaViola & Poupyrev, 2001) showed that the ray is efficient when users are requested to point and select objects that are out of the reach area, but there are also several limitations mainly imposed by the visualization, that is, by the way the 3D objects are designed and placed within the space. For instance, the ray is not efficient for the selection of very small objects (which are either too far or too little or -of course- both) and it has limited rotation capabilities, since rotation is allowed only around the ray axis. Two contributions of this thesis indicate that the ray casting technique is also ineffective when occluded objects have to be selected (cf. section 11.3 and 11.4).

Obviously it is possible to create 3D scenes where a simple algorithm automatically re-adjusts the positions of objects, preventing their total occlusion. But such a decision can be very risky, because it implies the alteration of the spatial relationships of the objects placed in the scene, which is not user-determined, but system-determined.

An example would be an interface that displays the trajectories of aircraft. Trajectories are simple lines that display the route of an aircraft. Every trajectory has to be provided with some visual means (a dot, sphere, a cube or any other shape) representing the waypoints, that is, the intermediate points of the line (trajectory) that can be manipulated so as to modify the trajectory path.

The distribution of the waypoints within the trajectory is naturally determined by the aircraft' routes and its modification can only be decided by the users (the controllers). If some waypoints are totally occluded, an algorithm that determines the alteration of their distribution could be implemented. But this means that the aircraft' trajectories would be automatically modified by the system (and not by the users). This, in turns, would determine a mistaken visualization of the *actual* air traffic.

This example suggests three things. First, the suitability of any interaction technique depends on the visualized content of the scenes and, more in general, on the purposes of the 3D interface. Second, the limits of an interaction technique cannot always be solved by actions that are system-determined. Third, when system-determined solutions cannot be chosen, other approaches have to be found to resolve the interaction problems. The description of an alternative approach is presented in sections 11.3 and 11.4.

9. Some words on methodology: analysis of users

In HCI there are several methodological views on how to capture users' requirements, how to analyze needs, abilities and how to support a user-centered design and development. The choice among methods is sometimes restricted by several reasons imposed by theoretical constraints but also by practical restrictions, limited resources, etc.

Providing an account of these constraints it is not in the scope of this thesis. However, the following section attempts to explain the reasons why three different methodological approaches were used in the present work.

9.1 Quantitative and qualitative methods

Experimental psychology has the ambition to be a science in that it attempts to look for general principles explaining behaviors and events. The basis of psychological measurements are attributed to Fechner's studies on psychophysics, the "psychology's first extensive, fully programmatic experimental research effort" (Wozniak, 1999). Fechner's methods were indeed quite solid, after more than a hundred years they are still widely used in psychology. The same approach is also used in HCI, as it provides results in the form of quantities that can be analyzed by mathematical and statistical means. For example, the results found in literature concerning spatial memory tasks and ATC (cf. sections 7.1 and 7.3) are gathered with this approach.

The original studies of this thesis, entailing spatial memory tasks and Air Traffic Control, were based on the analysis of the empirical data found in the literature. Methodological coherence and logical continuity between them were necessary, therefore these studies were based on controlled experiments and the analysis of the results was based on quantitative data.

Nevertheless, the 'experimental approach' has several limits and it has been strongly criticized. For instance, the socio-cultural theory founded by Vygotsky (1930) attacked behaviorist research based on the analysis of stimulus-response chains.

Vygotsky claims that traditional experiments are not sufficient to analyze activities, because they discard from the analysis the context in which the subject operates and acts. Experiments eliminate from the study something that is not objectively measurable but can affect people's behaviors and actions in real environments, such as their past knowledge, their habits, their preferences, and their social interactions. Therefore, qualitative methodologies like observations, content analysis, questionnaires, or interviews are used. Even though these methods are difficult to standardize, they are widely used and shared within the HCI community (Holtzblatt & Jones, 1991; Wixon & Holtzblatt, 1990; Meyers, 1999). Qualitative assessments like interviews and questionnaires were used also in this thesis to bring to light some data that the 'experimental approach' could not point out. For instance it was through interviews, that the air traffic controllers provided information about their working experience, their habits and their preferences (cf. Paper V). Through the same approach, the physicians interviewed provided data about their habits and their social interactions with peers (cf. Paper VII), that were later used in the analysis.

A third method used in this thesis entailed a usability evaluation without the direct involvement of users in the assessment (cf. Paper IV). The objective of the evaluation was the identification of possible problems in a novel interaction metaphor. Since the implementation costs were high, it was necessary to carry out an assessment quickly, without a working prototype. The cognitive walkthrough (Polson, Lewis, Rieman & Wharton, 1992) was chosen because it allowed a fast and economic usability evaluation, using as a basis the conceptual specifications of the system proposed.

To summarize, the most suitable methods at hand were chosen, depending on the constraints imposed by the context of the research question at hand.

10. Original contributions of the thesis

The previous sections reported some results that support the idea of a relative utility of 3D interfaces, which is dependent on specific factors such as the *perceptual factors* (which are related to human perceptual abilities (and/or limits) that can enhance or cause detriment to some specific tasks), *factors related to semantics* (which are related to the semantic organization of the space, of the objects distributed within it and of the relation among the objects and the space) and the *contextual factors* (which is related to familiarity, past knowledge of the interface and therefore to the time used to get acquainted with the style and functionalities of an interface).

In the light of these elements the incoherency of the results from the empirical studies comparing 2D *vs.* 3D interfaces was explained and analyzed with regard to the specific domains of information storage and retrieval (with special attention to memory related tasks and to the medical information domain) and Air Traffic Control. The empirical studies reported in the thesis, however, did not tackle the problem from the same perspective of this thesis. That is, the control of the factors that might have affected the results of the 2D vs. 3D comparison was discarded from the studies and was confounded and mixed with the investigation of the sole spatial dimensions.

This thesis attempts to provide empirical data to the general comparison of the 2D vs. 3D interfaces, but trying to control and analyzing the factors that were so far discarded by the researches at the state of the art. Specifically:

Within the domain of information storage and retrieval, especially in relation to retrieval and memory tasks the thesis attempted to:

- Control the *factors related to semantics*. Differently from (Robertson et al., 1998; Cockburn & McKenzie, 2001; and Cockburn & McKenzie, 2002), we defined a more precise organization (namely, hierarchical) of the space for the interfaces that were used, so that the spatial organization was not user-defined, but kept equal for every user. In accordance with the ESP principle (Tulving, 1983), coherence was maintained between the type (and the contents) of the information to be retrieved and the type of cue. The type of the contents was systematically varied and both symbolic and pictorial data were used and analyzed.
- Control the *perceptual factors*. The size and the visibility of all the items to be learnt (across both 2D and 3D interfaces) were kept equal and

comparable. Moreover, differently from (Robertson et al., 1998; Cockburn & McKenzie, 2001 and Cockburn & McKenzie, 2002), the 3D interface was deconstructed into single visual properties and then the contribution of 3D pictorial cues was evaluated separately.

- Control the *contextual factors.* Differently from (Robertson et al., 1998; Cockburn & McKenzie, 2001 and Cockburn & McKenzie, 2002), we used an equal distribution of the learning time across the instances of 2D and 3D interfaces, so that all the learning process and the level of familiarity with the interfaces was equivalent for all the subjects.

The description of the studies, the details and the empirical results are presented in Papers I and II.

Within the domain of ATC, the thesis attempted to provide both survey and empirical results. With the survey, we attempted to:

- Use a different approach from the one deployed in (Wickens et al., 1995; Wickens et al., 1994 and Wickens & May, 1994). For instance, before implementing a 3D stereo interface, an activity-based investigation was carried out. Real controllers were involved in a survey study with the objective of discovering which types of activities could benefit from the functionalities of 3D stereo technology.
- During the survey study, controllers were also engaged in some training sessions with the 3D stereo equipment and with the special interaction device. The problems related to the interaction discovered during the sessions were further analyzed and a few proposals to correct those problems were done. These proposals consisted of some new interaction metaphors. The results of a usability inspection carried out for one the interaction metaphor is also provided.

For the empirical studies, we tried to:

- Control the *sensory factors.* Identifying a task could be enhanced by the support of the 3D stereo interface, teasing apart the detrimental perceptual factors that most negatively ATC tasks.
- Control the *contextual factors*. Perform empirical studies with actual controllers as the main subjects of the tests, in order to examine to what extent the past experience and acquaintance with 2D planar interface affected the performance with the designed task.
- Control the *factors related to semantics.* An empirical *post hoc* comparison with non-controllers was carried out to investigate if the semantics (the specific ATC nature of the 3D stereo interface) affected the performance or if the results could be generalized to a wider extent.

The description of the framework, the details of the study, the survey, empirical results, and the usability study are presented in the Papers III, IV and V.

Concerning the management of information within the medical domain, the thesis tried to:

- Design and implement a prototype of an interface that translates strings of logical propositions into a more useful form using graphics
- Using 3D to display data across three dimensions and defining color and shape codes to display multidimensional information within the 3D interface.
- Perform a subjective evaluation of the 3D interface with the end-users of the application (experienced general practitioners).
- For this study, no other empirical investigation was carried out and therefore it was not possible to control the factors mentioned above. However, the analysis of the subjective evaluation indicate that a *contextual factor* like familiarity and acquaintance, could be extremely important in defining the level of acceptance of new information tools.

The description study, the details of the 3D prototype and the results of the evaluation are presented in Paper VI and VII.

11. Summary of the papers

11.1 Paper I: 2D vs. 3D, implications on spatial memory

The work was inspired from (Robertson et al., 1998), in fact the 3D interface used for our tests depicted simple rectangles arranged over a surface like in the Data Mountain. Similarly to (Robertson et al., 1998) the 3D interfaces was compared to a 2D interfaces across a spatial memory task (learning the pace of an alphanumeric character). However, our study tried to approach the problem in a different manner compared to (Robertson et al., 1998; Cockburn & McKenzie, 2001 and Cockburn & McKenzie, 2002). For instance, in these earlier studies no precise organization of the space was provided and the web pages' position were user defined during the storage phase. In this sense, the 3D interface of (Robertson et al., 1998) (and of all the follow-up studies) did not control the semantic organization of the space. Actually it discarded an important semantic characteristic of many interfaces that allow for storage and retrieval of data, that is, a hierarchical arrangement of information. Two-dimensional tree-views (similar to the Windows Explorer) are widely used to store and retrieve information. We thought of using 3D to create an 'ecological' representation of hierarchically arranged data (cf. the tree-view in Fig. 2). The *z* axis (depth) was used to indicate the position of the nested elements; the rectangles were placed in perspective from the user's point of view (so, the elements placed in higher positions of the tree were bigger, while the deeper ones were smaller). The rectangles in the tree were also properly juxtaposed so as to represent their logical distribution within the nodes.

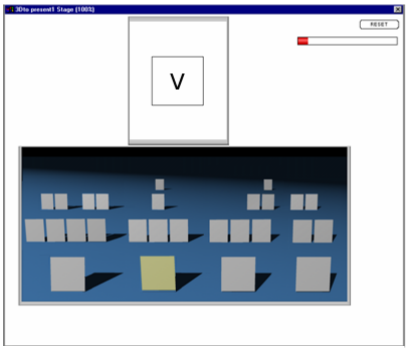


Figure 2. The 3D tree-view

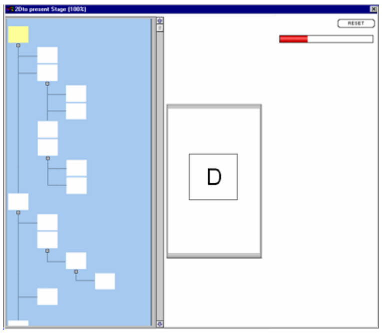


Figure 3. The 2D scrolling tree-view

This 3D tree-view was compared to a 2D tree-view (cf. Fig. 3), where the elements (folders or files) of the trees were displayed as small rectangles icons linked through thin dashed lines and the depth of the nested elements was expressed through their positions along the *x* axis. The semantics of the hierarchical arrangements was rigid and it was kept equivalent for both types (3D and 2D). Moreover, following the ESP principle, the same content was kept equivalent for every subject for the learning and for the retrieval phases and each of the rectangles corresponded to one of the 27 characters of the Swedish alphabet.

As previously stated, the perceptual factors involved in the task were also controlled. For instance, all the items displayed were evenly distributed and fully visible (while in the studies described in section 7.1 visibility was hindered by a misleading use of occlusion which could have created disparities in the way subjects could learn and locate the target tokens). Moreover, the content of the rectangles (the alphabet characters) was not texture-mapped on the rectangles (as it was for Robertson et al. (1998)), but displayed in a separated window, so as to eliminate visibility problems produced by textures mapped onto very small tokens. Both the interfaces were viewed by the subjects on a conventional computer monitor, but the graphical difference between the 2D and the 3D interfaces was extremely clear and explicit. The 2D interface was looking like a Window Explorer tree-view while the 3D interface really 'looked like 3D' since monocular cues like the relative size and the shadows were used to adequately provide a sense of depth and three-dimensionality. To reduce biases related to contextual factors, such as familiarity and time, an equal amount of learning time (two minutes) and time for the retrieval phase (five minutes) was allocated for each subject. In addition, between the learning phase and the task phase a questionnaire was used as a distraction task in order to assess spatial memory on a long-term basis. The study entailed two experiments. The first was a between-subjects experiment that compared a 2D scrolling version of the interface to the 3D interface described above, with 20 Swedish students as subjects. The results showed that subjects performed more accurately with the 3D display. However, the time allocated for the learning phase was very short and the subjects who performed in the 2D condition had to use a part of this time to scroll the window, while the subjects who performed the task with the 3D were able to exploit all the available time in the memorization task (since they did not have to scroll). This difference in the time allowance

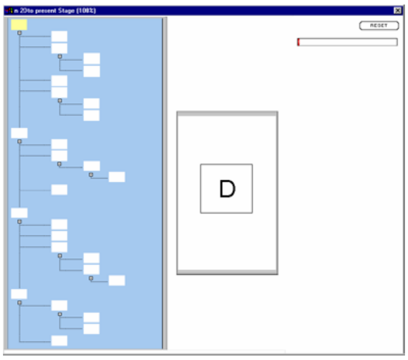


Figure 4. The 2D non-scrolling tree-view

could be the reason behind the poorer performance of the 2D group (cf. also section 7.1). Thus, in order to make the two conditions more fairly comparable, a second experiment was carried out. This second experiment was equivalent to the first one in terms of design, task, number of subjects, content and style of the interfaces, but here subjects were given three minutes for the learning phase and a 2D static, non-scrollable window showing the whole tree (cf. Fig. 4) was used. The results of this experiment showed, like the previous one, that accuracy was enhanced by the 3D interface, for the designated spatial memory task. However, the results of this study could have been influenced by factors that were not properly controlled. These issues were outlined and addressed in a second study.

11.2 Paper II: Deconstructing Space: Spatial Memory for symbols and pictures in 2D, 3D and in Deconstructed Displays

The results from the previous experiments indicated that a superior memory performance was possible with the 3D interface. However, a number of factors might have contributed to the results in an unintended manner. For instance:

- The population used was volunteers recruited amongst social science students at the university level, a fairly specialized one.
- A slight oversight in the randomization procedure increased the likelihood of a participant being selected to participate in one condition over the other if he or she didn't reply when called up the first time.
- The distraction task used might not have been distracting enough to prevent rehearsing.

Moreover, some conceptual questions were left unanswered by the first study. One important such question was that the rectangles varied in projected size when placed at different distances (the 3D size was kept constant). One possible explanation for the effect found could thus have been that these different 2D sizes acted as retrieval cues rather than the impression of depth per se. Another important question was whether the content of the placeholders (the letters of the Swedish alphabet) mattered. To check this, two new experiments were set up.

This time the assignment of subjects to each condition was randomized after the sample of subjects used had been fully collected. A more rigorous distraction task, backward counting, was used to totally prevent rehearsal of learnt items. Moreover, we tried to control the factors mentioned above, by doing the following.

First, the subjects were extracted from a different population, staff from the office of an international organization placed in Paris, France. This means that the subjects were slightly older and had a background from several countries.

Second, we decided to ‘deconstruct’ the 3D interface into single visual properties. Specifically, we took the most prominent cue of the 3D interface, the relative size (orderly arranged according to the linear perspective rules), we extracted it from the interface and then rebuilt it into a scrambled tree-view (cf. Fig. 5). This new representation presented the *y*-position of the rectangles disorderly arranged, so that the perspective effect was lost.

Then, two new types of information, iconic and symbolic, were used for the new experiments. In the first test, 27 small icons depicting every-day objects were used, in the second test the 26 characters of the English alphabet. Like in the previous study, the tree-view type was an in between-factor in the test, which involved 39 subjects (13 for each condition – 2D, 3D, scrambled). Besides the differences just described, the same procedure and materials used for the experiments reported in section 11.1 was used for these new experiments. Again, an equal amount of time was given to every subject both for the learning (three minutes) and for the retrieval phase (five minutes). But, it is possible that, in the previous experiments, the subjects of the 2D condition performed faster (before the five minutes ran out) than the ones in the 3D, with a negative impact on their accuracy performance. Thus, this time, the results comprised both the time spent and the accuracy of each subject. The results of the first experiment indicate that the subjects who performed with the 3D interface were significantly more accurate when compared to both the 2D and the scrambled conditions. The difference found between the scrambled and 3D tree-view was also surprising. Both of them are composed of rectangles with different sizes, which could have been used as mnemonics to perform the task. But the results indicate that -if the results hold- the relative sizes were more efficient mnemonics when ecologically displayed on a 3D surface. Moreover, no significant differences were found in the time performance among conditions, suggesting that the less accurate performances were not caused by a shorter learning time.

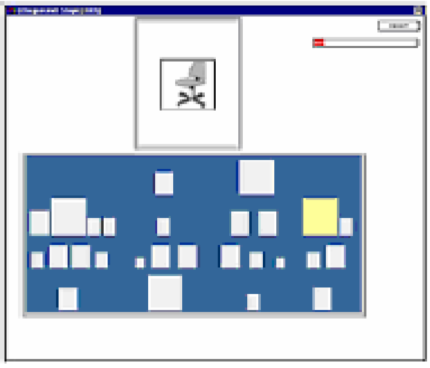


Figure 5. The scrambled tree-view

It is also interesting to remark that some subjects stated that they tried to find semantic links among the icons and that they attempted to discover similarities between the objects’ positions in the interface and the ones they know in real life. Perhaps the planar slanting surface provided by the 3D interface as well as the perspective arrangement of the rectangles, helped the subjects in visualizing the spatial associations among the objects.

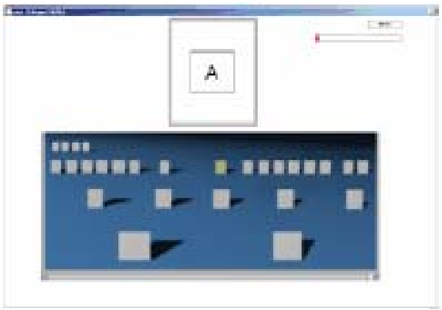


Figure 6.The new3D tree-view

The second experiment aimed, again, at comparing 2D, 3D and scrambled interfaces. But his time the contents to be memorized were different because the 26 characters of the English alphabet were used. In addition, a new 3D tree-view structure made of 26 rectangles was created using a random process. This new tree structure (cf. Fig. 6) was then visualized using the following generation rules:

- Every rectangle had to be placed (horizontally) on a planar surface.
- For every row and for every node, every rectangle had to be placed on the *x-axis* so that the centers of all items are equidistant from one another.
- A rectangle may not be occluded by another one.
- Every rectangle has a different 2D size according to its position in depth.
- Every rectangle should be juxtaposed so that items within the higher nodes should be at the center of rectangles composing the lower node.

The same hierarchical structure was applied to all three conditions. The same subject pool as in the previous experiment was used. Out of the 39 subjects 33 agreed to participate one more time and they were randomly assigned to the three conditions, eleven per condition.

The results showed once again a superior performance in the 3D condition compared to the 2D. This time no difference was found in terms of accuracy (the median were 14 and 15 correctly placed items, respectively), but the subjects in the 3D condition performed their task significantly faster (median values of 168 seconds for the 3D group was obtained and 196 seconds for the 2D group). It seems as though subjects prioritized speed over accuracy in performing this memory task a second time. The median time in the scrambled group was also higher than in the 3D group (187 seconds vs. 168 seconds), but this difference was not significant at the 5% level.

Several explanations can be provided for the different results between the previous experiments and this. First, a group of subjects extracted from the previous test group, took part to this second test. It is possible that boredom or other uncontrolled carry over effects affected the results. Second, the stricter distraction task imposed in the experiment could have more efficiently affected the rehearsal of characters compared to icons and thus, their memorization. Third, the new 3D tree-view created with a random process presented a more cluttered distribution of the rectangles than in the previous experiment (cf. Fig. 2 and 6) presenting a high concentration of items in the same row (the third), where the majority of the rectangles were placed) and perhaps producing higher interference effects in the recall task.

Finally, the icons displayed real and tangible things, which are far more concrete than the characters of the English alphabet. Also for this experiment, like in the experiments reported in Paper I, some subjects revealed that they had tried to use semantic strategies to recall the positions of the characters, like the creation of acronyms or pseudo words. Such strategies would probably work better in the two horizontal trees (the 3D and the scrambled conditions) and the efficiency of these strategies could have been unevenly distributed, due to the fact that subjects had different mother tongues.

11.3 Paper III: Stereo displays for Air Traffic Control: a multidisciplinary framework and some results

This paper is a general introduction of the Virtual Sky project. This project aims to empirically assess the suitability of 3D stereoscopic applications for the ATC domain and deploys an original framework of investigation. The paper discusses three main issues.

First, it gives a general introduction to the ATC field and a description of the main ATC tasks and of the main problems faced by controllers during their activities. The most important task of Air Traffic Controllers is the maintenance of aircraft separation. The term ‘separation’ refers to the safety distance that has to be kept both vertically and horizontally between aircraft (a/c) thus to avoid eventual collision. These safety distances are called *vertical* and *lateral separation minima*. To perform these basic tasks, controllers have to extract information from 2D top-views display. In these displays, the aircraft are visualized as simple moving dots within the monitor. For every aircraft the basic information, strictly necessary for the monitoring tasks, (like the aircraft call-sign its altitude, its speed, etc…) is displayed in ‘aircraft labels’ by means of symbolic expressions, that is, with simple text and numbers. Visualization is an important issue for ATC: current radar screens display not only aircraft tracks and labels but also other flight related information such as flight plans, airspace (sectors, routes and beacons) as well as the interfaces of all decision support tools that attempt to assist controllers in their tasks. In the current two-dimensional displays, controllers often report the problem of the ‘cluttered picture’ referring to radar display encumbered with too much information represented by too many graphical objects. The visual clutter is often perceived as disturbing, distracting and a cause of detriment to the monitoring tasks. Figure 7 shows a generic sector with two levels of traffic density, the first one with low traffic and the second one with high traffic. It is evident that the control tasks are more difficult in the second situation. Three-dimensional displays could provide some support to the visualization problem described, because they take advantage of the depth as an extra dimension to display the data. However, the actual applicability of the 3D stereo technology has to be investigated. This consideration leads to a second issue discussed in the paper, that is, the modality of investigation and the research framework.

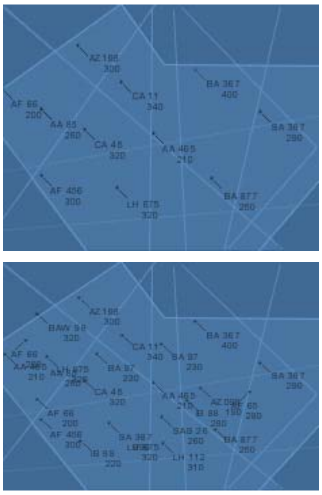


Figure 7. Low and high traffic density

The research framework is articulated into 3 main components: the visualization (that aims to define how ATC objects should be graphically represented in a 3D stereo display); the *interaction* (that aims to investigates the appropriateness of interaction metaphors and devices with a stereoscopic display) and the *human factors* (that aims to carry out users' evaluation to assess the suitability of visualization and interaction design choices). The interaction seems particularly interesting in this context. As a matter of fact, interaction with 2D interfaces makes use of the WIMP paradigm (Windows, Icons, Menu and Pointer), which is commonly used by most computers users. However, the 3D stereoscopic displays may require replacing the mouse and keyboard-based interactions with *ad-hoc* 3D input devices. As a matter of fact, the equipment available for the Virtual Sky project comprises both special visualization and interaction technologies, specifically a 3D stereo equipment composed of a projection system; shutter glasses to visualize stereo scenes and a special interaction device called the Tracked Wand.

The third issue discussed in the paper reports some preliminary results of a special session carried out with nine controllers. The session was twofold. First, we wanted the controllers to become acquainted with the stereo equipment and to train them to the special Interaction device (called the Tracked Wand) used in the stereoscopic environment. Second, controllers went through an interview aiming to identify some ATC activities that could effectively make use of the 3D stereo technology.

This session was very important because it allowed the discovery of problems concerning the interaction device. We were aware that our future research and forthcoming empirical investigations had to be carried out using this device; therefore it was also necessary to fix the problems and to provide some solutions.

The first problem concerned the ergonomics. During the training the controllers declared to feel tiredness to the upper part to the right arm caused by the intense movement of the arm and the wrist to manipulate the interaction device.

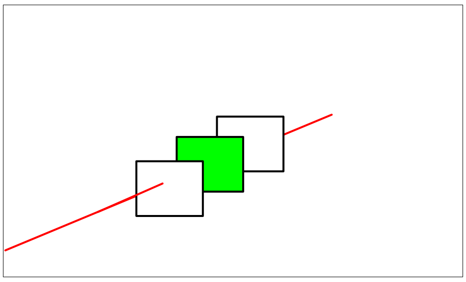


Figure 8. Ambiguous collision detection

The second problem concerned the usability of the interaction metaphor used in the Tracked Wand. The Wand is a special tool that can be held with one or two hands and from which a red ray (colored in black in Fig. 8 and 9) emanates. This red line is used to point at objects that can be selected by pressing a button located on the Wand. The selection of the object is determined by the collision between the ray and the object. When a user wanted to select a target object (like the gray rectangle in Fig. 8) occluded by other objects, the ray intercepts all the objects lying within the same trajectory and this causes several collisions to be detected. However, only one object at a time can be selected and sometimes another object than the target object was selected. In order to overcome this problem, three new interaction metaphors were elaborated. One of them was called the Elastic Wand. In the Elastic Wand the ray can turn and its curvature can be adjusted (cf. Fig. 9).

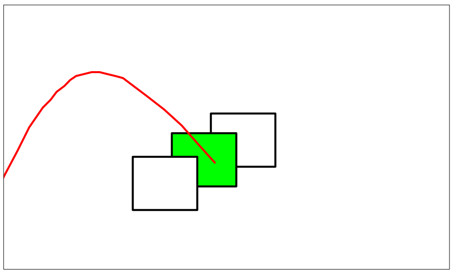


Figure 9. Elastic Wand

As in the current implementation, an end of the ray is fixed to the Tracked Wand and it moves with it. But the ray could be curved according to the user needs, by interacting with a pin-like joystick that is placed at the center of the Wand. In this manner users could more easily reach the desired objects without encountering the problem of ambiguous collision. The other two interaction metaphors (the Transparent Sphere and the Transparent Cylinder) implement more sophisticated solutions, e.g. transparent geometrical shapes and floating menus aiming to overcome the occluded objects problem. At the end of the paper, some preliminary results of the interviews were reported. It is worth pointing out that these interviews also had the goal to investigate the possible application activities of the 3D stereo technology. Controllers declared that they cannot envision a 3D display for standard monitoring tasks, like the approach or the en-route, but that other types of activities could be supported by the stereo 3D, like: tower operations, training, stack management and the traffic allocation within a sector.

11.4 Paper IV: Usability Inspection of a 3D Interaction Metaphor

In 3D stereoscopic environments the traditional mouse and icons interactions are sometimes obsolete and new ways of interaction may be required. During the studies carried out within the Virtual Sky project a new set of 3D interaction metaphors was proposed. These new metaphors were thought to overcome interaction problems encountered by the users during a previous training session. One of these metaphors, called the Transparent Cylinder, was evaluated by means of a usability inspection technique, the cognitive walkthrough.

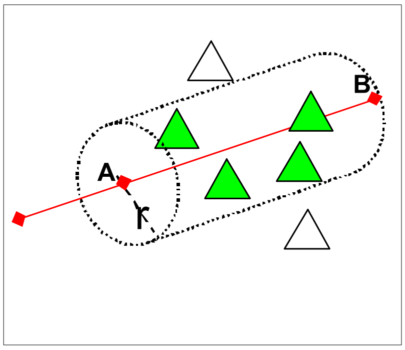


Figure 10. Transparent cylinder

The transparent cylinder, created by Dang (Dang et al., 2003) can be described as follows (cf. Fig. 10): it is a selection method that allows highlighting of all objects within a given radius [**r**] from the ray, the radius defines a circumscribed cylinder. When the desired object becomes one of the highlighted, a button can be pushed placed on the interaction device (the Tracked Wand). Then a menu appears at the closest distance from the cylinder. With the menu, the problem of selecting occluded objects (that caused ambiguity produced by multiple collisions) is solved. At the time of this investigation only a partial implementation of the cylinder metaphor was available, so it was compounded with a low-fidelity prototype (paper made) where all the system specifications were described. The cognitive walkthrough (Polson et al., 1992) method was chosen because it allows the identification of problems and flaws in the concept of a new system, at the early stage of its development, when only few specifications of a system are available.

The result of the inspection indicate that despite the apparent simplicity of the idea, the cylinder metaphor requires multiple steps to be done in order to perform selection tasks. Specifically, in order to select one object a hypothetical users has to:

1. Pass the cylinder through the scene in order to highlight the objects.
2. Press a button of the Wand to set and freeze the highlight on the area where the target object is and to open the menu.
3. Select an item of the menu by pressing a button on the Wand.

Moreover, the analysis revealed that some visualization details needed to be more precisely defined. For example, how should the length of [**r**] be determined? And also, according to what rule should the distribution of objects in 3D space map the ordering of the objects displayed in the menu? And, if an object is totally occluded, it may appear in the menu but it would not be visible in the scene, this situation could be puzzling for a hypothetical user.

Along with the identification of the problems, the paper presents some proposals that could be implemented in the final version of the interaction prototype. For example, it is suggested that the length of [**r**] could be customizable and adjustable, so that, the users can set the size of the highlighted area according to their preferences. Also, the order of the objects in the menu could map the geographical distribution within the scene (e.g. the waypoint which is farthest away from the user could be listed as the first one in the menu) or simply follow alphabetical order. Another idea could be the implementation of a ‘double highlight’, so that when the ray passes over the menu both the name of the item and the corresponding object in the scene could change their color (underlining the unique correspondence between them). Finally, the paper suggests the necessity of visually prompting the presence of occluded objects in the scene, in order to indicate that the menu possibly contains the name of a ‘hidden’ object. For example putting a small flag on the top of the object, or implementing semitransparent objects, so that occluded items can be seen through the occluding ones.

11.5 Paper V: 3D for ATC displays: ask controllers, they know better

This paper consists of the analysis of a survey study conducted with controllers and of two experiments carried out with subjects with and without operational experience.

The survey study present the results of the questionnaires and interviews performed during training sessions, at a previous stage of the research. The aim of the sessions was twofold. 1) Allowing controllers to get more familiar with the new system. The system requires some knowledge on behalf of the user about the way information is presented and accessed (visualization and special equipment to wear) and about the interaction modalities allowed by the system. 2) Believing that only some operational activities could benefit from the use of 3D stereo interfaces, we tried to gather direct feedback from controller on this issue. Nine former operational controllers were involved in the study. Information about the suitability of the interaction device was gathered, as well as a collection of possible application domains that could feasibly benefit from the 3D stereo technology. The results of the survey suggest that controllers judge the use of 3D quite negatively at least for the tasks requiring standard monitoring activities like en-route and approach. Yet, for some other activities, controllers were able to envision an efficient exploitation of such technology.

Specifically, the activities proposed were:

- Tower operations. In towers the control is done more 'by eye' than in 2D planar interfaces. 3D would then be a useful support to offer a clearer "out-of-the-window" view of the traffic situation, in poor or misleading visibility conditions, caused by bad weather or by nocturnal vision.
- Training purposes and traffic allocation within sectors. 3D would be profitable for training purposes, as it offers an explicit and ecological visualization of the four-dimensional problem exposed by air traffic. Traffic allocation activity requires controllers to have a 'global vision' of the traffic, which probably can be achieved without getting into exact metric judgments that could jeopardize safety maintenance (cf. the biases produced by the exact lateral estimation of distances described in section 7.3).
- Stack management tasks, which do not require controllers to make judgments on lateral distances but vertical arrangements. Nowadays stack management is done on 2D planar interfaces and the vertical information is extracted by means of data labels associated to each aircraft. 3D might give a more intuitive and clear visualization of the vertical positions of every aircraft.

Since it was important to gather some more precise data on controllers' performance with the 3D stereo interface, after the survey the traffic allocation activity was selected as the target of our research. Then, with the help of a controller, a very simple task related to such activity was defined. This task was very simple and motivated by the following reasons. First, this task was inherent to the scenario identified. For instance, the vertical separation is the basis upon which safety is maintained; consequently, when controllers have to balance traffic among sectors, it is crucial to perform this balancing by primarily focusing on vertical separations. Second, this task took into account an important semantic property of the traffic, its altitude, which was naturally displayed in 3D (by using graphical means and not only by means of simple symbolic characters). Third, this task attempted to control the perceptual factors, tearing apart tasks involving lateral judgment that, according to the literature results, are usually biased with 3D displays. We also decided that having actual controllers as subjects for the test was crucial. In fact, as previous literature studies showed, performance with 3D interfaces could be affected by contextual factors like the past experience and the acquaintance with 2D planar displays.

Yet, it was necessary to maintain a certain degree of realism with the content of actual traffic scenes, thus a controller helped to design some generic traffic scenes (cf. Fig. 11; the original scenes were not in black in white but colored).

Fourteen former operational controllers were engaged in a judgment task: they had to identify critical flight levels (that is, traffic flying at the same level) within each scene. The task was evaluated across two conditions: 3D stereo and 2D. After each test, controllers were asked to fill a questionnaire where they had to self-estimate their performance in terms of time responses and accuracy with the two interfaces. The objective measurements indicated that controllers were quicker in identifying the target in the 3D stereo than in the 2D condition and that accuracy was equivalent in both conditions. Perhaps surprisingly, the self-assessment data showed the same pattern: controllers estimated that their performance was quicker (but again not more accurate) with the 3D stereo interface.

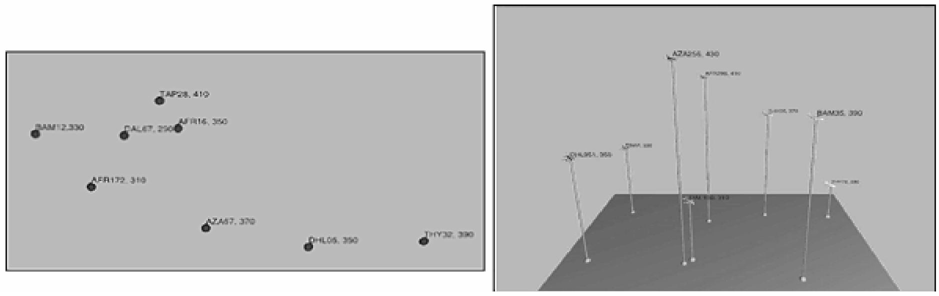


Figure 11. 2D planar and 3D stereo display

The quicker performance with the 3D scenes seemed to show that the controllers made use of the natural representation of altitude causing a decrease in time performance. However, at the beginning of the test we suspected that the acquaintance with the 2D planar interfaces could negatively affect the performance with the 3D stereo display; this hypothesis was not confirmed.

An explanation could be that the controllers who took part in the test are not operational nowadays; there is a temporal gap between their actual activities and practice on 2D ATC interfaces. Whether this gap affected the results was unclear and had to be further investigated. It was not logistically possible to involve operational controllers in the study; thus we decided to run more tests with subjects lacking operational background. This choice was motivated as follows. First, the controllers, despite their detachment from operational activities, still have operational knowledge; therefore it was possible that the specific ATC content might have played a role in the task; having non-controllers as subjects allowed us to gather more evidence in relation to this hypothesis. Second, we wanted to test if the specific semantics of the scenes (based on ATC content) could have affected the performance, or, by way of contrast, if the 'ecological' properties of the 3D scenes affected the performance, independently from their specific ATC content.

A second experiment was therefore set up. All the variables of the experiment were kept equal, but this time thirteen subjects with no operational background were engaged. Both the subjective and the objective measurements were consistent with the evaluation of the first experiment. The results again suggested that for the designed task the 3D stereo interface supported a quicker performance. No significant difference in terms of accuracy was found between the two groups of subjects.

We can also speculate that the judgments made along the vertical axis can be enhanced by the use of the 3D stereo. However this type of judgments also involves the height constancy bias (c.f. section 7.3). The scenes used in the experiment did not display a very deep area, so the height constancy bias was not very striking (but still present as it was revealed by the comments of some subjects involved in the study).

Such problems could be controlled if we still think about the original activity scenario, traffic allocation among sectors. Since this activity requires an exploration of the traffic, we can envision overview and zooming functions to access local information, limited to precise geographical spots. In this way, the interaction may help to resolve the height constancy bias.

11.6 Paper VI and VII: Using 3D to visualize medical data II & Representing medical information: a subjective evaluation

These papers present the specifications of a medical database that was designed to support doctors and medical experts in searching and retrieving information about diseases. The database was designed in cooperation with two medical experts who helped to define the users' requirements. It uses a quite new approach in that it implements a structured definition of the disease according to six defined categories, the etiology, the pathogenesis, the clinical symptoms, the therapy, the history, and the differential diagnosis. All information related to each category is briefly explained in short sentences. Moreover, these sentences can be linked to images or text files that allow a user to access more detailed data. The database also implements a 'temporal simulator' that provides a temporal description of every disease according to the *symptoms*, their *frequency* and their *development over time*. The description of temporal development of disease uses logical connectives. An example is provided by the following proposition: *if* ($\mathbf{S9} \wedge \neg\mathbf{S1}$) then $\mathbf{Th5} \wedge (\mathbf{Th6} \vee \mathbf{Th7})$, where **S** is symptom and **Th** is therapy. But logic can be difficult to deal with and it may not represent a useful option when information has to be grasped at a glance.

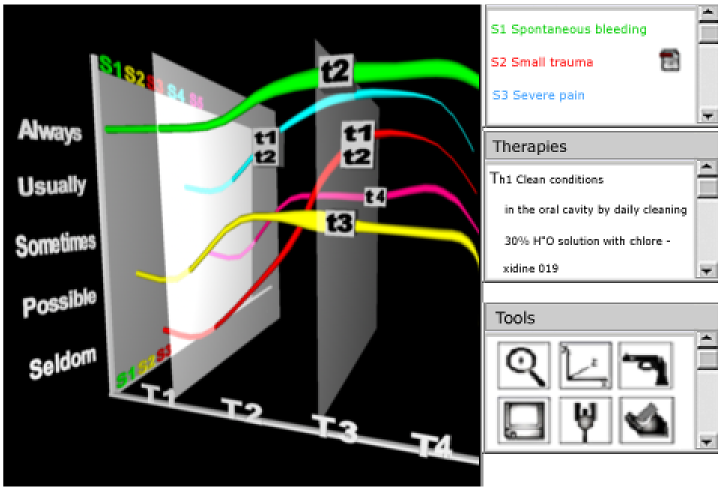


Figure 12. The 3D temporal simulator

The graphical solutions that are commonly used to graphically display logical relations, like the Venn diagrams, have too little expressive power since they only allow the representation of simple two-dimensional data. In the simulator every disease is described across three main axes (symptoms, frequency and development over time), therefore a more expressive graphical means had to be found. Three-dimensional models can offer more expressive possibilities.

Thus, a prototype of a 3D interface able to translate the logical notations into graphical objects was designed. The prototype implemented the simulator as a three-dimensional matrix (cf. Fig. 12). On the x-axis the temporal units (T1, T2, etc.) are listed. On the z-axis are the symptoms, which were differently colored. On the y-axis are the qualitative attributes that define the frequency of a symptom to be found, the attribute *always* placed at the higher position of the y-axis, while the attribute *seldom* takes the closest position to zero. The severity of every symptom (strong and acute, for example) is defined by the thickness of the snake-like item.

The 3D temporal simulator had to be evaluated, but it was not possible to carry out an objective evaluation, for example comparing subjects' performance with the 3D interface and with a textual version of the logical strings. As a matter of fact, on one hand, it was necessary to evaluate the prototype with real physicians, since the information contained in the interface is very specific to the medical domain and needs to be precisely understood. On the other hand, the physicians contacted for the study did not have any knowledge about formal logic. An objective assessment would have required the physicians to be trained in logic prior any objective evaluation, which was not logistically possible. Therefore a subjective evaluation involving six experienced physicians was carried out. For the evaluation a demonstrator of a well-known disease was implemented. The demonstrator displayed the two instances (textual/logical and 3D) of the disease' temporal simulator. The physicians were shown the demonstrator and the main purpose of the database and of the simulator was explained. Then they were interviewed and asked to fill a questionnaire. The questionnaire aimed to gather information about the physicians' background, their level of acquaintance with logic, their working activities, the tools that they usually use to gather information about their field, their opinion about the utility of the medical database, their opinion about the appropriateness of the structure of the database and about the 3D temporal simulator

The results of the evaluation indicate that the physicians spend a relevant amount of time in gathering medical information to be up to date with new medical discoveries. They estimated also that their experience and knowledge is usually sufficient to perform correctly their activities with patients.

They also stated that they like using books, papers and chats with colleagues to quickly access to the desired information. Alternative sources like the Internet are also used, but limited to trusted sources like databases of abstracts and articles.

Finally, they shared the idea that integrating medical information into a single database would be good, but they suggested that it would be useful only for limited and specific pathologies like skin or ear and throat diseases. In fact, these diseases have a more specific and codified structure, which would be easily described through the database categories. Moreover, especially the skin diseases have a more structured development over the time and the temporal simulator could nicely provide this sequential development of the symptoms. However, the choice of 3D to display the simulator was strongly questioned. The 3D interface was judged unusual, strange, and weird. Despite their lack of background on logic, the physicians estimated that it would be easier to understand the logic notations than grasping the information from a 3D graphical representation.

The number of subjects used for this study was small; therefore it is difficult to claim that these results are generally valid. However, they provide some useful hints and some speculative interpretations of the results can still be done.

All the physicians had a very long time experience as general practitioners (about 27 years). When they were asked to judge whether the logical strings describing the disease were easier (or more difficult) to understand than the 3D, they replied already having a precise knowledge of that disease. It is possible to think that the knowledge content was confounded with the means deployed to express this knowledge, determining the underestimation of the difficulty that logic implies. Also, the medical knowledge acquired over the years was mostly based on texts. Perhaps the textual nature of the logical strings inspired a higher degree of confidence and familiarity than the weird 3D interface. Familiarity can be defined as the “the extent to which a user’s knowledge and experience in other real-world or computer-based domains can be applied when interacting with a new system” (Dix et al., 1998, p.261) and it could plausibly explain the physicians’ negative judgment concerning the 3D interface. Over the years, the physicians had been able to create a sort of ‘knowledge deposit’ that they strongly trust and which is composed of the acquired medical information and of all the tools used to gather this information (e.g. the books, the talks with the colleagues, or the medical articles and abstracts). After so many years, the gathering of information through these tools becomes a routine process. The introduction of an unusual technology in this secure and trusted process is perhaps judged as disturbing. Then, it is not surprising that one physician advised to continue the evaluation with young doctors or with medical students who might be more used to these “bizarre computer things”.

12. Discussion and conclusions

This thesis attempts to explore the utility of 3D interfaces mainly for two application domains, information retrieval activities involving spatial memory tasks and Air Traffic Control. The choice of these two domains is not purely coincidental. For instance, these domains are characterized by an always-increasing amount of information (more and more documents to be managed and denser traffic densities to be managed). Many existing studies suggest that 3D interfaces could be the 'interfaces of the future' for these two domains, since the third dimension allows the displaying of higher amount of data simultaneously, when compared to 2D interfaces. Therefore, 3D could better support the users in their activities. However, is it really true that '3D is better than 2D?' Despite a large amount of empirical data this question cannot be clearly answered. The results are incoherent and contradictory. The explanation put forward in this thesis is that is because the question is badly formulated. The analysis performed suggests that the 2D vs. 3D issue can be tackled only if a more accurate question is formulated: given a particular domain of application, under which circumstances and across which tasks can a 3D interface contribute to an improved performance? The original studies of the thesis try to address this question. Three groups of factors that should be considered in the definition of the utility of 3D interfaces were identified.

In this section they will be presented in relation to the domain of application chosen for the thesis, first dealing with information tasks and then with ATC. To do so, we will take into account the results of the analysis presented in section 7 and the results of the original studies carried out for this thesis.

The first group of factors refers to human perceptual abilities and has a major influence in the comparison of 3D vs. 2D. This group of factors defines the minimum requirements to be respected if a fair comparison is to be done. The analysis provided in section 7.1 (concerning retrieval and spatial memory tasks) suggests that the design of the 3D interfaces should be very precise, so as to allow a realistic comparison with 2D interfaces. The displaying of monocular cues should be correctly designed and controlled and the difference between the types of interfaces should be visually obvious. Likewise, any source of memory interference (such as the confusing low discriminability of certain objects to be memorized) should be avoided. The projected size determines the visibility of the objects' details and their degree of occlusion. Logically, learning the positions of cluttered objects cannot ease memory tasks, it would only reproduce the problem that is the main cause of inefficient information retrieval with 2D interfaces: being unable to retrieve things easily. When an explicit graphical difference between interfaces is provided, differences in spatial memory performances can be found and, at least under some circumstances, this difference can be favorable to 3D interfaces. These circumstances are suggested by the experimental results presented in Paper I, and II. Besides the interface type, other factors have to be taken into account when spatial memory tasks are at stake, first the semantic structure. The results presented in Paper I suggest that when presented in a spatial structure that also had a semantic function (that is, a hierarchical tree view), the positions of Swedish characters were recalled more accurately in a 3D view than in a 2D. Similarly, a time advantage for the 3D tree-view in comparison to a traditional 2D view appeared in the second study -Paper II- despite the several constraints imposed by the experiment (e.g. the same subjects taking part to the test, the harder distraction task, the choice of the English alphabet joined with the uneven users' acquaintance with those characters).

Also the specific content of the objects played a role in memory tasks. The analysis of (Robertson et al., 1998) and (Czerwinski et al., 1999) suggest that the relative contribution of pictorial content was important for the effective accomplishment of the memory task. An indication in this sense is suggested in the Paper II. The idea that the content matters and that concrete things are more easily recalled than abstract things is certainly not new (Paivio, 1965). The interesting issue is how to effectively couple the meaning of objects with the meaning of space, especially of 3D spaces. The results of (Robertson et al., 1998) and (Czerwinski et al., 1999) showed that people made a strategic use of the space since every information cluster was arranged in a special portion of the space, which was later used during the retrieval. The same behavior was noted in the first test of the second study - Paper II-, as subjects declared to have a tendency to create associations between the content of the items and their position in space. Some of the subjects attempted to discover links among the pictorial tokens of familiar objects, creating associations between the spatial structure of familiar spaces and the ones displayed in the 3D interface. Perhaps, the distribution of the objects in the 3D interface preserved some properties of real world settings. This also suggests that memory could be enhanced by 3D interfaces, if there is a meaningful coherency between the objects and their places.

Familiarity is another important issue to consider. Familiarity has to be considered in a wider sense than the experience with computer systems and their interfaces alone and it is not a neutral factor in the performance with any type of interface. Support for this can be found in (Czerwinski et al., 1999) where the increased amount of time spent in training and re-learning progressively cancelled the 'cue type' effect. Familiarity is also the factor that probably affected the physicians' opinion about the 3D temporal simulator. They showed quite some self-confidence while declaring that it would be easier to learn logic and deal with it on a piece of paper rather than grasping information from a "weird" 3D interface. Medical activities are usually standardized and perhaps implies a habit with familiar and trusted tools. Thus, abrupt changes of these tools are considered simply useless if not detrimental.

A similar observation can be done for the domain of ATC, which share some common characteristics with the medical domain. First of all, the knowledge required for both activities is procedural and quite standardized; then, both types of users seem rather conservative in judging the efficiency of the tools that they use; also they show distrust and diffidence towards novel and (or) 3D interfaces. These factors may explain why controllers' performance is more negatively affected by 3D interfaces when compared to subjects with no (or little) operational background (Wickens et al., 1994). The suspicion of controllers towards novel displays is notorious in the ATC community. If the amount of experience is large, learning of different kinds will have taken place and procedural knowledge has been built up as well as an increased performance due to perceptual learning.

But the utility of 3D interfaces for ATC has to be assessed taking into account other elements besides familiarity. The most important group of factors are perceptual. The results reported in section 4.2 and 7.3 clearly show that tasks requiring metric judgments on distances along the line of sight are distorted. Therefore, trying to assess the utility of 3D interfaces when this type of tasks is involved is useless.

This is valid for standard-monitoring tasks typical of en-route and approach sectors, where it is often required of controllers to quickly produce precise estimation of lateral distances. If there is a utility of 3D for ATC it can only be envisioned for tasks that do not require such precise estimations of distances.

What seems to lack in most current approaches is the investigation of ATC activities that do not suffer from such perceptual shortcomings. Thus, the work presented in Paper V proposes a number of activities that do not fall within the category of metric judgments.

By way of contrast, when spatial judgments are done on the same vertical axis, 3D could show some utility. That was the case for the subjects who were requested to identify traffic flying at the same level. In this case, the ability to extract information was quicker if the symbolic description of the flight levels was compounded with their 3D representation (cf. experiment in Paper V). In addition, these results seem to generalize, independently from the specific content of the scenes and from the background of the subjects involved in the study. However, once again, this type of task could be limited by another perceptual factor, namely the height constancy problem, which can strongly bias the judgment of vertical distances if they are done across widely separated vertical axes.

We are currently working to implement overview and zooming functions to access limited spatial spots, in order to resolve the height constancy problem.

Yet, it is necessary to remark that the task defined for the experiment was very basic; that it represents a small building block of a complex activity where more qualitative estimations are required. As a matter of fact, the traffic allocation activity requires controllers to "explore" the traffic. The activity requires organizing volumes of traffic rather than estimating exact distances.

Another interesting activity concerns the tower operations. The controllers interviewed claimed that tower control is done mainly in rules based on sight, rather than in 2D interfaces. They also stated that 2D offers a very simplified picture of the traffic situation. The 3D could offer a suitable "out-of-the-window" view of the traffic, and correct common problems created by poor visibility conditions (such as bad weather).

Finally, 3D could help trainee controllers to understand the semantics of air space. Air space is four-dimensional, including time, and it is characterized by meaning, in that the objects within it (the aircraft, but also flying procedures) have to comply with rules that follow spatial and temporal requirements. Nowadays the understanding of these spatial rules is acquired with the support of 2D top-views, thus they are indirectly extracted from an artificial and unnatural representations of the problem. The level of acceptance of this practice cannot be precisely estimated, but it is worth mentioning that a few years of training are required before controllers become fully operational. It was mentioned above, that ATC monitoring tasks are highly standardized and that controllers develop a sort of dependency of the 2D planar interfaces that they use. But how efficiently would they perform, if their training also would entail 3D interfaces (at least for tasks not requiring the perception of metric 3D structure)? The results reported in section 7.3 and the survey reported in Paper V suggest that such use is possible and perhaps useful. However, general empirical evidence regarding this issue is still missing.

13. Future work

To further enhance knowledge with regards to the issues at hand, more investigations need to be performed. For instance, it would be interesting to investigate more in depth the role of familiarity in the 3D temporal simulator. Moreover, it would be interesting to evaluate if the characteristics of the 3D simulator could meet the preferences of younger users, such as medical student.

In addition, concerning the retrieval and spatial memory tasks:

- Factors related to semantics could be investigated further. For instance, the position of the pictorial tokens could be manipulated to objectively assess if the spatial and semantic proximity has an impact on memory and if it also affected by the interface style.
- The previous studies used very abstract symbolic information and concrete pictorial information. The inclusion of concrete nouns in the comparison could assess the effect of concrete symbolic contents across 2D and 3D interfaces.
- Finally, the studies did not take into account information density, which can play a major role in memory tasks. In future this issue needs to be addressed.

With regards to the ATC domain,

- It would be interesting to carry out additional experiments that include dynamic and denser traffic scenes. The increased traffic could emphasize more strongly the height constancy problem reported in the literature.
- In Paper V is hypothesized that for the particular activity envisioned (the traffic allocation within sectors) the inherent problems could be solved by special functions like dynamic interaction and zooming. However, this has to be assessed. These functions require the implementation of special interaction tools. A partial report of these tools was given in Papers III and IV. The implementation of the interaction tools and dynamic objects is currently under development.

Acknowledgments

I wish to thank Werner Schneider who accepted me as a PhD student at DIS. Werner, thank you for giving me the chance to start this adventure, which has brought me to so many places, and finally back to Sweden.

Thanks to Mats Lind, who accepted to supervise this work and to follow the course of this thesis. Mats, thank you for your patience, for teaching me so much, for your support and for providing detailed comments, advice and constant assistance during the whole ‘thesis-time’.

Thanks to Vu Duong, who accepted me as a student in the Innovative Research Area at Eurocontrol. Thank you Vu for giving me this opportunity, for your support and for having the courage to think in a creative manner in the ATC domain.

I also want to thank Andreas Hamfelt and Anders Christoffersson for being understanding and supportive at several stages of my studies.

Thanks to Patrizia Marti who offered me the chance to start working in the ATC domain. If I had not received her call few years ago, the course of my life would have been much different. Special credits to many people working at Eurocontrol, especially to Alessia, Alessandro, Andrea, Angelo, Chiara, Cristiana, Deirdre, Dario, Giuseppe, Maria, Massimo, Paolo, and Willi. Thanks also to Horst Hering and Martina Jurgens for daily support and help at Eurocontrol. Thanks to all the Eurocontrol students who provided opinions and contributions during our Friday meetings, especially my team-mates Ha and Thong, but also to Claus, Fred, Magnus, Nabil, Peter, Sophie, Stephen, Tonia, etc. Also thanks to Alistair Jackson for providing suggestions and insightful comments on some of the thesis works.

Thanks to Ray Dowdall for his contribution to the ATC paper, for sharing his ATC knowledge, providing so many suggestions and proofreading some of my papers! Ray, thank you for being supportive and cheering me up. Thanks to Hugh O’Connor for ATC discussions and for the ‘rugby afternoons’ at the Coolin. Thanks to the terrific Geraldine Flynn for allowing me to discover how difficult and charming the domain of Air Traffic Control is.

I am also grateful for having the chance of spending some time at DIS. Thanks to Veronika Wagner and to Sabine Koch for sharing their knowledge on medical informatics with me. Thanks to all the nice people who work, have been working at or simply paid a visit to DIS, as all of them helped me in many ways, both for human and work related matters: Eva Enefjord, Gunnilla Klaar, Anne Håkansson, Andreas Kjellin, Camilla Forsell, Ingrid Lukell, Maria Hägglund, Thomas Snäckerström, Lars-Göran Svensk, Vladislav Valkovsky, Andy Soehnel, Janne and Giordana Eriksson, Pierre Hjälm, Anneli Edman, Stefan Seipel, Göran Svensson, Andrea Andreucci, Anders Ågren, Alexander Allard, etc.

A special thanks goes to Else Nygren for doing the first seminar on this thesis and for providing feedback and advice, and to Stefan Karlsson, who 'was the committee' during the seminar and proofread this thesis.

A super 'thank you!' to Isabella Scandurra and Milena Ivanova for their friendship and for hosting me in a warm and friendly home, during my trips to Uppsala.

My grateful thoughts also go to Rodger Romero Ramirez and to Meftun Ahmed, Eric Finney, and to Redar Anderson for letting me experience some Uppsala students' life. Thanks also to all the staff of "La Comedia", Mario, Patrizio, Franco, Luis, Mimmo, etc.

Grazie to my Italian friends, especially to Gianmaria, Luca, and Manola for being caring and supportive.

Thanks also to Lida and Antonio for their nice friendship, their support, and the lovely time we spent together in Sweden. Your presence in Uppsala really meant a lot to me. Also thanks to Adela and Zeynep, with whom I spent very funny moments in France, or in our old place in Flogsta.

Thanks also to my dear friends, Brahim and Timour. I was very lucky and I am very grateful for meeting you guys. Together, we shared so many moments, happy and sometimes sad. You were my family in Sweden. No matter where life will take me, you will always have a special place in my heart.

My greatest thanks to my family, my parents Lida and Mario, my grand parents Nerino and Rosa and my sister Maricla. Thank you for the love and your constant presence, no matter the distance. Your support gives me the strength to go on and find my way in life.

Bibliography

Andrews, K. (1995). Visualising cyberspace: information visualization in the Harmony Internet Browser. *Proceedings of the IEEE Symposium on Information Visualization*, GA, 97-104.

Atkinson, R. C. & Schriffin, R. M. (1968). Human memory: A proposed system and its control processes. In W. K. Spence and T. .J. Spence (Eds.) *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 89-125.

Banks R. & Wickens C.D. (1997). Commanders' display of terrain information: manipulations of display dimensionality and frame of reference to support battlefield visualization. *Technical Report ARL-97-12/ARMY-FED-LAB-97-2*, Savoy, IL: Aviation Res. Lab. http://www.aviation.uiuc.edu/UnitsHFD/TechPdf/97-12.pdf Accessed 2004-10-23

Barreau, D. & Nardi, A.B. (1995). Finding and reminding, File organization from the desktop. *SIGCHI Bulle*tin, 7, 3, July, 39-43.

Boyer B. S, & Wickens C. D. (1994). 3D Weather Displays for Aircraft Cockpits. *Technical Report ARL-94-11/NASA-94-4*, Savoy, IL: Aviation Res. Lab. http://www.aviation.uiuc.edu/institute/research/arl/technical-reports/3Dweather/3Dweather.html Accessed 2004-10-23

Bowman, D., Kruijff, E., LaViola, J. & Poupyrev, I. (2001). An Introduction to 3D User Interface Design. *Presence*, 10, 1, February, 96-108.

Bray, T. (1996). Measuring the Web. http://www5conf.inria.fr/fich_html/papers/P9/Overview.html Accessed 2004-10-23.

Burnett, M.S & Barfield, W. (1991). Perspective versus plan view air traffic control (ATC) displays: Survey and empirical results. *Proceedings of the Sixth International Symposium on Aviation Psychology*, Columbus, Ohio State University, 448-453.

Bruner, J. (1966). *Towards a theory of instruction*. Norton & Company, New York.

Card, S.K., Robertson, G.G. & Mackinlay, J.D. (1991). The Information Visualizer, an information workspace. *Proceedings of the SIGCHI Conference on Human factors in computing systems*, 181-188.

Card, S.K., Mackinlay, J.D., & Robertson, G.G. (1991). A morphological analysis of the design spaces of input devices. *ACM Transactions on Information Systems*, 9, 2, 99-122.

Card, S.K., Robertson, G.G. & York, W. (1996). The WebBook and the Web Forager: An Information Workspace for the World-Wide Web. *Proceedings of the SIGCHI Conference on Human factors in computing systems*, 111-117.

Carrol, J.M., Mack, R. L., & Kellogg, W. A. (1988). Interface metaphors and user interface design. In M. Helander (Eds.) *Handbook of Human Computer Interaction*, North-Holland, 75-76.

Cockburn, A. & McKenzie, B. (2001). 3D or not 3D? Evaluating the effect of third dimension in a document management system. *Proceedings of the SIGCHI conference on Conference on Human factors in computing systems*, 434-441.

Cockburn, A. & McKenzie, B. (2002). Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments. *Proceedings of the SIGCHI conference on Conference on Human factors in computing systems*, 203-210.

Costabile, M. F., Malerba, D., Hemmje, M. & Paradiso, A. (1998). Building Metaphors for Supporting User Interaction with Multimedia Databases - A Demonstration. *Proceedings of 4th IFIP 2.6 Working Conference on Visual DataBase Systems - VDB 4*, 47-65.

Czerwinski, M., Van Dantzich, Robertson, G. & Hoffman, H. (1999). The contribution of thumbnail image, mouse-over text and spatial location memory to web page retrieval in 3D. *Proceedings of Interact 1999*, 163-170.

Dang, N. T., Hong, H. L., & Tavanti, M. (2003). Visualization and interaction of flight trajectory in a 3D stereoscopic environment. *Proceedings of AIAA/IEEE 22nd Digital Avionics Systems Conference*, Indianapolis, Indiana.

De Kerckhove D. (1993). *Brainframes, Mente, tecnologia, mercato*. Baskerville, Bologna.

De Valois, R. L., & DeValois, K. K. (1975). Neural coding of color. In E.C Carterette & M.P. Friedman (Eds.) *Handbook of perception*, Academic Press, New York.

Dix, A., Finlay, J., Abowd G., & Beale R. (1998). *Human Computer Interaction*, Prentice Hall, London.

Eco, U. (1992). *La production des signes*. Librairie Générale Française, Paris.

Ellis, S.R., McGreevy, M. (1987). Perspective traffic display format and airline pilot traffic avoidance. *Human Factors*, 29(4), 371-382.

Ellis, S.R. & McGreevy, M., (1983). Influence of a perspective cockpit traffic display format on pilot avoidance maneuvers. *Proceedings of the Human Factors Society 27th Annual Meeting, Human Factors Society*, Santa Monica, CA, 726-766.

Erickson T., D. (1990). Working with interface metaphors. In Laurel B. (Eds.) *The art of Human-Computer Interface Design*, MA: Addison-Wesley, 65-73.

Gershon, N., Eick, S. G. & Card, S. (1998). Information visualization. *ACM Interactions*, March-April, 9-15.

Gapp, K.P. (1995). *Object Localization: Selection of Optimal Reference Objects*. citeseer.ist.psu.edu/gapp95object.html Accessed 2004-10-23.

Gleitman, H. (1995). *Psychology, 5th Edition*. Norton & Company, New York.

Gibson, E. J. (2000). *An ecological approach to perceptual learning and development*. Oxford University Press.

Gärdenfors, P. (1996). Mental representation, conceptual spaces and metaphors, *Synthese,* 106, 21-47.

Hart, R. A. & Moore, G. T. (1973). The development of spatial cognition: A Review. In Downs, R. M., and Stea, D., (Eds.) *Image and Environment: Cognitive Mapping and Spatial Behavior*, Aldine Publishing Company, Chicago, 246-295.

Haskell, I.D. & Wickens, C.,D. (1993). Two and three-dimensional displays or aviation: a theoretical and empirical comparison. *International Journal of Aviation psychology*, 3(2), 87-109.

Hewitt, T.T., Baecker, R., Card, S., Gasen, J., Mantei, M., Perlman, G., Strong, G., & Verplank, W. (2004). *ACM SIGCHI curricula for human computer interaction*, http://sigchi.org/cdg/cdg2.html Accessed 2004-10-23.

Hirtle S.C. & Jonides, J. (1985). Evidence of hierarchies in cognitive maps, *Memory and Cognition*, 13, 208-217.

Holtzblatt, K. & Jones, S. (1991). Contextual Inquiry: A Participatory Technique for System Design. In D. Schuler & A. Namioka (Eds.) *Participatory design: Principles and Practices*, Erlbaum, 177-210.

Hunt E. & Waller, D. (1999). Orientation and wayfinding: a review. *ONR technical report N000 14-96-0380*, Arlington, VA: Office of naval Research, 1999.

Intersense. http://www.isense.com/ Accessed 2004-10-23.

International Organization for Standardization (1998). *Ergonomic Requirements for Office Work with Visual Display Terminals (VTDs) Part 11: Guidance on Usability.* ISO/IS 9241, Geneva.

Johnson-Laird, P. N. (1993). *La mente e il computer*. Introduzione alla scienza cognitive, Bologna, Il Mulino, p. 236.

Jones, W.P. & Dumais, S.T. (1986). The spatial metaphor for use interfaces: experimental tests of references by location versus name. *ACM Transactions on Office Information Systems*, 4, 42-63.

Kanizsa, G., Grammatica del vedere. Saggi su percezione e gestalt, Feltrinelli, 1980.

Kay, A. (1990) User interface: a personal view. In B. Laurel (Eds.) *The art of Human-Computer Interface Design*, MA: Addison-Wesley, 194-196.

Kelley, C. M. & Jacoby, L. L. (2000). Recollection and familiarity process-dissociation. In E. Tulving & F. I.M. Craik (Eds.) *The Oxford Handbook of Memory*, Oxford University Press, 217.

Lakoff, G. & Johnson M. (1980). *Metaphors we live by*. University of Chicago Press, 1980.

Larkin, J. H., Simon, H., A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-100.

Laurel, B., *The art of Human-Computer Interface Design*. In B. Laurel (Eds.) MA: Addison-Wesley, 1990.

Lind, M., Bingham, G., & Forsell, C. (2003) Metric 3D structure in Visualizations. *Information Visualization*, 2, 51-57

Meyers, M. D. (1999). Investigating information systems with ethnographic research. In *Communications of AIS*, 2, 23, 3-4.

Michard, A. (1982). Graphical presentation of Boolean expressions in a database query language: design notes and an ergonomic evaluation. *Behavior and Information Technology*, 1, 3, 279-288.

Miller, N. E., Hetzler, B., Nakamura, G. & Whitney, P. (1997). The need for metrics in visual information analysis. *Proceedings of the Workshop on new paradigms in information visualization and manipulation*, Las Vegas, Nevada, 24-28.

Miller, N. E., Wong, P. C., Brewster, M. & Foote H. (1998). Topic islands- a wavelet-based text visualization system. *Proceedings of the IEEE conference on Visualization '98*, 189-196.

Molendi, G. & Patriarca, M. (1992). Virtual Reality: medical researches. *Technical Report December 1/92*, Universita' degli Studi di Milano, Dipartimento di Scienze della Informazione.

Nelson T. H. (1990). The right way to think about software design. In B. Laurel (Eds.) *The art of Human-Computer Interface Design*, MA: Addison-Wesley, 237.

Norman, D. (1990). *La caffettiera del masochista*. Giunti Editore, Firenze.

Norman, D. (1993). *Le cose che ci fanno intelligenti*. Feltrinelli, 1993.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 4, 32-38.

Perrone, J.A. (1982). Visual slant underestimation: A general model. *Perception*, 11, 641-654.

Polson, P.G., Lewis, C., Rieman, J. & Wharton, C. (1992). Cognitive walkthroughs: A method for theory- based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36, 741-773.

Poupyrev, I., Weghorst, S., Billinghurst, M. & Ichikawa, T. (1998). Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. Computer Graphics Forum, 17, 3, 41-52.

Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D. & van Dantzich, M. (1998). Data Mountain: using spatial memory for document management. *Proceedings of UIST'98*, ACM, CA, 153-162.

Robertson, G., Van Dantzich, M., Robbins, D., Czerwinski, M., Hinckley, K., Risden, K., Thiel, D. & Gorokhovsky, V. (2000). The Task Gallery: a 3D Window Manager, *Proceedings of the SIGCHI conference on Human factors in computing systems*, 494 – 501.

Roediger, M. L. & McDermott, K.B. (2000). Distortions on memory. In E. Tulving & F. I.M. Craik (Eds.) *The Oxford Handbook of Memory*, Oxford University Press, 149-162.

Satava, R. M. (2001). Surgery a technological framework for the future. *Proceedings of Virtual Reality International '93*, UK, 101-103.

Saussure de, F. (1959). *A Course in General Linguistics*. Philosophical Library, New York.

Schwamb, K. B. (1990). *Mental models: A survey.* citeseer.nj.nec.com/schwamb90mental.html Accessed 2004-10-23

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualization. *Proceedings of IEEE Symposium on Visual Languages '96*, CA, 336-343.

Sparacino, F., Davenport, G., & Pentland, A., (2000). *City of News*. http://xenia.media.mit.edu/~flavia/Papers/CityOfNewsKOS.pdf Accessed 2004-10-23

Stenning, K., & Oberlander, J. (1997). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*, 19, 97-140.

Stone, R. J. (1999). The opportunities of virtual reality and simulation in the training and assessment of technical surgical skills. *Proceedings of Surgical Competence: Challenges of Assessment in Training and Practice*, Royal College of Surgeons and Smith and Nephew Conference, London.

Stone R. J. (1991). Advanced human system interfaces for tele-robotics using virtual reality and telepresence technologies. *Proceedings of the 5th Int. Conference on Advanced Robotics*, Pisa, Italy.

Tham, M. & Wickens, C.D. (1993). Evaluation of perspective and stereoscopic displays as alternative to plan view displays in air traffic control. *Technical Report ARL-93-4/FAA-93-1,* Savoy, IL: Aviation Res. Lab.

Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 77, 593-60.

Tversky, B. (2000). Remembering spaces. In E. Tulving & F. I.M. Craik (Eds.) *The Oxford Handbook of Memory*, Oxford University Press, 363-372.

Vygotsky L. (1930). *Mind and Society.* http://www.marxists.org/archive/vygotsky/works/mind/ch04.htm Accessed 2004-10-23

Ware, C. (1990). *Information visualization: perception for design*. Morgan Kaufmann Publishers.

Wickens, C.D., Campbell, M., Liang, C.-C. & Merwin, D.H. (1995). Weather displays for air traffic control: the effect of 3D perspective. *Technical Report, ARL-95-1/FAA-95-1,* Savoy, IL: Aviation Res. Lab. http://www.aviation.uiuc.edu/institute/research/arl/technical-reports/weathdisp/weather2.pdf Accessed 2004-10-23

Wickens, C.D. Miller, S. & Tham, M. (1994). The implications of data-link for representing pilot request information on 2D and 3D air traffic control displays. *Technical Report, ARL-94-2/FAA-94-1*, Savoy, IL: Aviation Res. Lab.

Wickens, C.D., & May, P. (1994). Terrain Representation for Air Traffic Control: A Comparison of Perspective With Plan View Displays. *Technical Report, ARL-94-10/FAA-94-2*, Savoy, IL: Aviation Res. Lab. http://www.aviation.uiuc.edu/institute/research/arl/technical-reports/terrain/terrain.html Accessed 2004-10-23

Winn, W. (1995). Learning in Interactive and Immersive Environments. *Proceedings of Media and Higher Education*. National Institute for Multimedia Education, Chiba, Japan.

Wixon, D. & Holzblatt,. K. (1990). Contextual design: an emergent view of system design. *Proceedings of SIGCHI Conference on Human factors in computing Systems*, 329 – 336.

Wozniak R. H. (1999). *Classics in Psychology, 1855–1914: Historical Essays*. http://www.thoemmes.com/psych/fechner.htm Accessed 2004-10-23

Zhang, J. (1997). The nature of external representation in problem solving. *Cognitive Science*, 21(2), 179-217.

Acta Universitatis Upsaliensis

Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Social Sciences
Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is either a monograph or, as in this case, a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences*. (Prior to July 1985, the series was published under the title "Abstracts of Uppsala Dissertations from the Faculty of Social Sciences".)

Distributor:
Uppsala University Library,
Box 510, SE-751 20 Uppsala, Sweden
www.uu.se, acta@ub.uu.se

ISSN 0282-7492
ISBN 91-554-6102-X