Gaze Interaction With Zoomable Interfaces

A comparative study of gaze and mouse navigation

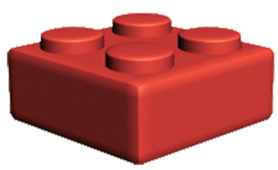
Emilie Møllenbach and Thorarinn Stefansson

IT University of Copenhagen

Master Thesis

Supervisor: John Paulin Hansen

September 1, 2006



Gaze Interaction with Zoomable Interfaces

A Comparative Study of Gaze and Mouse Navigation

Emilie Møllenbach and Thorarinn Stefansson

IT University of Copenhagen

Master Thesis

Supervisor: John Paulin Hansen

September 1, 2006

Abstract

This thesis examines gaze-controlled navigation in a pan and zoom interface. Pan/zoom navigation was chosen on the assumption that it in many ways resembles the way eyes are used in everyday orientation.

In the main experiment described in this paper, a test environment was constructed with two information spaces; one large with 2000 nodes ordered in semi-structured groups, the other was smaller and designed for precision zooming. Four navigation techniques were tested; two using a mouse and two with gaze-control. Gaze was implemented as a pointing device by employing an eye-tracker and combined with keyboard controls for zooming. Two existing mouse implementations were chosen for comparison.

Testing the performance of gaze- and mouse-control in broad search tasks with pan/zoom navigation, but without selection, shows that the performances of the most efficient mouse and gaze implementations were indistinguishable. Furthermore, in tasks involving zooming in on a limited number of known locations, the most efficient mouse-control proved to be about 16% slower than the gaze-control.

The results indicate that gaze-controlled pan/zoom navigation is a viable alternative to using a mouse in tasks concerning visual inspection and target exploration in large static multiscaled environments; supplementing mouse control with gaze navigation holds interesting potential for interface and interaction design.

Preface

This is a master thesis written by Emilie Møllenbach and Thorarinn Stefansson at the IT University of Copenhagen in September 2006 and supervised by John Paulin Hansen.

We would like to acknowledge and thank Aza Raskin for his insightful observations and for allowing us to use his source code to the zoomable interface demo. We also thank Haakon Lund from the Royal School of Library and Information Science for his extensive help and granting us full access to eye tracking equipment and testing facilities.

Finally we would like to thank John Paulin Hansen for his inspirational and dedicated supervision of this thesis.

Supplementary file

A supplementary file to this thesis can be found at *http://thorarinn.com/files/thesis/exp2.html*. It opens in an ordinary web browser and requires a Flash Player, version 8 or higher.

The file *exp2.html* is a slightly simplified version of the test environment used in our second experiment. More on-screen information has been added, and tasks no longer require the activation of a moderator. The tasks presented are generalizations of the task types used in our test, whereas the original version only presented one specific task at a time. Otherwise all navigation techniques behave as in our original test.

Please note that this is a large file, over 6MB, so it takes a while to load.

Also note that the gaze versions can be explored using a normal mouse (although the cursor is initially hidden, a representation can be activated pressing the “C” key on the keyboard). Viewing the file in full-screen mode is recommended, in most browsers this is done by pressing F11.

Authors

The authors can be contacted via email:

emilie AT itu DOT dk / thorarinn AT thorarinn DOT com

1 Introduction **4**

1.1 Problem statement 5
1.2 Structure of this thesis 5

2 Related Works **7**

2.1 General background 7
2.2 Large information spaces: Visualization and navigation 9
2.3 Alternative input devices 22
2.4 Naturalness versus familiarity 36
2.5 Our design decisions 38

3. Empirical studies and test-design **40**

3.1 Experiment 1 40
3.2 The context of the second experiment 49
3.3 Test design for the second experiment 53

4 Empirical results **66**

4.1 General results 66
4.2 Comparisons 70
4.3 Qualitative evaluations of the navigation techniques 74
4.4 Summary of results 76
4.5 Discussion of results 77

5 Conclusions **82**

6 Future research **86**

Literature **90**

Appendixes **96**

1 Introduction

Following the unprecedented computing capability available to scientists and knowledge workers the amount of data being worked with has grown at an explosive rate.[1] And while the improvements in CPU performance and bandwidth capabilities have mostly kept up with this evolution, we humans have had trouble keeping up using the current interaction methods:

> "Nowadays, the main bandwidth bottleneck in an interactive computer system occurs in the link between computer and human, not between computer components within the system." [2]

In the last decades the evolution of computer displays has greatly increased the amount of information displayable, both with increased colour-depth and increased resolution. While the first PC displays were capable of displaying 128,000 pixels in monochrome around 1981, today the most common resolutions are around 1 million pixels and high-end monitors are exceeding 2 million pixels, capable of displaying several millions of colours. Despite this evolution the amount of data to be processed grows much faster than the capabilities of the displays.

Added to this is the growing trend towards mobility with miniature displays and the need for alternatives to traditional input devices. Therefore the research of information visualization and of navigation in large information spaces, including the use of alternative inputs, is currently more relevant than ever before.

These issues have been substantially explored within the field of Human Computer Interaction (HCI). In this thesis we attempt to establish a historical and theoretical framework by presenting a range of previously conducted experiments along with the observations and contemplations of visionaries who have worked in the field. This framework is an essential part of defining the boundaries of our own experiments.

We explore some aspects of using gaze-interaction for navigating in large nodal information spaces and describe two empirical studies which we have conducted. In our first round of experiments we attempted to clarify some of the general challenges of a pan/zoom navigation technique by developing and testing a mouse-controlled environment. In a second experiment we tested our

---

[1] Ma, Kwan-Liu (2000)
[2] Funkhouser, T. et al. (2000) p. 20

implementation of a gaze-controlled navigation technique in a semi-structured visual information space, comparing its performance to traditional mouse-controlled techniques. This latter experiment and its results form the backbone of this thesis.

The main focus of this thesis is the relationship between zoomable interfaces and gaze; specifically in regard to navigation. The notion behind the types of tasks we research can be exemplified by the act of searching in a library for a specific book. Our main interest lies in finding the right shelf and book, stopping short of actually picking the book off the shelf.

1.1 Problem statement

Our main problem statement for this paper is:

> **How can gaze-interaction be implemented as an efficient navigation technique in visual information spaces?**

This key question involves many different topics which have to be explored if we are to gain an understanding of the issues involved. This includes questions such as:

- What elements from our natural visual perception can be applied in gaze-based interaction?
- What type of information suits a gaze-controlled navigation technique?
- How can we measure and compare gaze-controlled navigation techniques, with classic mouse-controlled techniques?
- Is gaze-interaction a future substitute for the mouse?
- Is there a way where gaze-interaction can be implemented using low resolution cameras and be effective as a supplement to existing input devices?

Aspects of this problem statement are underlying in the presentation of previous works and in the discussion of our results.

1.2 Structure of this thesis

Our theoretical framework is primarily based on two fields of interest;(1) the challenges of visualizing and navigating in large information spaces, and (2) the area of alternative input devices, especially with regard to eye-tracking and gaze interaction. This framework is presented and discussed in Chapter 2, concluding with our remarks on how observations from previous works relate to the design decisions made in our experiments.

The process leading up to the results presented in this thesis can be described as being explorative and iterative in nature. Starting from a general problem statement our main focus has been on designing and conducting empirical research, including the learning process of developing a conceptual prototype.

This part of the process has been conducted in two rounds of experiments, described in Chapter 3:

(1) In the initial test design phase we explored different theories, empirical studies and practical implementations in order to gain a fundamental understanding of the research field. Observations done on information visualization, navigation strategies and the technical possibilities for designing a usable test scenario, all originating in this phase, have been instrumental in reaching our results.

(2) The experience gained from designing and implementing our initial test, helped establish a clear direction for the second phase. Here we turned the focus to gaze interaction and specifically zoomable interfaces, comparing the gaze navigation with existing mouse techniques.

By creating original content and tasks we attempted to eliminate as many variables as possible and isolate the navigational aspect in both gaze and mouse interaction.

The results from the second phase are presented in detail in Chapter 4, as this experiment is the source of our main results. We compare different navigation techniques, e.g. a mouse controlled version and a corresponding gaze-controlled, and discuss the major findings.

In Chapter 5 we draw our conclusions on gaze-controlled pan/zoom navigation of large information spaces, in light of our empirical results.

Finally we discuss possible future studies extending our works and on gaze-controlled interaction in general in Chapter 6, where we propose several interesting research areas and present potential future scenarios of gaze interaction.

2 Related Works

In the related works section of this paper we will present and discuss the most important concepts in connection to our area of research, establishing a taxonomy which will be used throughout the paper. An overview of noteworthy previous works in this area will be presented by discussing theories and articles of special interest. We will also introduce relevant connections between theory and our experiments.

2.1 General background

The overall field in which we are working is Human Computer Interaction (HCI), an interdisciplinary subject with computer science at its core and branching out into other fields of research.

The interaction between man and machine is conducted through an *interface*, which translates input to output. Currently the most common instances of human computer interaction involve the *graphical user interfaces* (GUI) known from most personal computers, with keyboard and mouse as input devices and the output from the computer displayed on a monitor. But an interface is by no means limited to the manipulation of a graphical user interface, as clearly expressed in Jef Raskin's definition: "The way that you accomplish tasks with a product – what you do and how it responds – that's the interface" [3]

While computers essentially break data into binary form, the human mind favours combining elements into groups and performing visual analysis. To facilitate working with large amounts of data some kind of visualization is required:

> "Visualization transforms large quantities of raw data into graphical representations that exploit the superior visual processing capability of the human brain to detect patterns and draw inferences." [4]

Information visualization deals with this process of transforming data into an appropriate visual representation. These visualizations can range from a simple table or graph, to representations of large-scale data structures. In the context of this paper information visualization is typically seen as dealing with large-scale nodal hierarchies, such as the index system of a library. Clearly the visual

---

[3] Raskin J., The Humane Interface (2000) p. 2
[4] Ma, Kwan-Liu (2000) p. 16

representation chosen is dependent on the type of data being visualized and existing relationships within it. What we refer to as an *information space* is the combination of data and its underlying relationships, for example a hierarchically structured dataset. Any given information space can then be visualized in a number of different ways, dependent on the intended interaction with the data.

Any interaction with an information space involves some form of movements within the structure and potentially the selection of some subset of data for further manipulation. The interaction taking place up until a possible selection is what we define as a *navigation technique* (NT). As an example, scrolling within a word processor document using the vertical scroll-bar is one navigation technique, another example could be in a strategy games (e.g. Age of Empires) where the player can zoom out and gain perspective on the entire known map or zoom in to observe placement of specific units.

The above mentioned concepts are closely intertwined; the information visualization is dependent on the data, appropriate navigation techniques are dependent on the information visualization chosen, and the nature of the data dictates what forms of navigation need to be available.

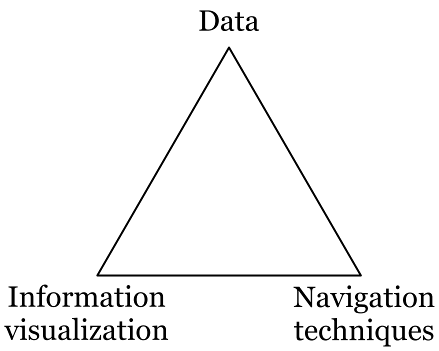


Figure 1: The relationships between data, visualization and navigation. Changes in one potentially affect the whole triangle.

The concept of an information space is in fact implicit in the triangle shown in Figure 1. In order to come up with any sort of useful visualization of a particular dataset, a study of the data and the inherent relationships within that data (the information space) is necessary. And since visualization is based on the structured presentation of data, so are the navigation techniques which apply. The idea of an information space, including both the data and underlying relationships within it, is in fact central to our triangle.

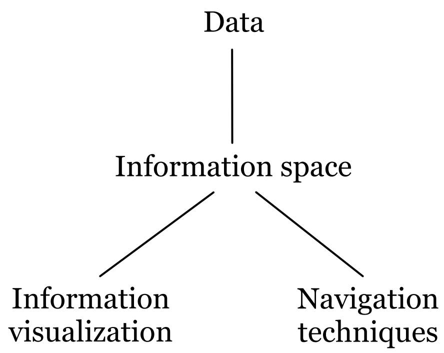


Figure 2: The concept of an information space can be seen as an intermediary between data, visualization and navigation.

In essence: We see information visualization and navigation techniques as being so closely related that they are almost inseparable, any design decisions regarding information visualization therefore have to include a consideration of the corresponding navigation techniques, and vice versa. The choice of information visualization, and implicitly navigation techniques, is of course dependent on the nature of the data to be visualized and the tasks to be solved.

2.2 Large information spaces: Visualization and navigation

> "Big information worlds cause big problems for interfaces. There is too much to see. They are hard to navigate." [5]

The larger the information spaces to be visualized and navigated, the more problems arise. Attempts of trying to display all details at once quickly exceed the capabilities of any available display technique, and overly focusing on detail presentation may lead to sacrificing the option of an overview of the whole dataset, which is necessary for detecting patterns and relationships within it.

This type of information spaces are sometimes referred to as "larger than screen datasets", but we prefer the concept presented by Furnas and Bederson (and later taken up by Baudisch et al.[6]) of *multiscale interfaces*. The expression multiscale refers to datasets where the order of magnitude between an overview and a useful detail view is too large for simultaneously representing both in a single view. As an example, a doctor examining an X-ray needs to study close-ups of vertebrae pairs

[5] Furnas, G. W. et al. (1995) page number missing
[6] Baudisch, P. et al. (2002)

as well as the full spine X-ray, and a visualization of a power network needs to present both details about the status of each node and an overview of the whole network.

Examples of multiscale data are abundant, and so are the attempts at creating interfaces for coping with the challenges of these types of data, each with its strengths and weaknesses.

Distortion interfaces versus zoomable interfaces

The most common techniques for dealing with multiscale data can be grouped into two categories; *distortion techniques* and *zoom techniques*.

The idea behind distortion techniques is to simultaneously display the area of interest (the focus) and the context (global view) in which it exists as an attempt at guaranteeing the visibility of all elements of interest, as Nekrasovski et al. put it: "Ensuring that regions of interest remain visible independent of navigation actions, possibly as compressed landmarks rather than being shown in full detail, is termed guaranteed visibility." [7]

The distortion techniques integrate the focus and context regions into a single view, using distortion and non-linear magnification [8], and are therefore commonly referred to as *focus+context techniques (F+C)*. But when comparing F+C interfaces to zoomable interfaces it is the distortion element that most clearly differentiates them and therefore we choose this as the basis for our categorization.

On the other hand zoom techniques give the appearance of scaling the whole information space, with elements outside the area of interest disappearing out of view when zooming in. The benefits and drawbacks of zoom techniques will be further discussed shortly.

A picture says more than a thousand words, and the following images display some well known distortion implementations of a focus+context interface.

---

[7] Nekrasovski, D. et al. (2006) p. 2
[8] Nekrasovski, D. et al. (2006)

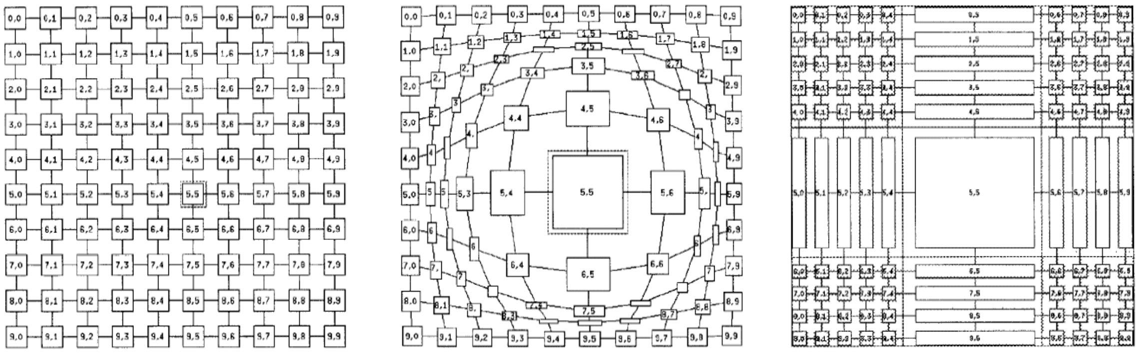


Figure 3: A symmetric layout stretched by two rubber sheet implementations; spherical and linear.[9]

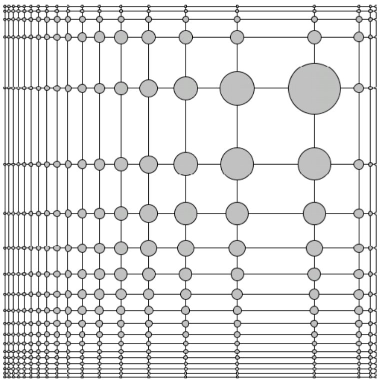


Figure 4: A schematic presentation of a fisheye graph.[10]

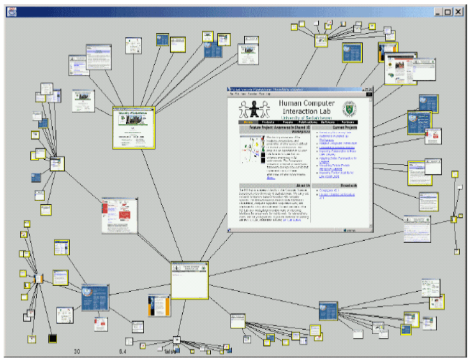


Figure 5: An interactive fisheye view for browsing a web site.[11]

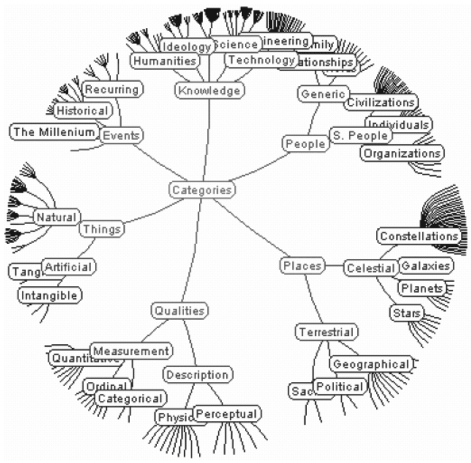


Figure 6: A hyperbolic representation of a tree structure.[12]

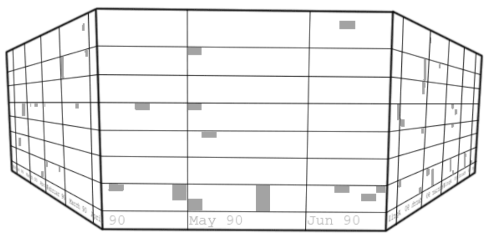


Figure 7: A perspective wall, here displaying a calendar application.[13]

---

[9] Sarkar, M. et al. (1993)
[10] Heer, J. et al. (2005)
[11] Gutwin, C. (2002)
[12] Heer, J. et al. (2005)
[13] Adapted from: Mackinlay, J. D. et al. (1991)

As evident from these pictures, distortion techniques display a non-distorted point of focus and then visually compress the surrounding context. The compression chosen depends on the information space being presented, but most common are implementations of fisheye, rubber sheet and other hyperbolic techniques.

One of the strengths of these information visualization techniques is that the context and the focus are simultaneously displayed, rendering the need for a separate overview redundant.

The downside of distortion visualizations is that they sacrifice the proportions between information elements which are important in many tasks. The quote below can be applied to distortion techniques in general.

> "Fisheye views give detail and context in a single view but, the fisheye browser severely distorts the image and requires constant reorientation. The distortion is a severe problem for applications where size and geometry are important." [14]

Zoomable interfaces

Multiscaled information can also be explored by a zoom navigation technique (NT). The most commonly implemented version of this is the *geometric zoom*, which maintains the visual proportions between the information objects by scaling the entire information space linearly.

Geometric zoomable interface is only an approximate definition of the concept of zoomable interfaces in general. Different approaches in information visualization techniques and additions in navigation techniques, such as pan, spawn other expressions: *semantic zooming* and *pan/zoom navigation* (PZN), these will subsequently be elaborated upon. These various available options make a precise definition difficult, as Hornbæk et al. point out:

> "While zoomable user interfaces have been discussed since at least 1993 [...], no definition of zoomable user interface has been generally agreed upon." [15]

The zoomable interface has three strengths that make it a viable alternative for viewing large nodal information spaces:

---

[14] Plaisant, C. et al. (1995) page number missing
[15] Hornbæk, K. et al. (2002) p. 4

(1) The first is *perspective*. Jef Raskin, author of the Humane Interface, had a vision for a zoomable interface called ZoomWorld. The idea was to rid interfaces of intricate intertwining menu systems, which he compares to mazes: “As legends and stories from ancient times inform us, humans have always been notoriously bad at mazes.” [16]

The zoomable interface would allow a dynamic view of the system; permitting the user to access any level of detail directly from a global view, without the need for entering sub menus.

(2) The second strength is *geometric integrity* which addresses the weakness of distortion interfaces. The problem with distortion interfaces, as mentioned, is that size, shape and the mutual relationship between elements of information are distorted.

In zoomable interfaces proportions within the information space are upheld. There are several real life examples where this geometric integrity is of importance; for instance a doctor viewing an x-ray. The image of a broken bone makes little sense in a distorted view where the fraction is out of proportion with the remaining leg.

(3) The third strength is *familiarity*. Some of the earliest developers and advocates of zoomable interfaces, among others Perlin and Fox [17], claimed one of its strengths to be “tapping into people’s natural spatial abilities”.[18] The argument is that the way we navigate the world is by gaining perspective and focussing in on targets of interest, which makes zoomable interfaces a familiar mapping of our natural navigational interaction with the world.

> ”In natural scenes, new objects rarely appear abruptly; they are more likely to appear by progressive disocclusion from behind other surfaces (J. J. Gibson et al., 1969).” [19]

The strength of perspective and geometric integrity also potentially constitute the greatest weakness of zoomable interfaces. If we assume that most tasks are actually solved in the detail view rather than in the perspective view. The fact that the perspective is not maintained could lead to users loosing a sense of context and continuously having to zoom in and out to reorient.

There have been two main approaches suggested in solving this potential problem. (1) The most explicit and well known approach is creating an *overview* in a separate field. This way the overview

---

[16] Raskin J., The Humane Interface (2000)
[17] Creators of Pad and Pad++: early implementations of a toolkit for making zoomable interfaces
[18] Bederson, B. B. et al. (1996)
[19] Franconeri, S. L. et al (2003)

(global view) is presented simultaneously with the detailed view. Overviews have in regard to other information visualization techniques proved to be a valuable asset in navigating large nodal information spaces. [20]

The overview (global view) "gives a view of the entire universe that could be explored. One purpose of the global view is to give a sense of what information will be in the image - and what is not". [21]

The fact that users have a constant reminder of their location has also been said to "improve subjective satisfaction [...], and efficiency [...]"[22]. We will subsequently explore whether this is the case for zoomable interfaces.

(2) A second attempt to allow a greater contextual understanding of the information space in zoomable interfaces is *semantic zooming*; here information has different representations depending on scale level. Instead of seeing an object scaled down, when zooming away from the information space, it is substituted for a different representation. Thereby the visualization allows the content to dynamically change according to scale level.

> "For example, we can use semantic zooming to change the way things look depending on their size. As we mentioned, zooming provides a natural mechanism for representing abstractions of objects. It is natural to see extra details of an object when zoomed in and viewing it up close. When zoomed out, instead of simply seeing a scaled down version of the object, it is potentially more effective to see a different representation of it." [23]

A common example of semantic zooming is employed in online geographical maps. In a Danish implementation of this, called Krak [24], 8 predefined levels of zoom are available. At the highest level all of Denmark is shown, with only labels for the 5 biggest cities. The next scale levels reveal more cities and natural landmarks such as lakes, at the fourth scale level train stations and parks become visible; at the fifth level street names of larger roads can be viewed and so on. A visualization of this can be seen from Figure 8a through Figure 8d. The red circles represent the area being zoomed in on.

---

[20] Baudisch, P. et al. (2002) / Frøkjær, E. et al. (2001) / Hornbæk, K. et al. (2002)
[21] Plaisant, C. et al. (1995) page number misssing
[22] Hornbæk, K. et al. (2002) p. 2
[23] Bederson, B. B. et al. (1996) p. 8
[24] www. krak. dk

Figure 8a: Semantic zooming: The Krak map of Denmark at scale level 1. Only major cities shown.

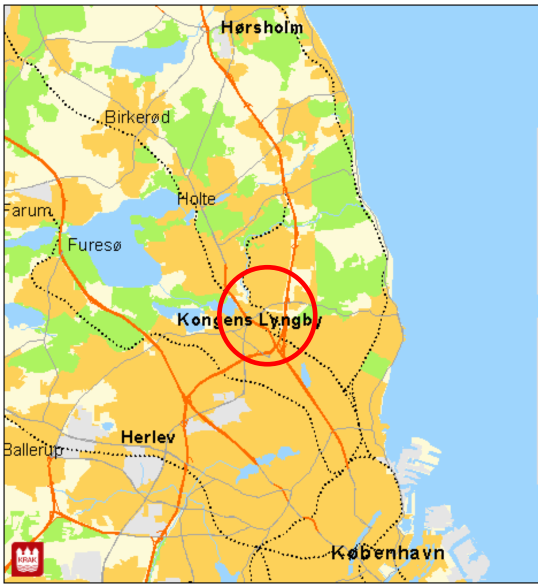


Figure 8b: Semantic zooming: The Krak map of Denmark at scale level 3. Smaller cities and natural landmarks.

Figure 8c: Semantic zooming: The Krak map of Denmark at scale level 4. Quarters within cities and train stations.

Figure 8d: Semantic zooming: The Krak map of Denmark at scale level 7. Detailed road names and landmarks.

This allows for a perceived closer proximity between the detail level and overview, as the users don't necessarily need to enter the detail level of an information object in order to gain comprehension of its content.

Zooming, as mentioned, enables the user to move from one level of detail to another. When coupled with pan another potential weakness is avoided. Zoom without pan doesn't enable the user to explore the immediate context of an equally scaled detailed area without leaving the particular scale level, by zooming out and in again.

Pan can be defined as planar movements, allowing navigation both horizontally and vertically in the information space. Panning is a well known form of navigation in information spaces, in regard to its most common implementations Plaisant et al. observed:

> "We observed four different panning implementations. The most common is vertical and horizontal scroll bars. Another way is with a "sticky hand" which grabs the picture when the mouse button is pressed (first used in MacPaint). The picture then follows the cursor until the mouse button is released. The "sticky hand" metaphor is only appropriate when a real time update of the image is possible. Arrow keys are another method used for panning. Finally when an overview is present, panning of the detail view can be accomplished by moving the field-of-view indicator in the overview." [25]

(For future reference, what is called "sticky-hand" in the quote, we will refer to as a grab navigation technique.)

A final element which needs to be introduced is the behaviour of zoom. There are two distinct manners in which zoom can take place. The first is *stepwise* (jump zooming) and the second *continuous* (animated):

> "In jump zooming, the change in scale occurs instantly, without a smooth transition. Jump zooming is used in Pad [], Schaffer et al.'s [] experimental system, and commercial systems such as Adobe PhotoShop or MapQuest. In animated zooming the transition from the old to the new scale is smooth []." [26]

Combining the planar movements of pan with the depth effect of zooming gives a 3D-like navigation; even in 2D information spaces. This allows for a flexible interaction and presents a form of navigation which has a strong impact on how information can be presented; employing the potential of depth in a two-dimensional representation of information spaces.

Although pan/zoom navigation emulates familiar elements from 3D navigation, it does differ from a "true" 3D motion. This is most easily demonstrated by a camera analogy:

---

[25] Plaisant, C. et al. (1995) page number missing
[26] Hornbæk, K. et al. (2002) p. 3

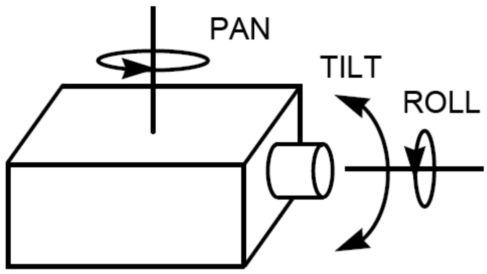


Figure 9a: In addition to moving, a camera is typically described as being capable of pan, tilt, roll and zoom (not displayed).[27]

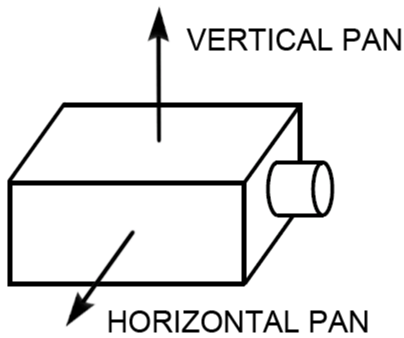


Figure 9b: In pan/zoom navigation the pan movements emulate moving a camera in a plane perpendicular to the lens.

Pan/zoom navigation is comparable to moving a camera in a plane and from there zooming in on details. Zooming in on an object emulates the familiar behaviour of an increasing amount of detail becoming visible as we approach a real life object.

But without allowing the "camera" to be moved towards objects and tilting it, this kind of movements can not be described as fully emulating 3D movements; for instance when it comes to viewing objects located behind other objects. Also note from Figure 9 that the definition of pan used in this paper, and in 2D interface design in general, is somewhat different from the definition of pan used in cinematography.

All the zoomable interfaces explored in this thesis are implementations of pan and zoom movements within 2D spaces.

Relevant studies on Zoomable interfaces

As previously mentioned the main problem with zoomable interfaces is that they potentially leave the user lost in the information space. This lack of context awareness in detail views may cause inefficient navigation.

As a consequence, this dilemma has been the subject of several studies regarding zoomable interfaces in recent time.

Support of added overviews has been described in several studies regarding information visualization. Evidence has shown that in many cases the addition of overviews makes for more

[27] Adapted from Fjeld, M. et al (1999)

efficient navigation patterns in large nodal information spaces. Hornbæk et al (2002) explain three reasons why this is the case:

> "First, navigation is more efficient because users may navigate using the overview window rather than using the detail window []. Second, the overview window aids users in keeping track of their current position in the information space []. Third, the overview gives users a feeling of control []." [28]

For zoomable interfaces the results are somewhat different. Nekrasovski et al. put overviews and zoomable interfaces to the test in a study comparing a rubber sheet navigation (RSN) and pan/zoom navigation (PZN) with and without an overview. Two of the initial three hypotheses of this experiment, and their subsequent results, are of relevance. The most important was: 1) "For PZN, the presence of an overview results in better performance." [29] The second hypothesis of interest was: 2) "RSN interfaces perform better than PZN interfaces independently of the presence or absence of an overview." [30] But the experiment turned out to reveal different results.

There were 4 different navigation techniques: 1) RSN without overview, 2) RSN with overview, 3) PZN without overview and 4) PZN with overview. The test consisted of search tasks in large data trees, inspired by evolutionary biologists, whose work "relies heavily on visual inspection and topological analysis of large trees" [31]. The dataset used in the test was a "binary tree" consisting of 5,918 nodes, some of which were colour coded; a small extraction of the tree is seen in Figure 10. The 24 subjects, who participated in the study, were asked to solve tasks of searching for the colour coded nodes within the data tree. The context and nature of the information space and the tasks within this test influenced the design of our second experiment.

---

[28] Hornbæk, K. et al. (2002) s. 364
[29] Nekrasovski, D. et al. (2006) p. 3
[30] Nekrasovski, D. et al. (2006) p. 3
[31] Nekrasovski, D. et al. (2006) p. 2

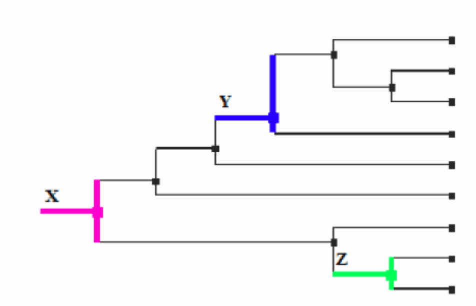


Figure 10: A small extraction of the content used for a study by Nekrasovski et al. on rubber sheet and pan/zoom navigation with and without an overview. Tasks involved analysing the distances between nodes in a hierarchal tree, as in counting the nodes from X to Y.

The results showed both of the earlier stated hypotheses to be incorrect:

The statement regarding hypothesis 1: "for PZN, having an overview made no significant difference in terms of completion times, navigation actions, or resets. Having an overview was, however, reported to reduce physical demand." [32]

The statement regarding hypothesis 2: "PZN interfaces performed better than RSN interfaces in terms of completion times, navigation actions, and resets. Mental demand was also reported as lower in PZN". [33]

The overall findings of the test showed that the overview, in this case, did not improve the efficiency of task solving, but was none the less preferred by many subjects. It also showed that in this type of tasks, rubber sheet navigation required too many actions and was subsequently both less efficient and less preferred.

In a similar study Hornbæk et al. examined the usability of zoomable interfaces with and without an overview. Here 32 subjects were asked to solve 2 types of tasks in maps of geographical content with eight different types of map objects (e.g. cities, airports, lakes etc.), each defined by individual colours.

The maps were of Washington and Montana. All in all there were 833 map objects on the Washington map, separated into three scale levels to accommodate semantic zooming; on the Montana map there were 806 map objects and it had an overview in the top right corner (see

[32] Nekrasovski, D. et al. (2006) p. 8
[33] Nekrasovski, D. et al. (2006) p. 8

Figure 11a and Figure 11b). The subjects were required to have spent less than 2 weeks in either location to ensure that the tasks were solved based on exploration rather than familiarity. According to the study, one of the advantages of using geographical maps is the fact that they contain *nested objects* (as found in hierarchies) and *data networks* (a connection between points of information). *Familiarity* can be an influential and unpredictable factor when the content is a geographic map but this could be avoided by keeping the content of the map completely abstract.

The subjects navigated the map by zooming and panning. Zooming was done by using the left and right mouse key. Holding down the left mouse key, resulted in zooming in, after a delay of 400 ms; this delay was implemented as the left mouse key also functioned as a panning control, by dragging the map with the mouse. The right key was used for zooming out. There was also a button implemented on the screen called “zoom out”, which, when clicked, brought the user back to the initial view of the map.

The first of the two task types required the subject to find a specific map object; in the context of the test this was referred to as a navigation task. The other task type was a browsing task, which required the subjects to locate a sequence of objects on the map.

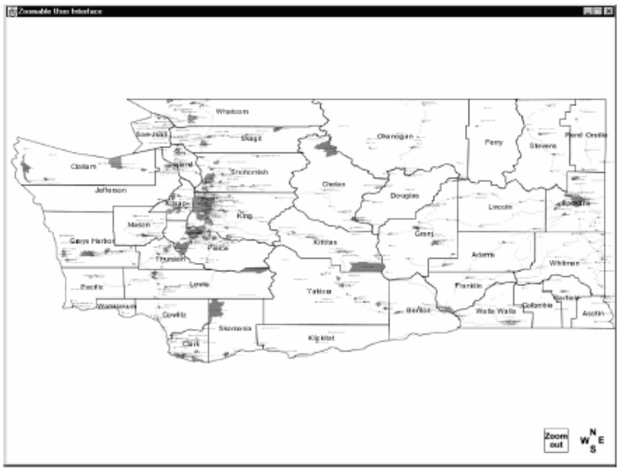


Figure 11a: A map of Washington without an overview, nodes supporting semantic zooming, used as content in a usability study on zoomable interfaces with and without and overview.

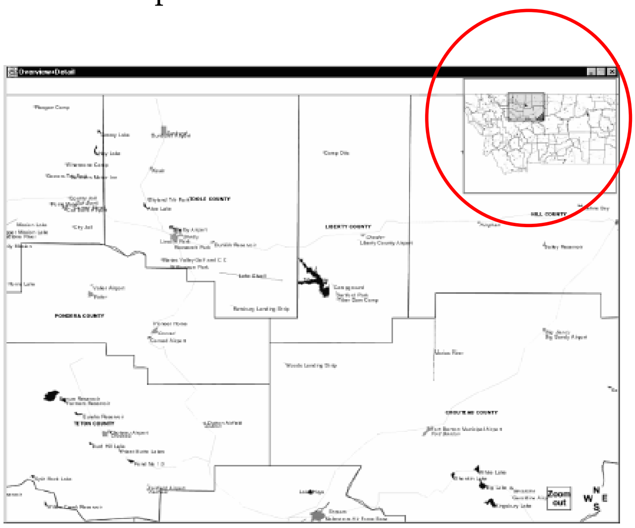


Figure 11b: A map of Montana with an overview (in the red circle), used as content in a usability study on zoomable interfaces with and without and overview.

In this study there were 2 hypothesis of relevance:

1. “Subjects would prefer the overview interface, because of the information contained on the overview window and the additional navigation features.”

2. "The overview interface would be faster for tasks that require comparison of information objects and scanning large areas." [34]

The subjects again preferred the detail+overview map, which may have led them to have the sense of control previously mentioned. However, the tasks given were solved faster in the map with no overview. This could, however, have been a consequence of the implicit overview given by the information space because it supported semantic zooming. [35]

These results seem to indicate that overviews have no bearing on the efficiency of zoomable interfaces, but they were preferred by the subjects in both cases. What neither of these studies answer is whether or not the continued use of overviews would be preferred by long time users of zoomable interfaces.

The results of Nekrasovski et al. show that to their surprise the zoomable interface was preferred to the distortion interface. One reason for this can be the familiarity in navigation patterns which was mentioned earlier as one of the key strengths of zoomable interfaces.

The 3D-like navigation ability of a pan/zoom navigation (PZN) technique allows flexible interaction with a 2D static or dynamic information space.

We maintain that in creating equilibrium between the constant increase of the information spaces in which users need to navigate, and the subsequent demands made on human cognition, choosing the right navigation technique and visualizing the data appropriately in the information space is essential.

Our opinion is that the strengths of PZN; perspective, geometric integrity and the familiarity in navigation, far outweigh the weaknesses; potential loss of contextual comprehension. Others have made bolder statements in this regard.

> "The zooming interface paradigm can replace the browser, the desktop metaphor, and the traditional operating system." [36]

---

[34] Hornbæk, K. et al. (2002) p. 7
[35] Hornbæk, K. et al. (2002)
[36] Raskin J., The Humane Interface (2000) p. 164

In the context of this paper it is our opinion that in regard to large multiscale interfaces the single view pan/zoom navigation (PZN) technique has legitimate strengths and is therefore a suitable choice for our further research.

> "[T]he smooth scrolling plus rapid and continuous zooming remains the secret of success for single view browsers." [37]

Because PZN represents an efficient way of accessing an area of interest in the information space, the potential application of this navigation technique to gaze interaction is the focus of our study. Because of the inconclusive results regarding overviews in zoomable interfaces, leaving these out when designing an interface could prove to be a valid option.

2.3 Alternative input devices

*Input device*s are used for manipulating digital information. Input has been and is conventionally exclusively done by hand interaction. Hands-discrete input devices [38], such as keyboards, and hands-continuous input devices [39], such as the mouse, have for decades dominated the world of PCs and as a consequence the field of Human Computer Interaction (HCI).

As mentioned earlier the size of information spaces is continually increasing. As tasks and data structures become more complex and ubiquitous in nature, the need for alternative input devices increases.

> "We can view the basic task of human-computer interaction as moving information between the brain of the user and the computer. Our goal is to increase the useful bandwidth across that interface with faster, more natural, and more convenient communication mechanisms." [40]

There are many situations where the conventional hand input devices are inadequate and alternative inputs can be necessary. In order to determine what kinds of situations require or could benefit from alternative inputs, three basic parameters should be considered. (1) What type of task needs to be solved; searching in visually represented data compared to textual data requires different approaches. (2) The context of the task; e.g. the user is physically restricted, as a doctor

---

[37] Plaisant, C. et al. (1995) page number missing
[38] Jacob, Robert J. K.(1996)
[39] Jacob, Robert J. K.(1996)
[40] Jacob, Robert J. K et al. (2003) p. 12

performing surgery or a paraplegic who only has use of her eyes. (3) What type of input device is being employed; e.g. the user is working with the small screen of a mobile device or a large digital whiteboard.

Similar to our discussion about visualization techniques in Chapter 2.1 it is possible to draw up a triangle showing how the task type, usage context and the choice of display and input devices all influence each other.

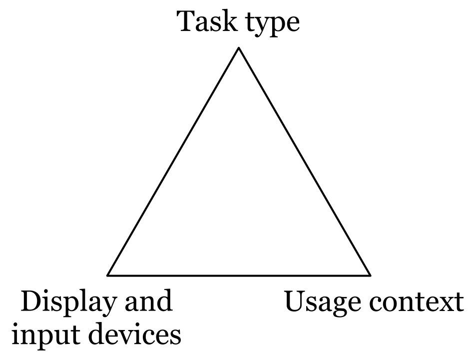


Figure 12: The relationships between task type, usage context and display and input devices. Changes in one potentially affect the whole triangle.

There are numerous variations of *alternative input devices*. Many of which use other parts of the body and do not require the use of hands such as; voice recognition, controls based on head movement and eye-tracking etc.

Eye-tracking and gaze

Eye-tracking has been an area of research for more than 100 years. Initially the process of tracing eye movements was an intrusive procedure employing mechanical objects being placed directly on the cornea. In the 1950's eye-tracking was conducted using head mounted equipment and in the 1970's the big leap came with: "the discovery that multiple reflections from the eye could be used to dissociate eye rotations from head movement".[41]

At present the invasive equipment has all but vanished and substituted with increasingly precise and fast eye-trackers. One implementation projects infrared light from several different points,

41 Jacob, Robert J. K; Karn, Keith S (2003): Eye-tracking in human-computer interaction and usability research: Ready to deliver promises p. 2

creating a stable reflection, essentially red-eye, then an infrared sensitive camera collects the information allowing a computer to track the direction of the *gaze*. [42]

This type of equipment is still quite expensive and therefore has accessibility issues; however, attempts have been made to make eye-trackers as accessible as possible. Low resolution gaze-trackers have been made with standard digital video cameras from Sony [43]. These systems are not as robust as the commercially produced ones. However, if the accompanying software applications are designed to tolerate noisier input, the systems work.

We use the term *gaze* because it is a more precise term than e.g. sight or vision as it deals not only with what is visible to us, but what exactly we are looking at. The tracking of eye movements has become a technology which allows different types of *gaze interaction* with the computer.

Workings of the eye

The eye is a light sensitive organ, which absorbs light impulses and directs them to the brain. Important elements of the eye are the transparent *cornea* and *lens* that direct light onto the *retina*. The retina consists of light sensitive cells that are different in both density and function, and are called *cones* and *rods*.

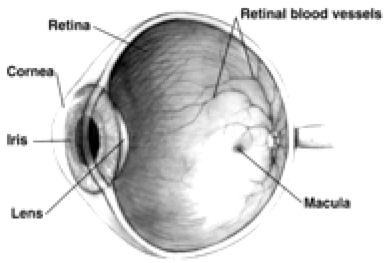


Figure 13: The structure of the eye. Incoming light is directed onto the light-sensitive retina with the highest density of cone-cells in the macula.

Rod cells exist mainly at the edge of the retina, and are responsible for our *peripheral vision*. For high definition viewing the cone cells are used. They are densest in an area called the *macula*, which is on the axis of the eye; this area is responsible for the 5° angle which is our central vision. In the centre of the macula is the *fovea*, which contains the highest density of cones, and is

---

[42] Product Description Tobii 50 Series
[43] Hansen, John Paulin et al. (2004): Eye-tracking off the Shelf

responsible for what is called *fovic vision*, which represents an angle of about 1°. We are essentially only seeing a very small part of the world in high definition. [44]

There are two basic temporal states in which the eye moves; *fixations* and *saccades*. Fixations are moments that last between 200- 600 ms,[45] in which the eye is relatively still and the central, fovic vision is used to view an object clearly. After a fixation the eye moves to a different fixed position; this type of movement is called a *saccade*. The saccades are ballistic in nature, which means, that the motion is predetermined, and the eye moves in a straight line from one point to the next. [46]

During this motion, which can cover between a 1° and 40° angle, most of the visual input ceases; this is called *saccadic suppression*. With saccades lasting between 30-120 ms, this suppression of information is a way of ensuring smooth visual input.

Saccades are reactions, forced or unforced, to visual information perceived in areas of lower resolution than that of the fovea; it is the eyes' way of making up for not being entirely made up of cone cells.

As an object cannot be perceived clearly unless it is seen by the limited area of the fovea, the position of a person's eyes is a good indication of that individual's point of interest.

Evolution of eye-tracking

For the better part of the eye-tracking history it has been used in the field of usability and cognitive psychology. Some of the earliest work includes studies done by Fitts et al. in the 1950's where the eye movements of 40 pilots were examined while landing a plane. The information was collected by using a combination of mirrors and a camera.[47] Needless to say the amount of data requiring analysis must have been overwhelming without automatic processing capabilities.

The complexity of the equipment, combined with the amounts of information which needed to be processed, made eye-tracking in usability studies an arduous task. Even so eye-tracking has been seen as promising since the 1950's and still is.

---

[44] Zeki, Semir (1993)
[45] Jacob, Robert J. K.(1991)
[46] Jacob, Robert J. K.(1996)
[47] Jacob, Robert J. K.(1991)

> "While there has been considerable use of eye tracking in usability engineering over the 50+ years since Fitts' pioneering work [...]. We see however, that just in the past ten years, significant technological advances have made the incorporation of eye-tracking in usability research much more feasible." [48]

This technological advance, mentioned by Jacob, has not only benefited eye-tracking in the field of usability research, but has since the beginning of the 1980' been a subject for visionaries in human computer interaction; using the technology as not only a tool for collecting data on eye movements, but as a real time input device.

In the beginning of the 1980's Bolt was one of the first who began to conceptualize how the use of eye-tracking equipment could be implemented in real time. In 1981 he wrote:

> "While the current application of eyes at the interface is in a sense a special case, eyes-as-output has a promising future in interactive graphics, especially where the graphics contemplates the possibilities arising out of the orchestrated use of videodiscs." [49]

One of the first evaluations of eye-tracking as an input device came from Ware in 1987. Three types of on screen selection techniques were tested; (1) *dwell time selection*, where a target is selected by the user fixating on it. (2) A *screen button*, which required the subjects to look at an object and subsequently look at a onscreen area which functioned as a button and finally (3) *a hardware button* where the subject pressed a physical button while fixating on an object.

In this test a video camera captured the eye movements by registering the boundary of the pupil based on a reflection of infrared light from the cornea; this information was then digitally processed. The system required the head of the subject to be fixed.

Two experiments were conducted; the first in a structure which was intended to resemble a menu. There were two representation of the fixation point; one large which reflected the current point of fixation and one smaller which represented the previous point (see Figure 14).

---

[48] Jacob, Robert J. K.(1991): The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get p. 11

[49] Bolt, Richard A. (1981): Gaze-Orchestrated Dynamic Windows p. 10

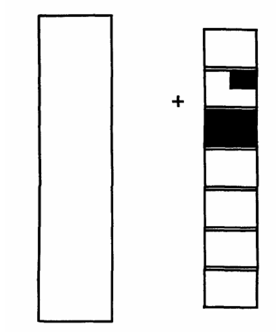


Figure 14: The onscreen layout of the Ware experiment in 1987. The rectangle to the left is the on-screen button, the field on the right is the menu representation.

The first test required the subjects to select targets in a static version of Figure 14. The result of this experiment showed dwell time selection to be equally as fast as hardware selection and with lower error rates in the dwell time selection than in the hardware button. In the second test the subjects were asked to select targets of different sizes. In this experiment the hardware button was faster, but the dwell time selection was more consistent. However the error rates for dwell time selection were the lowest. (The onscreen button was by far the most ineffective in both tests). The overall conclusion of this test was:

> "...where speed is of the essence, cost is no object, sizes are moderate, and it is important that the hands be reserved for other activities, the eye-tracker, may be the input device of choice." [50]

One area of HCI which has taken this ability of onscreen selection to its core is the field of *assistive and alternative communication*; where eye-tracking often is used as the sole input device for selection and navigation. The hands can in some cases be restricted, not only because of other activities, but as a consequence of an innate or acquired physical disability.

GazeTalk and Dasher [51] are two examples of gaze typing systems, which allow users to communicate solely by gaze. Systems like these are giving disabled people the possibility to independently interact with the world, some for the first time. GazeTalk not only supports typing but includes other options such as online browsing.

Another way of using eye-tracking as a real time input is found in the field of *attentive interfaces*. Here the input received from the user is augmented to include subliminal information which the

---

[50] Ware, C. et al. (1987) p. 6
[51] Itoh, Kenji et al. (2006)

system will then respond to. Attentive interfaces can react automatically to commands without interrupting the user and attentive phones can communicate to others the availability of the user.[52]

One implementation of attentive gaze-controlled interfaces which has received a lot of attention is gaze-contingent multiresolutional displays.[53] This method attempts to solve the bandwidth and processing problems of large high-resolution displays; by only showing the viewer's point of attention in full resolution (see Figure 15).

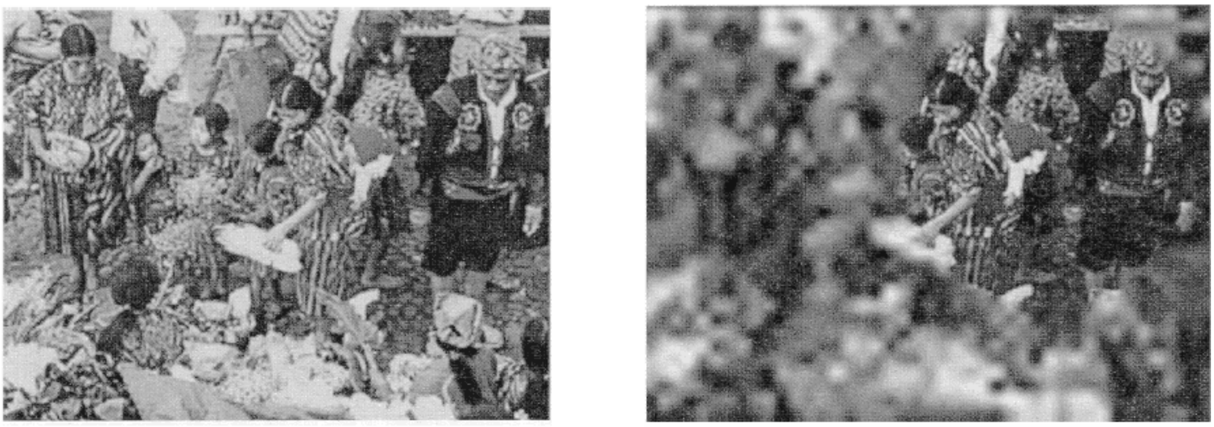


Figure 15: Gaze-contingency: An image displayed in complete high-resolution (left) and displaying high-resolution only at the point of attention of the viewer (right).

Gaze interaction when used as a selection and navigation input device has the advantage of extending the existing behaviour of orientation and locating an area of interest, to also involve a direct and active response, eliminating the additional action of moving the mouse.

The eye is in almost constant use during our waking hours and therefore represents a very durable form of interaction. This is due to the nature of muscles controlling the eye, they offer near fatigue-free pointing.[54] Gaze interaction could therefore be of assistance to many people who suffer from repetitive muscle strain.[55]

On the other hand gaze as sole input has the problem of only having one mode; the eye-tracker recognizes that someone is looking at the screen, not the intention behind the gaze; the input device is always on.

---

52 Hyrskykari, A. et al. p. 7
53 Reingold, E. M. et al.
54 Bates, R. et al. (2002) p. 119
55 Zhai, S. et al. (1999)

This results in a potential Midas' Touch problem. Midas was a character in a Greek myth, after doing service to a god, he was granted a wish. He wished that all he touched would turn to gold. When this wish was granted, he quickly found a flaw in his plan, as all he touched, including food and drink, indeed turned to gold and he almost starved to death.

Because of gaze's single modality of always being on there is the risk of selection and activation when it is not required or wanted:

> "Everywhere you look, another command is activated; you cannot look anywhere without issuing a command. The challenge in building a useful eye-tracker interface is to avoid this Midas' Touch problem." [56]

One of the solutions to this problem has been dwell time activation, which is intended to ensure that the user is actually fixating on a specific object and not just passing it by with a glance. [57]

There might also be a potential problem in using the eye both as an input modality to the user, looking at feedback from the system, and as an output modality from the user, moving a pointer on the interface. [58]

Finally two major problems with gaze in navigation and pointing are that we are not used to having objects, onscreen or otherwise, react to our sight and that we rarely need to employ the type of precision required for controlling information spaces with gaze, as it is relatively difficult to control eye position precisely as a pointing device. [59]

The limitations of eye tracking were the foundation for a study done by Zhai et al. in which the possibilities of using the eye tracker as a supplement to the mouse, rather than a substitute, were examined.

> "We believe that many fundamental limitations exist with traditional gaze pointing. In particular, it is unnatural to overload a perceptual channel such as vision with a motor control task." [60]

---

[56] Jacob, Robert J. K.(1991) p. 13
[57] Hansen, J. P. et al. (2003)
[58] Bates, R. et al. (2002)
[59] Jacob, Robert J. K. (1991)
[60] Zhai, S. et al. (1999) p. 1

They designed and implemented the concept of MAGIC (Manual and gaze input cascaded) pointing. The basic idea was to maintain selection and targeting as motor controlled tasks, only using gaze indirectly by allowing the cursor to follow the point of interest (gaze) and then performing the actual targeting and selection by motor control (e.g. using a mouse).

They implemented two different designs, a liberal and a conservative. The liberal pointer placed the cursor within the vicinity of every object looked at by the user; the user could then decide to select or ignore the target and move on to the next. The conservative pointer wouldn't move to the target until the manual input device (a miniature pointing stick as used in laptops) had been set into motion (see Figure 16).

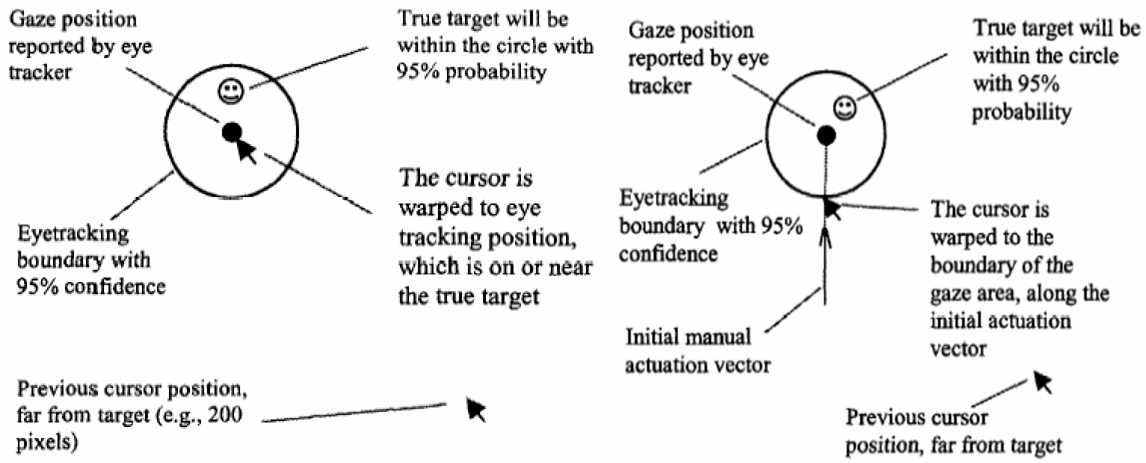


Figure 16: The MAGIC pointer: Explanation of the behaviour of the *liberal* MAGIC pointer (left) and of the *conservative* MAGIC pointer on the right.

Nine subjects participated in a trial where both designs were tested and compared with a purely manual input. Subjects were asked to point and click targets, which randomly appeared on the screen (see Figure 17).

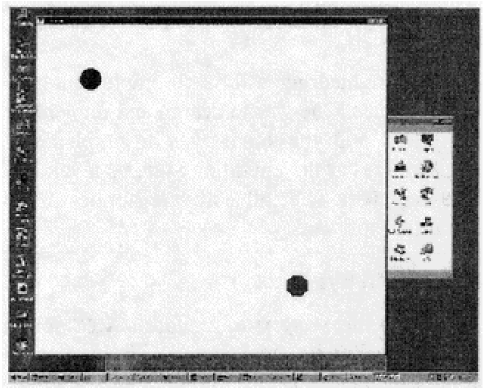


Figure 17: The MAGIC pointer tasks: Subjects were asked to click on targets of varying sizes appearing randomly on the screen.

The results showed that of the two gaze designs the liberal pointer was a bit faster than the completely manual input, and the conservative pointer slightly slower than the manual input. Overall the conclusion was that gaze pointing as a supplement to manual control showed great potential.

Zoomable interfaces is another way of attempting to solve the problem of precision in gaze controlled target selection; by allowing objects of interest to be enlarged and thereby require less selective or navigational precision.

Research on gaze and zoomable interfaces

The zoomable interface viewed in regard to gaze interaction makes for a valid design choice for two distinct reasons. (1) The possible *perspective* gained works at alleviating one of the main problems with gaze interaction which is precision; both in terms of equipment and the natural movements of the eye. (2) The naturalness of eye movements is extended with the *familiarity* of zooming.

The idea of coupling gaze with zooming actions has been explored since eye-tracking was first conceived of as a real time input device. However, considering the time-frame surprisingly few have researched this specific area. Below are different studies, which have all sought to implement gaze into zoomable interfaces.

In 1981 Bolt envisioned the concept of Gaze-Orchestrated Dynamic Windows. The basic idea was to allow for dynamic interaction with multiple windows in a single large display. The eye-tracking technology used was in the form of a pair of eye-tracking glasses used to inform the display of where the user was looking.

A system imitated a real time streaming of several television episodes, which were represented simultaneously on the display. Depending on which television episode the user was looking at, all other stations were additively muted and after focus had remained on the same image for a certain amount of time it would zoom out and fill the entire display. However, the technology at the time did not support a real time application of the system. [61]

A similar idea was recently explored by Fono et al.[62] in a study that evaluated gaze selection in, what they call, zooming windows.

They conducted 2 experiments, 12 participants took part in the first test and 10 in the second. The first was to determine choice of selection techniques, four different techniques were tested; eye-tracker with key activated selection, eye-tracker with automatic selection, mouse with click activated selection, and hotkeys.

Part of the tasks were based on stimulus and response. The subjects were presented with a grid onscreen (Figure 18a). When a field in the grid changed colour (Figure 18b) the subject had to select that field and the response time was registered. Upon selecting the field a text string would become visible (Figure 18c) and the subject was then to mirror the text in an activated field immediately below (Figure 18d).

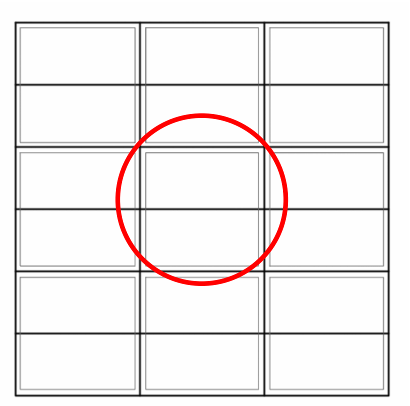


Figure 18a: The zooming windows study: Basic grid for first experiment tasks.

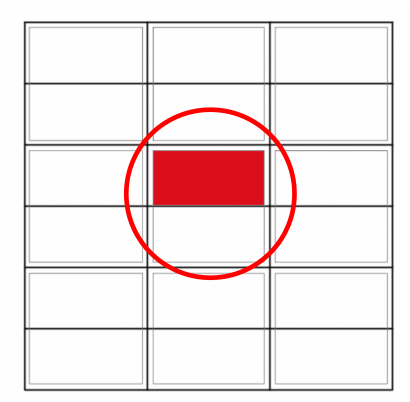


Figure 18b: Initial stimuli, a field turns red.

[61] Bolt, R. A. (1981)

[62] Fono, D. et al. (2005)

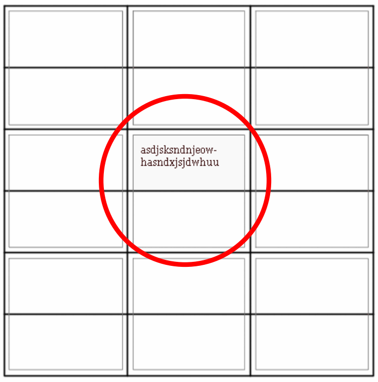


Figure 18c: The field is selected and reveals a pre-written text.

Figure 18d: The field below becomes active and the subject has to retype the text string above.

Their results showed that eye-tracking with key activated selection was approximately 35% faster than mouse activated selection and was subsequently chosen to be the sole selection method in the second test, which compared zooming windows with static windows.

Similarly to the first test the tasks required the subjects to interact with a grid. In the zoom windows version a field would turn red, upon selection the field would zoom out and the subject was able to enter a text string (Figure 19b). In the static version the fields in the grid remained the same size. Results showed that zooming windows were up to 30 % faster to work with than static windows.

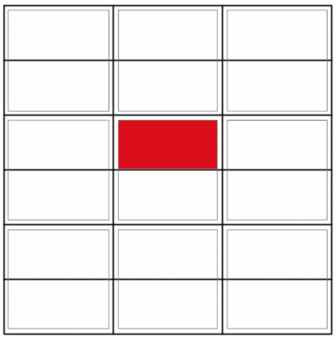


Figure 19a: The zooming windows study: The field turns red, and can be selected.

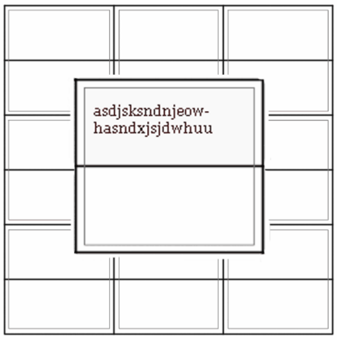


Figure 19b: After selection the field chosen zooms up.

The second notable study of a gaze-controlled zooming interface was done by Pomplun et al. in 2001. Their implementation employed two windows, A and B, of the same size. In window B subjects could view a magnified part of the whole image shown in window A. This was achieved by fixating on a part of the image in A for at least 120 ms and then looking over to window B, where a magnified view of the point of fixation was then showed. To zoom in on another part of the image this process was repeated. This gaze behaviour was compared to a mouse technique where clicking in window A (on the left) resulted in a magnified view in window B (on the right) (see Figure 20).

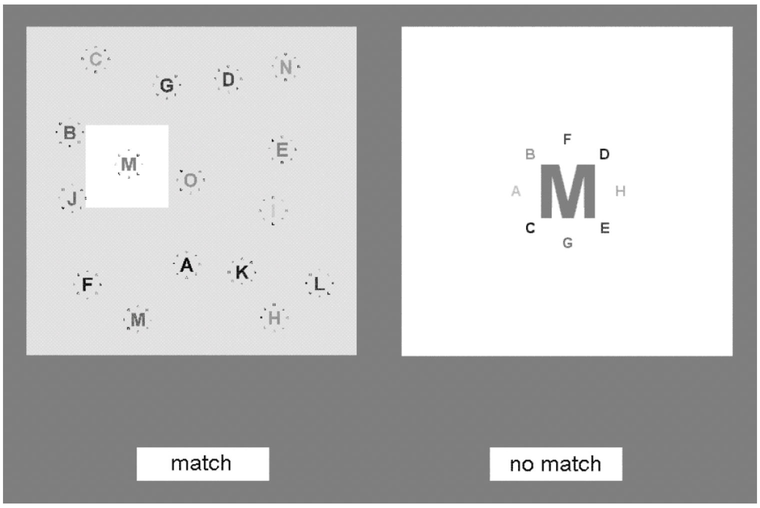


Figure 20: A two window implementation of a gaze-controlled zooming interface by Pomplun et al.: Fixating on the letter M (indicated by a highlighted square) in the window on the left (window A) and then gazing to the window on the right (window B) shows a magnification of the activated area.

The interface only provided two magnification levels. Subjects were asked to look for a specifically coloured small letter, for instance if there was a small green B on the map. In order to solve these tasks, subjects had to repeatedly zoom in on the large letters. Subjects completed the tasks by selecting either the "match" or "no match", using dwell-click or mouse. In 50% of the tasks no match existed, requiring the subjects to search the whole map. These "no match" tasks are the ones discussed in the results.

They only had four subjects, but repeated the tests over a longer period of time. The main results were that despite more magnifications being triggered in the gaze conditions than in the mouse conditions, the difference in response time was less than 10% in the favour of the mouse:

> "[...] demonstrating that, despite occasional unintentional selections for magnification, using the unfamiliar gaze interface and wearing the eye tracker headset did not considerably slow down participants' task performance as compared to the familiar mouse interface." [63]

This and the lack of noticeable practice effect for the gaze implementation lead them to conclude that eye movements can in fact substitute manual input, and that appropriately designed gaze-

[63] Pomplun, M. et al. (2001) p. 5

controlled interfaces could replace conventional interfaces in areas of application where hands-free interaction with machines is desirable.[64]

A final experiment of great relevance to our study is that conducted by Bates et al.[65], comparing different pointing devices in zooming and non-zooming environments based on a traditional WIMP (windows, icons, menus and pointing devices) style of interaction.

A standard eye mouse and a zoom eye mouse, in their terminology, were compared with a standard head mouse, an instrument that translates the movement of the head into input. The tasks had their origin in 'real world' computational interactions; one being word processing, the other web browsing. 6 test subjects participated in the experiment, none of which had physical disabilities.

Their results showed that the standard head mouse was superior to the standard eye mouse; however, the addition of zoom allowed the eye mouse to surpass the head mouse in efficiency.

The zoom environment also seemed to ease the workload and subjects deemed it better than the head mouse for ease of use.

The overall conclusions of these three experiments show that gaze interaction and zoomable interfaces in combination lead to an increase in efficiency for navigation and target selection.

Where target selection is the main focus of the above mentioned texts, little research has been done which centres purely on the navigational aspect of gaze interaction in zoomable interfaces, which is the focus of our main experiment. Navigation should here be understood as a process of reaching the target, rather than the actual selection.

Information visualization uses the strength of our perception to mould the information space, as an aid in the decision-making process. With the possibility for stable and accessible gaze interaction, our perception could also be directly applied to navigation.

In the implementation tested in the second test phase, we chose to sidestep the issue of selection/activation and focus instead on gaze interaction as a navigation technique. We therefore combined the gaze control with a pan and zoom navigation technique, studied its performance and compared it to more traditional input; a classical mouse.

---

[64] Pomplun, M. et al. (2001)
[65] Bates, R. et al. (2002)

2.4 Naturalness versus familiarity

Reoccurring themes many HCI matters are discussions about to which extent an interface idea is *intuitive* or *natural*. Examples of this can be found in relation to practically all the concepts we have introduced in the preceding chapters; be it information visualization, navigation techniques, alternative inputs or gaze interaction.

Although both terms seem self explanatory and relevant for interface designs, care should be taken when using them.

In his book, The Humane Interface, Jef Raskin discusses the problems with using these terms, and instead proposes using terms like *familiar* or *easily learned*:

> "Many interface requirements specify that the resulting product be intuitive, or natural. However, there is no human faculty of intuition, as the word is ordinarily meant; that is, knowledge acquired without prior exposure to the concept, without having to go through a learning process, and without having to use rational thought." [66]

In our case the use of eyesight to gather information is clearly natural, but any attempts at using gaze-direction for navigation or activation can at best be described as easily learned or based on a familiar analogy. In fact the only thing that humans naturally activate using eyes are other humans, and perhaps to some extent other mammals.

Our visual perception includes automatic processes observing other peoples' eyes, allowing us to notice when someone looks at us, even across a crowded room. We use gaze to non-verbally communicate with other humans and make our intentions more clear. Infants seem to become aware of faces at an incredibly early age and quickly learn to use their gaze as an indication of intent, and even following the gaze of adults.[67]

So while it is natural to look at things we want to interact with, in nature inanimate objects do not respond to our gaze. But gaze-interaction can be made easily understandable by mapping behaviour we know from our natural environment, and one approach of this is combining gaze with zoom/pan navigation as we attempt in our experiments. Bederson et al. describe the

---

[66] Raskin, J., The Humane Interface (2000) p. 150
[67] Voss, J.C. (2005)

motivation behind the zoomable interface of Pad++ as "tapping into people's natural spatial abilities".[68]

Even the use of the familiar computer mouse may not be as natural as we tend to think. In the Humane Interface Raskin tells the story of a Finish scientist and her first encounter with a mechanical mouse. She begins by turning the object on its head and rolling the ball around, as with a modern day track ball. This has no effect, as the ball is pushing down on all the detectors, thereby cancelling the effect. She discovers this and turns the mechanical mouse the right way up, but instead of placing it on a surface she continues to control the mouse manually, moving the ball with her fingers. When shown how to use the mouse she quickly understood the mapping – easy, but not natural. [69]

In this paper we try to use the terms intuitive and natural only with caution, usually by using the concept familiarity instead.

Modality

An example of behaviour that we have become used to in human-computer interactions, but can hardly be described as natural is the concept of *modes* and *modality*.

Raskin defines modes in terms of responses to a *gesture,* which is "a sequence of human actions completed automatically once set in motion".[70] If the response of the system to a specific gesture is not always the same it is because the system has more than one mode; it is modal.

An example of this is the behaviour of the Caps Lock key; pressing it changes the state of the system so that instead of the gesture of pressing a key resulting in a lowercase letter, the result is now an uppercase letter. (Except of course if the system previously was in Caps Lock mode, in which case the behaviour is reversed.)

To avoid the inherent confusion of using of modes in interface design, Raskin recommends using quasi-modes that are kinaesthetically maintained:

---

[68] Bederson, B. B. et al.(1996) p. 2
[69] Raskin, J., The Humane Interface (2000)
[70] Raskin, J., The Humane Interface (2000) p. 37

Caps Lock leaves the system in a state of writing uppercase letters and needs to be deactivated; a mode. The Shift key also implements uppercase letters, however as soon as the user moves on and releases the key, the state of the system returns to a default state; a quasi-mode.

We have previously mentioned modality in our discussion of gaze-interaction in relation to the problem of making the computer aware of the intentions behind our gaze. In a gaze-controlled zoom/pan interface some additional information is needed to inform the computer of whether our fixation is an indication of us wanting to zoom further or if we are content with working at the current scale.

This question was discussed as far back as in 1981 by Richard Bolt when describing the implementation of the "World of Windows"-display technique:

> "There are at least two competing philosophies about what should cause you to be "zoomed-in" upon some window you are looking at: 1) zoom-in automatically, based upon some timing-out of how long you look at ("stare at") some certain window; 2) zoom-in upon the window you are eye-addressing contingent upon a deliberate action via an independent modality (e.g., joystick action; or a word spoken to the system)"[71]

The first option is what we would today call dwell-activation and the second involves using some input in addition to the eye-tracker. In our implementation of a gaze-controlled pan/zoom interface described in the following chapter we chose the latter option by using two keys on a keyboard for zooming in and out in a quasi-modal fashion.

2.5 Our design decisions

The basic premise for the environment of our experiments is that large nodal information spaces present challenges in terms of navigating them effectively and efficiently.

Pan/zoom navigation (without semantic zooming) has three valid strengths; geometric integrity, perspective and familiarity, which can all be applied beneficially to search and navigation tasks in large nodal information spaces.

---

[71] Bolt, R. A. (1981) p. 8

Gaze interaction has a cognitive advantage, in skipping the mouse as a pointing device, which could ensure its efficiency, when implemented in a visually represented information space and with a navigation technique that capitalizes on its strengths.

In light of these concepts and others introduced in the preceding text we can summarize the design decisions for the gaze-controlled implementation that we tested in our second round of experiments:

We chose to study navigation in a multiscale system using a geometric zoom/pan interface; maintaining the size proportions between all elements, without a separate overview presentation.

- For comparison testing we evaluate the efficiency of using gaze as a means of pointing against using a mouse.
- We focus on navigation, rather than selection, in large nodal information spaces and thereby avoid issues concerning Midas' Touch and dwell time activation.
- The zooming is controlled by a quasi-modal implementation; to zoom in or out a key on the keyboard is pressed, releasing the key stops the zoom motion.
- As an attempt of implementing panning in a gaze-controlled pan/zoom interface we try to map the familiar behaviour of gazing in the direction of an object of interest.

Hypotheses on gaze-controlled pan & zoom navigation

As a consequence and in light of the design choices we made, three general hypotheses were formed:

1.1. Eye-tracking can effectively supplement existing input devices.

1.2. Zooming interfaces have traits that correspond well with eye-based interaction.

1.3. A gaze-controlled navigation technique should be able to successfully entail many familiar aspects of sight and perception.

3. Empirical studies and test-design

In order to gain practical insight into the area of gaze interaction and pan/zoom navigation (PZN) interfaces, we conducted 2 rounds of experiments. In the first we tested different aspects of PZN and information visualization and the second had the more specific goal of comparing gaze and mouse interaction in PZN environments. The following chapter discusses the design and implementation of both experiments.

The results of the second experiment are our main area of interest; therefore they will be discussed and elaborated upon in chapter 4.

3.1 Experiment 1

In the initial experiment 4 different mouse-controlled navigation techniques (NT) were tested, the intention was to explore different aspects of pan and zoom. Three of the NTs dealt with variations of pan and one with zoom.

The data which constituted the information space was a playing card metaphor, chosen because of it being both abstract and easily recognisable. The four navigation technologies were tested by having subjects solve tasks that required them to either find one specific target or to explore a sequence of targets. The information space was presented in 3 different ways; structured, semi-structured and random. The script for conducting the tests can be found in Appendix D.

Information space and tasks

The data was chosen to be easily recognizable and keeping the cognitive processes required for solving the tasks at a minimum. Playing cards as a metaphor possessed the desired qualities and had three dimensions; colours, suits and numbers. To allow a larger information space, the four classic suits were supplemented with combinations of well known icons and basic colours. So in addition to the classic “seven of spades” (♠ 7), we also have “seven of squares” (■ 7), etc.

So as not to make the information space too visually confusing, only the numbered cards were used (1-10). The content was structured in three different ways, and then displayed in a map consisting of rows and columns:

1. **Fully structured content:**

- Eight columns, each containing a single suit: ♥, ♠, ♀, ■, ♦, ♣, ♂, ●
- Four to eight cards in each column (randomly chosen), six on average.
- Chronological order of numbers with the lowest at the top of a column.

♥ 1	♠ 1	♀ 2	■ 2	♦ 3	♣ 1	♂ 1	● 2
♥ 3	♠ 2	♀ 3	■ 4	♦ 5	♣ 5	♂ 2	● 5
♥ 5	♠ 4	♀ 4	■ 6	♦ 8	♣ 7	♂ 3	● 6
♥ 6	♠ 5	♀ 7	■ 9	♦ 9	♣ 10	♂ 4	● 8
♥ 8	♠ 7	♀ 8	■ 10	♦ 10	-	♂ 5	● 9
♥ 10	♠ 8	♀ 10	-	-	-	♂ 7	● 10
-	♠ 9	-	-	-	-	♂ 9	-

Figure 21: A fully structured map from Experiment 1: Suits in individual columns and numbers in chronological order.

2. **Semi structured content:**

- Eight columns, each containing a mix of two suits in the same colour: red (♥/♦), black (♠/♣), green (♀/♂) and blue (■/●).
- Four to eight cards in each column (randomly chosen), six on average.
- Random order of numbers and icons within each column.

♦ 9	♣ 2	♀ 10	● 1	♥ 5	♠ 1	♂ 3	● 7
♥ 8	♠ 5	♀ 3	■ 10	♦ 1	♣ 9	♀ 7	■ 9
♦ 2	♣ 5	♂ 9	● 10	♥ 1	♠ 8	♂ 10	■ 2
♦ 5	♠ 3	♀ 1	■ 1	♦ 6	♣ 7	♂ 6	● 6
-	♠ 6	♀ 6	● 5	♥ 2	♣ 4	-	● 2
-	-	♂ 1	-	♥ 9	♠ 2	-	■ 5
-	-	-	-	♦ 3	♠ 4	-	● 4

Figure 22: A semi structured map from Experiment 1: Suits of same colour mixed within columns and numbers randomly ordered.

3. Randomly ordered content:

- Seven columns.
- Four to eight cards in each column (randomly chosen), six on average.
- Fully random order of cards.

♦ 8	● 3	♠ 7	♂ 2	♣ 4	● 4	♠ 10
♥ 10	♀ 6	♣ 1	♥ 8	♠ 9	♦ 5	■ 8
♂ 10	♦ 2	● 5	♣ 3	♂ 5	♦ 10	♠ 5
♠ 3	♂ 4	♀ 1	♣ 7	● 10	♥ 6	♀ 2
♥ 9	♣ 9	♂ 8	♀ 9	-	♥ 2	♀ 8
■ 7	-	♠ 6	♥ 5	-	■ 3	♠ 1
-	-	♂ 6	♂ 3	-	■ 1	-

Figure 23: A random map from Experiment 1: Totally random order of cards.

We designed 12 different maps for the experiment: 3 differently structured maps for each of 4 navigation techniques. This was intended to prevent a learning effect and to ensure that the tasks were solved based on exploration rather than recognition.

The tasks were designed to produce quantitative as well as qualitative data, allowing comparison of the NTs. The tasks in this experiment (as well as in the second experiment) can be compared with Plaisant et al.'s taxonomy of visual browsing behaviour, used for instance by Baudisch et al.:

> "*image generation* (e.g., graphic design), *open-ended exploration* (e.g., submarine remote operated vehicle ROV operation), *diagnosis* (e.g., mechanical engineering), *navigation* (e.g., gamers, ROV), and *monitoring* (e.g., air traffic control)." [72]

The first type of tasks in our experiment required the user to find a specific card: Find the 7 of hearts, 2 of diamonds, etc. These tasks are comparable to the definition of *open ended exploration*:

[72] Baudisch, P. et al. (2002) p. 2

> “The overview of the space being explored is not always complete or even available since it is explored for the first time. The navigation needs to be fast and the user interface quickly mastered” [73]

There is a slight difference between this definition and the tasks in our experiment. In Plaisant’s definition the overview is the goal itself; in our experiment the overview is the means to an end, which is finding the specific target. We therefore chose to call this type of task, in our experiment, *target exploration*.

The second type of task was to find and count all cards of a specific type: How many 8’s are there? How many hearts are there? etc. These tasks required the subject to scan through the entire dataset to find the correct answer and can be directly compared to the definition of *diagnostics*:

> “Most of the time is spent panning the image, looking for patterns, therefore panning speed is crucial, so is coverage as skipping part of the image can result in the wrong diagnosis. But a complete automatic scan can also lead to boredom and errors.” [74]

All tasks in the test required planar movements (panning) to a large extent, as only 3x3 cards were visible at each time, although in one navigation technique these planar movements were only achievable by zooming in and out (showing more cards than 9).

Navigation techniques

As mentioned 4 navigation techniques were tested by solving tasks in the maps. Three of them employed navigation by panning and one by zooming.

Pan 1: Scrolling
This NT allowed the subjects to scroll through the information both horizontally and vertically, by placing the mouse cursor near the edges of the active field. The closer to the edge the cursor was placed, the faster the scroll speed. We chose to call this interaction “edge scroll” and discuss it further in Chapter 3.3. The reason for testing this type of pan was primarily that it seemed a viable idea for implementation in a gaze-controlled interface.

---

[73] Plaisant, C. et al. (1995) page number missing
[74] Plaisant, C. et al. (1995) page number missing

Pan 2: Grabbing

In this NT the subject could move the map by grabbing it with the left mouse button, and drag it around. On releasing the mouse button the movement stopped.

However, in our implementation using a Flash application within a browser window the active field was smaller than the screen, and the grab only functioned properly within the application. As a result releasing the mouse button outside the active field was not registered, requiring the subject to click once within the active field to "release" the map. Even though subjects were made aware of this when the NT was introduced to them, this behaviour still proved to be an annoyance.

Pan 3: Cylinder

The third pan NT was an attempt at exploring ideas about an navigation equivalent to a Rubik's cube. The subjects navigated through the information by controlling intertwining cylinders. There was a primary horizontal cylinder, controlled by panning left and right, which affected the entire map in a continuous loop. Vertical scrolling only affected single columns, also in a loop.

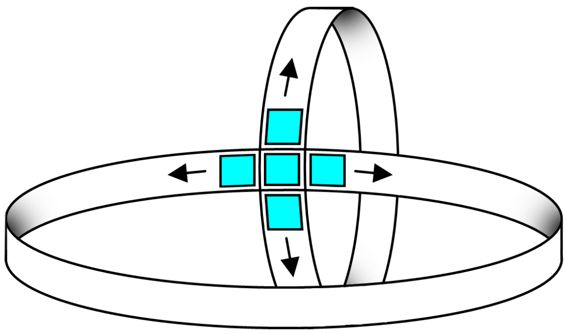


Figure 24: A schematic visualization of the cylinder navigation technique in experiment 1. Horizontal scrolling affects the entire content. Vertical scrolling only affects a single column at a time.

Zoom:

In this NT we implemented a zooming option, based on the code we received from Aza Raskin that later became fundamental to our second round of experiments. The subjects used both a mouse and keyboard; pointing with the mouse and zooming in or out by using the up and down arrows on the keyboard.

Our intention was to get a sense of zooming as an isolated navigation technique, and we therefore took measures to ensure semantic zooming. To begin with we included no option for panning and secondly we weakened the advantage of gaining a perspective, by exponentially scaling the numbers down when zooming out. As a consequence, having a perspective of the whole map meant that the numbers were invisible. This design forced our subjects to use zoom as a means of navigating from one location to another. The following images display this behaviour, with the dashed rectangle representing the area showed in the first image.

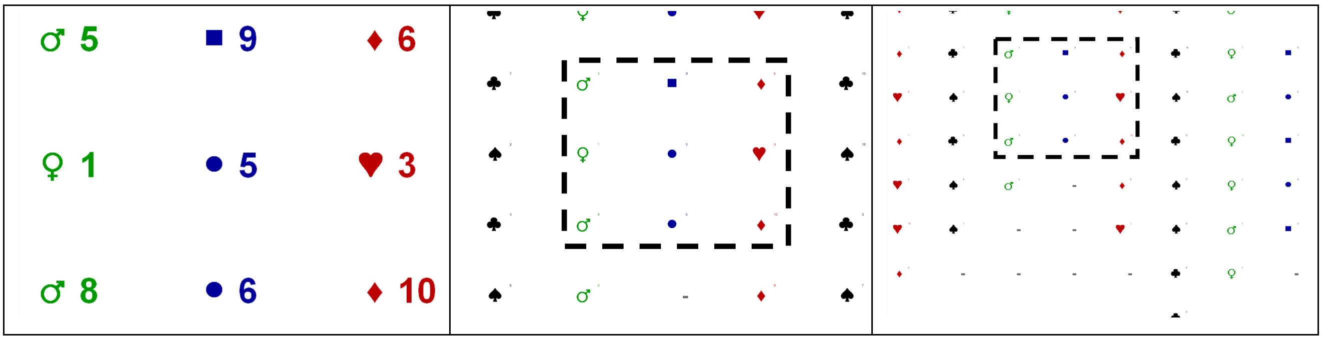


Figure 25: The behaviour of the semantic zoom in experiment 1: Initially the numbers and icons are equal in scale (left). On zooming out the numbers are scaled down exponentially faster than the icons (middle). Before gaining a perspective of the whole map the numbers have completely disappeared (right).

Apparatus

We conducted the study on a system running Windows XP with a Pentium 4, 2.0 GHz processor, 512 MB RAM, a 64MB GeForce4 MX 420 video card and a 17 inch monitor at a resolution of 1280x1024 pixels. The test itself was developed using Macromedia Flash 8 and ran in a Flash Player 8 within an Internet Explorer browser.

Procedure

There were 12 test subjects (4 females) between the age of 23 and 35 who participated in the study. All of our subjects were right handed. They were all frequent users of computers and needed no introduction to the use of mouse or keyboard.

The experiment took from 40 to 60 minutes depending on the subject. There was a general introduction welcoming the subject and explaining the overall procedure. The three different types of maps were presented along with a brief introduction to the type of tasks they would encounter.

There was one block of tasks relating to each of the 4 navigation techniques. At the start of each block the subject was introduced thoroughly to the NT and was allowed to experiment with it at will.

After each block the subjects were asked about their likes and dislikes with the technique and how well they had gotten a sense for the location of cards within the map. After completing the test subjects were asked a few questions regarding possible applications for the different NTs. (See Appendix D for full script)

All task times were measured by the subject starting and stopping each task and automatically logged. The moderator also logged approximate task times (in case of problems in the test process) and registered any errors in the answers given by the subjects.

Results

The results and general observations from this first experiment were influential on the context and design of the second round of experiments. Presented below are some of the most important findings, with more data presented in Appendix A.

For clarity the results are presented in charts without error bars, but standard deviations in our tasks were typically in the region of 20-40% of the mean results.

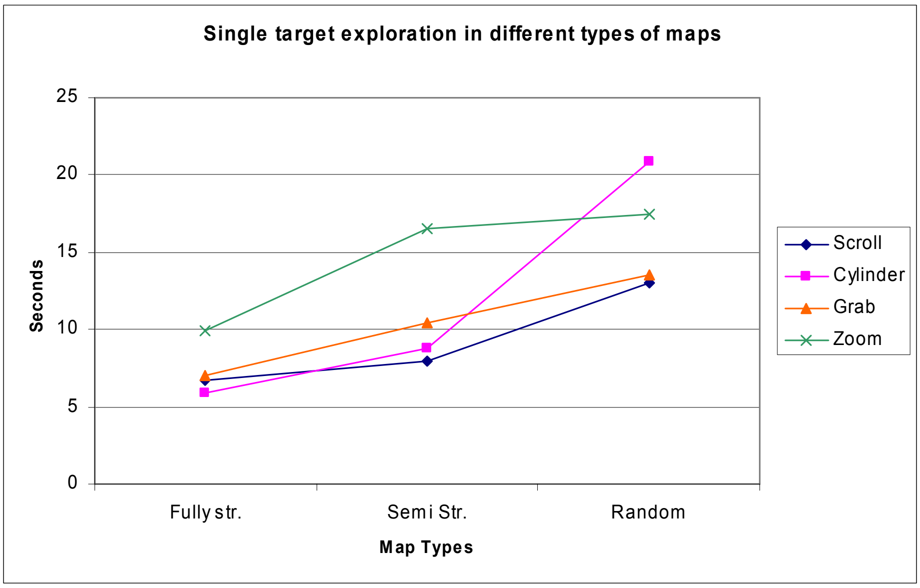


Figure 26: Experiment 1: Average task times for finding a specific target in different types of maps, using the four navigation techniques (scroll, cylinder, grab and zoom).

Figure 26 shows the average time for all single target exploration tasks, such as finding the 3 of diamonds. It reveals that, without the benefit of perspective and the ability to pan, zooming is relatively inefficient. It also shows that the cylinder NT is quite efficient in structured maps, but completely sub par in random maps. The results for panning with edge-scroll and grabbing are practically identical.

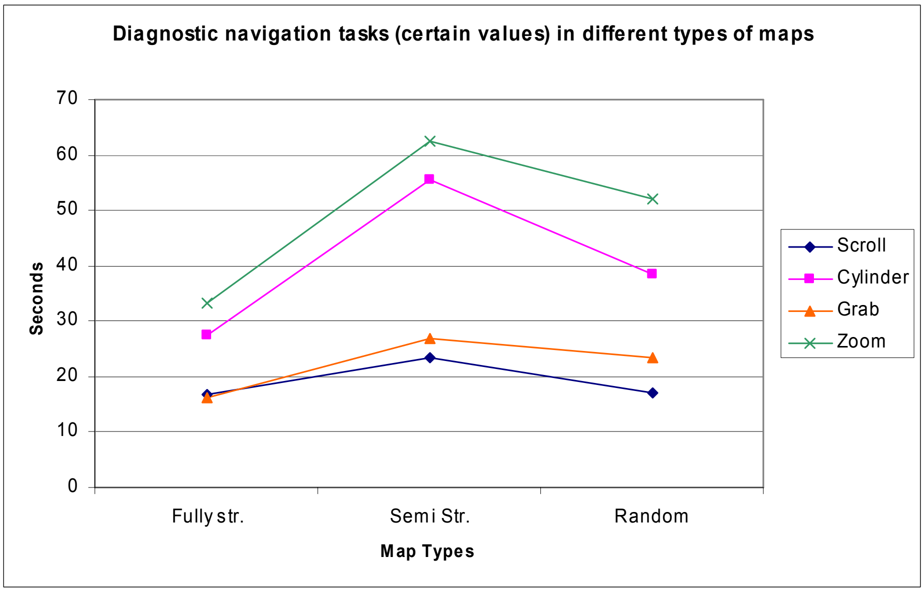


Figure 27: Experiment 1: Average task times for counting cards of a certain value in different types of maps, using the four navigation techniques (scroll, cylinder, grab and zoom).

Figure 27 shows the average times for tasks asking the subjects to count cards of a certain value, such as: How many 7's are there on the map? These were diagnostic tasks which required the subject to scan the entire body of information.

Once again the results for the grabbing and scrolling NT are almost indistinguishable. The semantic zoom function again shows the worst performance, closely followed by the cylinder NT. Even if the cylinder NT performed marginally better than zoom it was by far the least preferred by the subjects in general:

> "It was really terrible, it took me a long time to before I could orient myself"

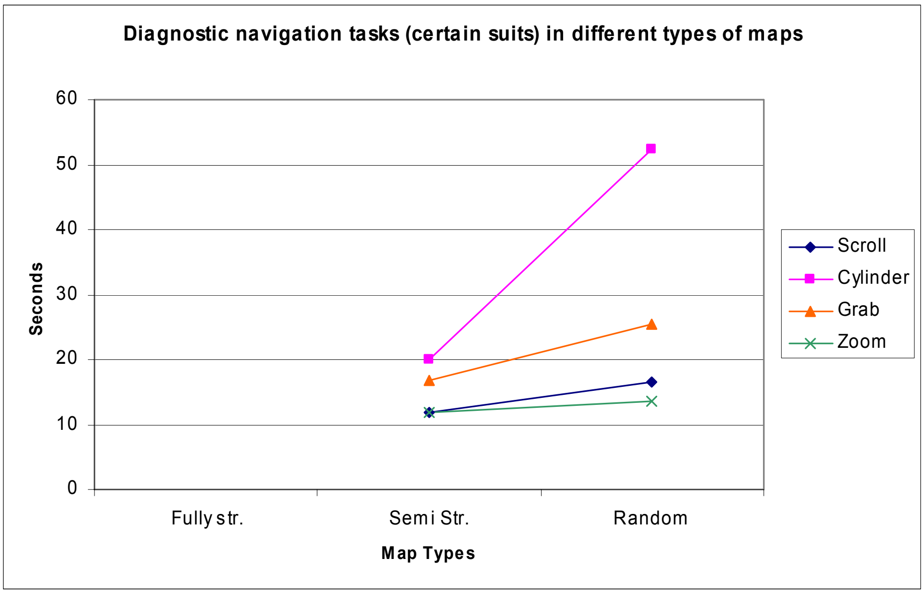


Figure 28: Experiment 1: Average task times for counting cards of a certain suit in different types of maps, using the four navigation techniques (scroll, cylinder, grab and zoom).

Figure 28 shows the average times for tasks asking the subjects to count cards of a certain suit, such as: How many hearts are there on the map? These were diagnostic tasks which again required the subject to scan the entire body of information.

As solving this kind of tasks was trivial in the fully structured maps, they were only conducted in semi structured and random maps.

Since the icons representing the suits stayed clearly visible when zooming out, the strength of viewing the whole map is clear in the results from the zoom NT.

The cylinder technique again proves to be an inefficient way of solving diagnostic tasks.

Discussion and findings

This first experiment was conducted in order to explore the field of navigation techniques, especially in relation to pan and zoom. The findings and the general observations from this study have been essential in the design of our second round of experimenting in which gaze and mouse interaction are compared.

The zooming in this test was implemented semantically in an attempt to isolate zoom movement as a navigation technique. It was no surprise that the limited benefit of the overview and the difficulty of a systematic search through the map without a pan option clearly shows in the results.

The idea of implementing the zoom NT semantically was intended to work against the approach of zooming out to an appropriate scale and answering all questions from there. Our experiences from this first experiment led to an attempt of designing the information space for the second round of experiments in a way that forced subjects to zoom in and out.

Comparing the results for the two panning options; edge-scrolling and grabbing, indicates that the differences between them were insignificant. This confirmed our intention of pairing them against each other in a potential gaze/mouse test, with the edge-scroll as a candidate for gaze-controlled panning. The subjects showed mixed preference; some preferred grabbing and others preferred scrolling.

The fact that the information in this test was laid out in a single plane and ordered in rows and columns had a great bearing on the navigation strategies employed by the subjects. This emphasized the need for considering the mutual influence between data, visualization and navigation techniques.

Some of the tasks in the experiment required the subjects to look through the entire dataset in order to solve them. This type of task is well known from real life scenarios (e.g. Plaisant's definition of *diagnostics*). However, we noted that the results from these tasks were greatly dependent on the individual subject's level of self-confidence. Some simply examined the data once and were quite sure in their answer, while some went through the data 2 or 3 times before answering.

In summation the first experiment revealed interesting facts about three distinct areas; zooming, panning and information visualization, all observations which were instrumental in designing the second experiment.

3.2 The context of the second experiment

Inspiration

The main focus of our second experiment was examining pan and zoom navigation (PZN) in gaze-controlled interaction and comparing it with an equivalent mouse-controlled navigation technique (NT).

Inspired by the initial test round we wanted to explore to a greater extent the possibility of applying a sense of depth in a zooming interface, in order to exploit the possibilities of zooming movements.

When designing the environment and navigation techniques for our second experiment, Aza Raskin's implementation of a zoomable interface was a great source of inspiration. Here the information is presented in different scales, giving the sense of multiple layers and facilitating exploration and discovery as the user interacts with the space. [75]

Below is an attempt to visualize the zooming experience of this particular design, which uses continuous geometric zoom in a multiscaled information space. The red circles represent the points which are explored on the subsequent images.

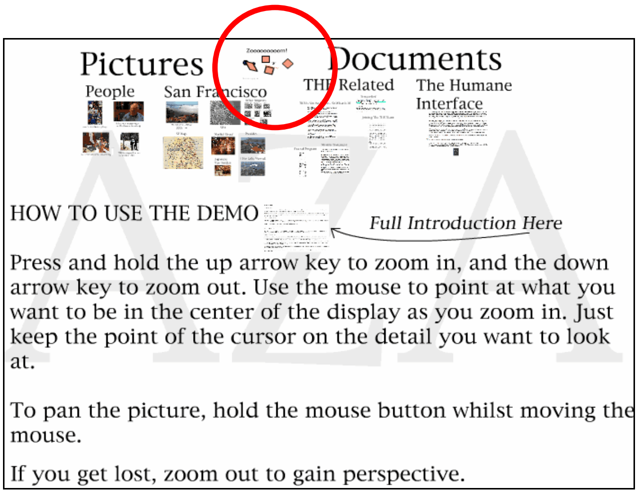


Figure 29a: The initial view of Aza Raskin's zoom demo.

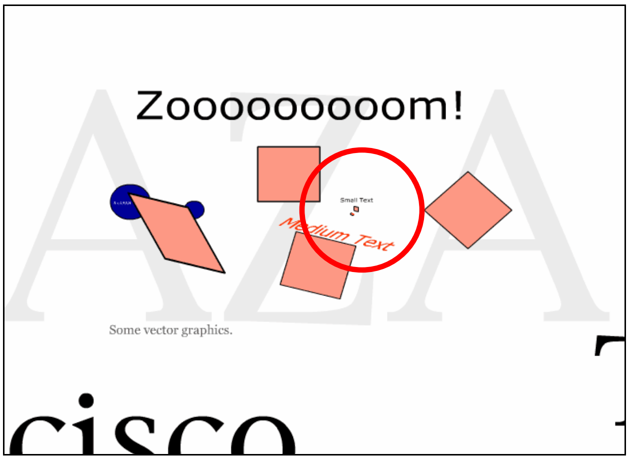


Figure 29b: Zooming in on an area of interest; more details of new information become apparent.

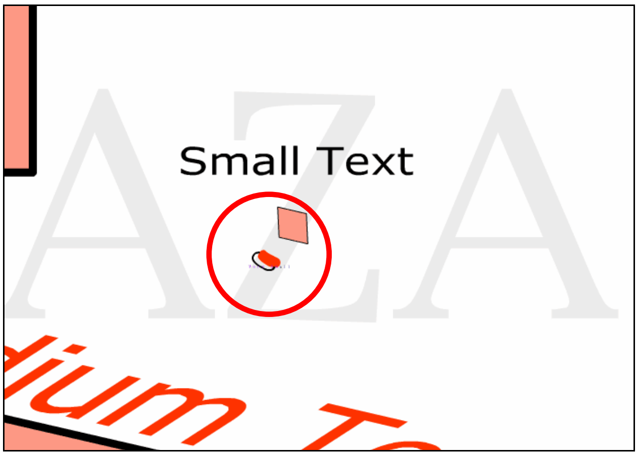


Figure 29c: Zooming further into the information space.

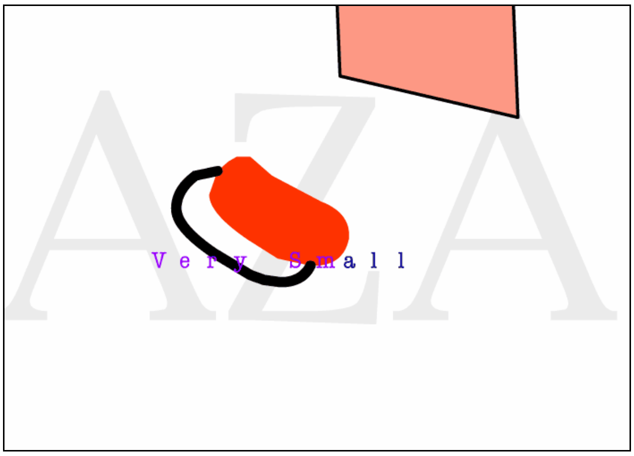


Figure 29d: The user reaches the lowest level of information detail in this particular area.

Our goal was to create a similar experience of exploration in designing our information space and implementation of navigation techniques.

---

[75] http://rchi.raskincenter.org/demos/zoomdemo.swf

Information space and tasks

The idea of abstract content prevailed, because it allowed the focus to be on navigation rather than the specifics of solving a “real-world” task with explicit content (e.g. biological data or other specific graphical data).

“Lego bricks” possessed the desired qualities that allowed us to work with highly visual multidimensional nodes. Each brick can be defined by three different dimensions of size; height (half or one third), width and length. They can be of various forms (e.g. slanted or curved) and facilitate both examination that can be done at a glance (e.g. noting colours) and requiring closer inspection (e.g. counting the pegs of large bricks). Adding different colours gives the possibility of designing an information space that is both familiar and complex.

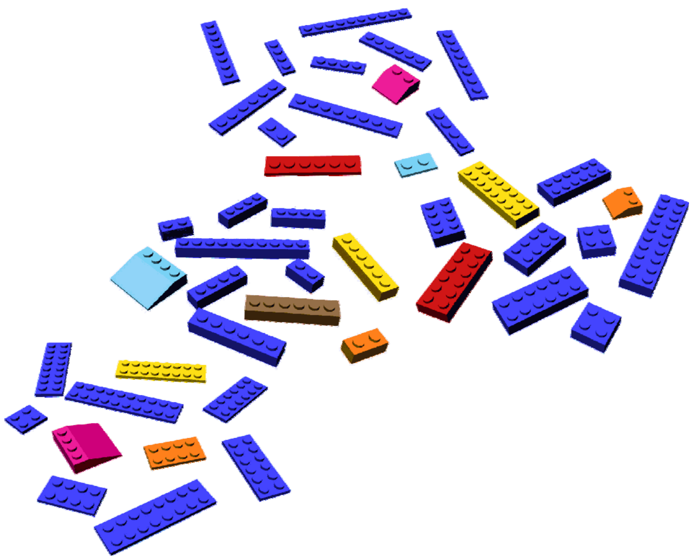


Figure 30: An example of the ”Lego bricks” we use as data in our second experiment, demonstrating the variety this dataset allows.

Inspired by our experiences from the first experiment, we again wanted the content and its structure to require as little analysis as possible. The gestalt grouping principles of colour, proximity and size were therefore used when designing the maps and subsequently the tasks. This resulted in tasks that only required a low level of cognition, allowing us to focus on interaction rather than comprehension.

To fully explore the potential of pan and zoom navigation, we tried to structure the information space so that subjects were required to move in the third dimension of depth when solving the tasks.

In the initial study the diagnostics tasks, requiring the subjects to search through the entire data content, turned out to be more dependent on the individuals' confidence level than the navigation techniques tested. Therefore we made the decision to move away from that type of tasks, replacing them with what we call *visual inspection tasks*.

As with the diagnostics tasks the subjects are required to look for patterns and nodes of relevance (in this case a sequence of specified bricks). However, unlike the diagnostics task, the visual inspection task only requires the collection of enough data of a certain criteria, rather than all data.

Pilot study

Before conducting our second experiment, we tested our implementation and tasks on a few subjects in a pilot study. The subjects of this pilot test attempted to solve some tasks using gaze interaction with and without a cursor indicator of the point of interest. The conclusion was to totally remove the cursor as it distracted the subjects from the underlying elements and revealed even the slightest miscalibration of the eye-tracker, disturbing the subjects. The same observations have been made in other studies of gaze interaction.

> ”Subjects explained that they were particularly annoyed when they noticed that the pointer symbol failed to follow their fixation points accurately.” [76]

The pilot study also revealed that subjects were quickly able to gain perspective of the whole map, so the map was enlarged several times to provide a more challenging navigation experience. Eventually the decision was made to make one huge map in which all tasks, in all NT, by all subjects were solved (see Figure 31).

The advantage of a single large map was that the tasks could be more complex and having all subjects solve tasks in the same information space, removed the variable of map design as a factor in the results. The disadvantage of a single map was a possible learning effect as subjects became more familiar with the content, but we hoped the size of the map (with approximately 2000 bricks) would counteract this. In fact, the results of the second experiment indicate that there was no noticeable learning effect.

The pilot study also demonstrated the subjects' preference for trying to solve the tasks at the surface of our map, that is; rather than explore the depth of the map they tried zooming out and

[76] Hansen, John P. et al. (2003) p. 6

finding the requested bricks in the topmost layer. As a consequence we thoroughly redesigned our map, ensuring that of the requested bricks, no more than half could be found without zooming into the lower layers. (The construction of our map is described in more detail in the next chapter.)

Since most of the tasks were based on colour recognition it was clearly necessary to ensure that none of our subjects had trouble distinguishing colours – so a colour blindness test was devised.

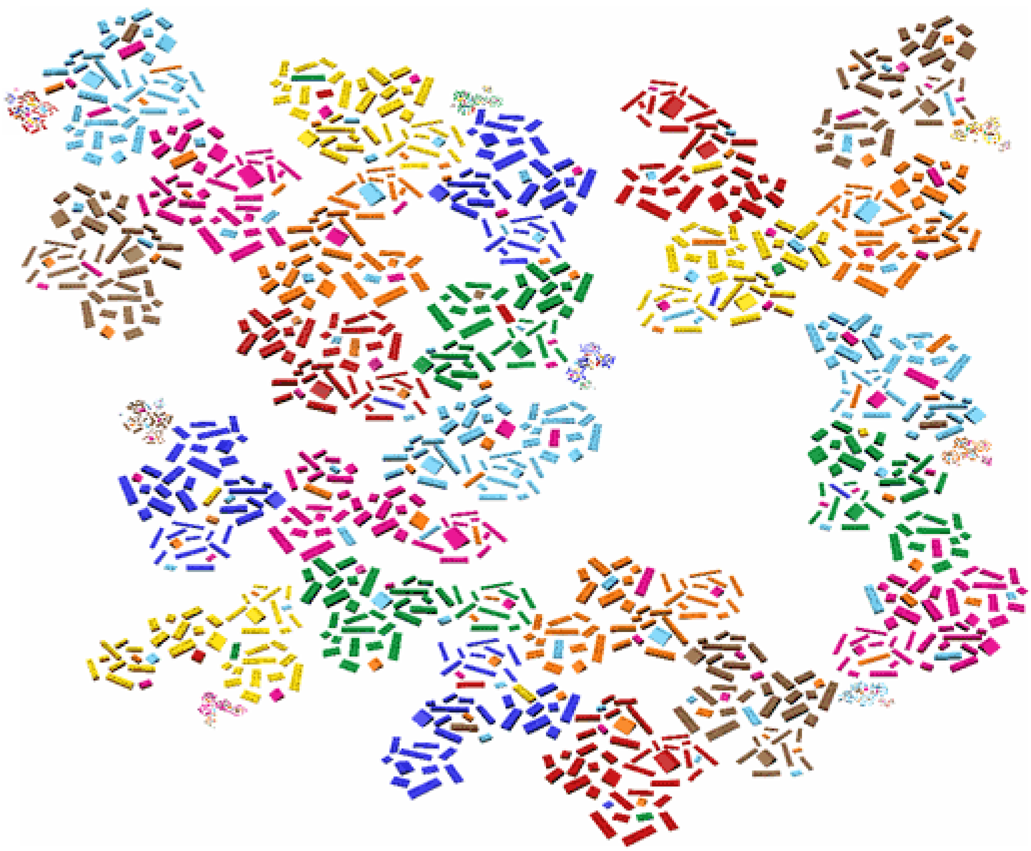


Figure 31: An overview of the whole map used in the second experiment, containing approximately 2,000 bricks in eight colours.

3.3 Test design for the second experiment

In our second round of experiments we tested four navigation techniques; two mouse-controlled and two gaze-controlled, with the intention of evaluating them pair wise to draw conclusions about the potential of gaze-controlled pan and zoom navigation.

Tasks and dataset

As previously discussed, we chose to conduct our testing using an abstract visual dataset of groups of "Lego bricks" in two different types of tasks:

1) Find-tasks (Visual inspection tasks):

The majority of the tasks in the test were what we choose to call "find-tasks", where subjects were asked to find seven bricks according to a given criteria, such as 7 yellow bricks with two pegs. The find-tasks were conducted on the large map with brick groups in primarily one colour but with some "misplaced" bricks of other colours also present in the group.

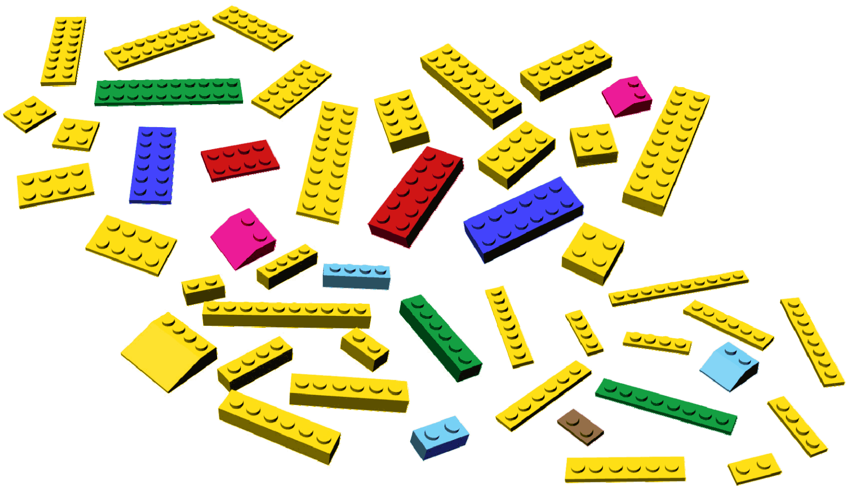


Figure 32: An example of a colour group, predominately yellow. This group contains three yellow bricks with two pegs, two misplaced red bricks, three misplaced green bricks, etc.

There were two kinds of find-tasks; finding 7 bricks of a certain colour with 2 pegs, and finding 7 misplaced bricks of a certain colour (see Figure 32). The bricks were of eight colours, whereas the tasks only involved four of the colours – the other four colours being used as noise. To avoid ambiguity all diagonally slanted bricks with two pegs were coloured in one of the neutral noise colours.

Besides the groups in our primary layer there were also groups on three "lower" layers, each layer approximately $1/4^{th}$ of the scale of the layer above. As a navigational aid, all groups in the lower layers were gathered in 8 clusters on the map, each cluster containing three groups in different layers.

As in many zoomable interfaces our definition of a 100% scale is arbitrary, but for the purposes of this paper we have chosen to designate the initial zoom level as 100%. Subjects started each task at the same location (and scale) of the map and navigated from there. By zooming out to a scale of approximately 25% the whole map became visible, and by zooming beyond 100% the “deeper” groups could be explored. Figures 33a to 33d give an indication of how the details of the map are layered; each image shows a further zoom into the centre of the previous image.

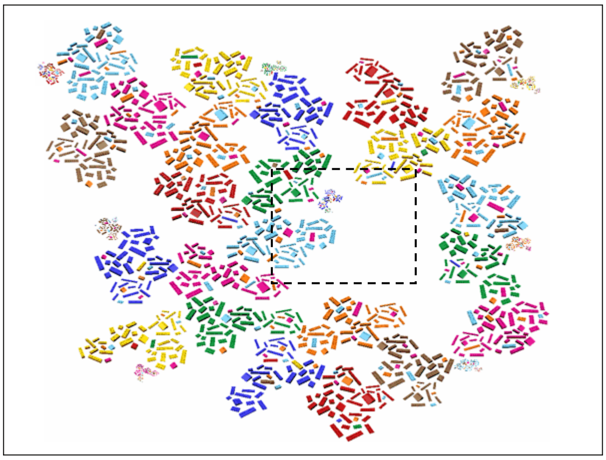


Figure 33a: Screenshot showing an overview of the whole map (25%)

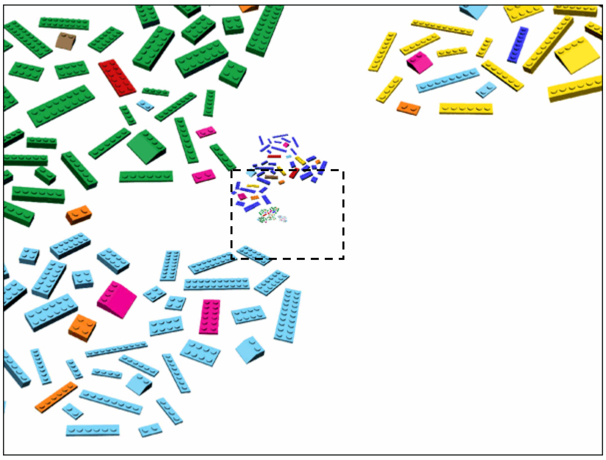


Figure 33b: The area marked in Figure Xa shown at initial zoom (100%)

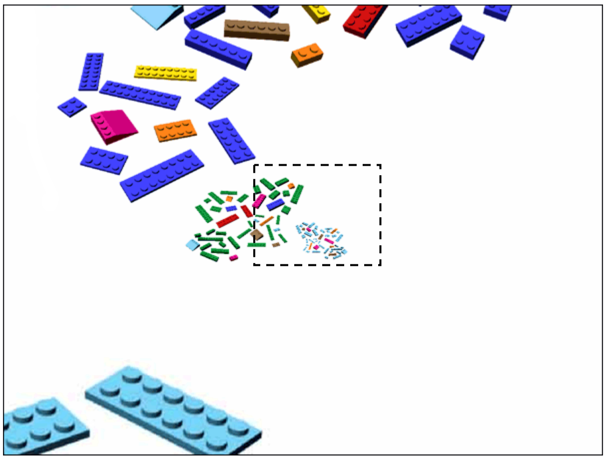


Figure 33c: A cluster at zoom level approx. 500%

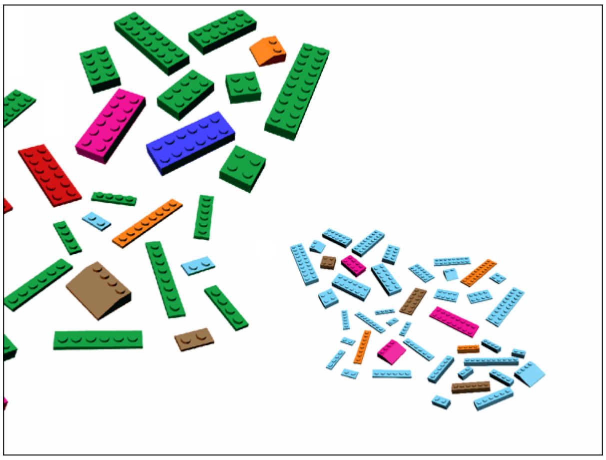


Figure 33d: A cluster at zoom level approx. 2400%

The map was designed to ensure that no task could be completed by only exploring the primary layer, exploration of at least one lower cluster was necessary. For each task there were 13 bricks of the requested type on the map, but subjects were asked to find 7 of them.

Subjects counted out loud the number of bricks they found and stopped the task themselves after finding 7 bricks, by pressing the spacebar. As discussed later, there was no formal registration of errors, but informal notes were taken by a test observer.

2) Target-tasks (Target exploration):

The other type of tasks we chose to call target-tasks. These tasks were designed to measure precision and efficiency without the element of exploration. Subjects were asked a question, which required them to take two steps, for instance: "How many misplaced bricks are in colour 1?" The subject then had to zoom in at the tip of a numbered arrow (in this case number 1) and find a coloured brick, after which to zoom in at the tip of the matching coloured arrow to find the group containing the answer.

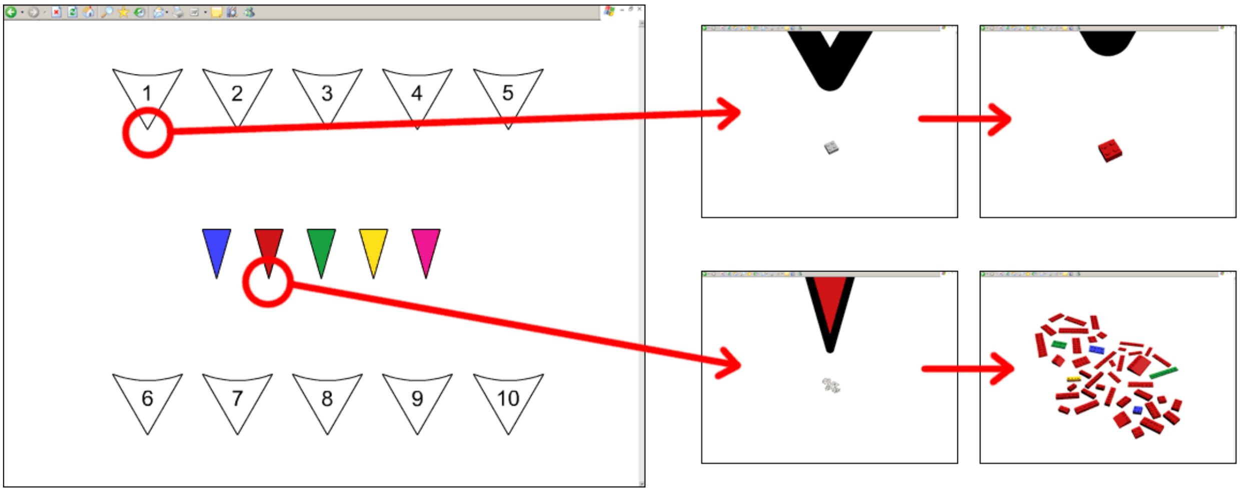


Figure 34: Schematic representation of a target task: At the point of arrow number 1 a grey brick reveals it colour when zooming past 20,000% magnification (in this case red). At the bottom of the correspondingly coloured arrow a blurred group of bricks reveals its details when zooming past 20,000% - allowing the subject to count the requested type of bricks.

In effect the target-tasks consisted of three steps:

1. Zooming in on a specific location to a magnification of approx. 20,000%
2. Zooming back out to approx. 100% in order to gain an overview
3. Zooming in on another location, based on the information found in step 1, to a magnification of approx. 20,000% and answering the posed question.

Figure 34 gives an overview of the steps involved in solving such a "target-task". The tasks either involved finding misplacements or counting the number of slanted bricks in a group. (For instance, in the bottom-right image in Figure 34 there are 5 misplaced bricks.) The subjects were asked to give verbal answers and then to stop the time by pressing the spacebar.

Types of mistakes encountered

As in any form of interaction, watching our subjects navigate within the map there are various types of mistakes that they can make:

- Misunderstanding the task at hand. This was counteracted by allowing the subjects to solve some practice-tasks and if a subject e.g. forgot what colour to look for the moderator helped restart the task.
- Getting lost. The subject may loose track of what areas of the map she has explored and which not. This may result in confusion or in the recounting of targets.
- Overlooking clues. The subject may not pay attention to a group of bricks or choose not to explore a group further, although there are indications that some of the requested targets might be found there.
- Overlooking targets. The subject may not notice bricks that fulfil the requested criteria even though they are visible onscreen.
- Recounting targets. Since there is no visible indication of which bricks have been counted, there is the chance that a subject may count the same brick twice without noticing.

Of these types of possible mistakes we feel that only the last one can bee seen as a direct error in the context of our tasks. As the subject is only asked to find about half of the bricks that are present on the map fulfilling each criteria, overlooking targets leads to less effectiveness in solving the task but does not constitute a wrong answer.

Since there is no direct selection of targets (e.g. by clicking on them), neither the testing application nor the subjects have any inconclusive way of determining which bricks have been counted and which have not, so the subjects have to use the location of groups to memorize which have been explored and which not. Although this is typical for many real-life tasks involving passive exploring it is clearly problematic when it comes to logging errors.

Keeping track of error-rates is a key element in all interaction research, but in the case of our test-setup, to non-ambiguously register all instances of recounting would have required us to combine:

- Recordings of the visible bricks, both in terms of location and scale.
- Registrations of the gaze of our subject.
- An audio recording of the subject counting out loud.

Although all of these are technically feasible, the reviewing of all factors combined would have been far too labour-extensive for the timeframe of our project. Therefore we chose to limit any registering of mistakes to written test notes.

Navigation techniques

We tested four navigation techniques; two mouse-controlled (NT1 & NT2) and two gaze-controlled (NT2 & NT3).

Aza Raskin's design was not only an inspiration in the design of the information space, but also served as the foundation for our view of a well-implemented pan and zoom navigation. It retains the advantage of perspective in a continuous zoom, controlled by the keyboard, and allows panning by grabbing and dragging the information space. This navigation technique (NT2: Mouse; pan+zoom) was deemed the *mouse-champion* as it represented an effective way of using mouse interaction in a zoomable interface. The reason we chose this and not a mouse navigation technique with edge-scrolling, was that we found it interesting to test gaze against an existing implementation.

For comparison we chose a more common PZN such as the one implemented in Adobe Photoshop and Adobe Acrobat Reader, where the user is required to select a tool and change either the mode or at least sustain a quasi-mode (e.g. Jef Raskin's discussion), by either selecting the pan, zoom-in or zoom-out option from a toolbar, accessing it from a sub-menu or familiarizing oneself with the shortcuts. This type of navigation technique (NT1: Mouse; toolbar) constituted our other mouse control. Because we tried to emulate how first time users would interact with the system, we chose a version were the subject selects the modes onscreen and not to test a modal pan/zoom navigation technique based on keyboard shortcuts.

For the gaze-controlled NTs the goal was to create a PZN which could rival the *mouse-champion*. The results from our first test showed that the difference in task times between the edge scrolling and grabbing with a mouse, were insignificant. The original intention of implementing the edge scroll as the panning option was upheld, thereby creating the *gaze-challenger* navigation technique (NT4: Gaze; pan+zoom), using the zoom behaviour of the *mouse-champion* and replacing grab with edge scroll for panning. This could also have been tested with stepwise zoom, to establish the preference of users. However, we relied on claims that a continuous zoom was superior because of the absence of potential discomfort.[77]

Results from the first experiment seemed to expose a shortcoming of zoom as the only means of navigation when systematically exploring a planar interface. A reason for this is the difficulty in exploring the immediate context of a target; which is where panning appears most useful. As an attempt of measuring the importance of panning, the second gaze navigation technique was a zoom implementation without a panning option (NT3: Gaze; zoom), otherwise identical to NT4. In both gaze-techniques the zooming mechanism is the same as in NT2.

We are effectively dealing with three dimensions; input devices (mouse/gaze), type of zoom and type of pan. The combinations of these dimensions that we chose to implement in our tests are displayed in the following table:

	Mouse		**Gaze**	
	Zoom:		*Zoom:*	
Pan:	Stepwise	Continuous	Stepwise	Continuous
Grab	**NT1**	**NT2**		
No pan				**NT3**
Edge scroll				**NT4**

Figure 35: The combinations of input device, pan and zoom techniques making up the four navigation techniques NT1 to NT4 in the second experiment.

Since there are many variables in our navigation techniques, designing a test that isolates each one would have been far too ambitious for our limited time frame. Therefore we chose to compare them pair wise; the two mouse NTs to support our selection of a *mouse-champion*, the two gaze NTs to measure the importance of pan, and finally the *mouse-champion* with our *gaze-challenger* to evaluate the potential of gaze controlled PZN.

Zoom:

All our zoom interactions are based on Flash ActionScript code by Aza Raskin. In his original implementation zooming in and out is done by pressing keys on the keyboard as we do in three of our navigation techniques. Each registered key press results in a magnification of 107% (or 93%

[77] Plaisant, C. et al. (1995)

when zooming out). As our application ran at 45 frames per second, zooming continuously in for one second resulted in a 2100% magnification ($1.07^{45} = 21.00$).

In Raskin's implementation the map is slightly shifted in relation to the position of the cursor when zooming in or out; if the cursor is held over an object on the right hand side when zooming in, the whole map is shifted to the left resulting in the object moving towards the centre. This behaviour helps users zooming in on any object of interest visible on the screen without having to adjust the zoom by panning. If the cursor is held in the centre of the screen while zooming no shift occurs. Zooming out this effect is reversed, so that by holding the cursor to the right of an object it moves "back" to the right. This behaviour surprised some of our subjects, but with the exception of the NT3 (gaze; zoom) navigation technique they could easily compensate by panning when needed.

Subjects had a choice of either using the up and down arrows on the keyboard or 'W-key' and 'S-key' for zooming. All subjects were right-handed and all but two chose to use W+S after trying out both options. Our choice of W+S was based on the classic WASD key configuration of controls in computer games, but this turned out to be flawed in that gamers usually use the same finger for moving forwards (W) or backwards (S) while our subjects kept two fingers on the keys, resulting in a slightly awkward positioning. Using for example W and D would probably have been more comfortable. In the navigation technique NT1 (mouse; toolbar) we used the same code but modified it for stepwise zooming.

Pan:

In our mouse based NT panning was done by mouse drags, resulting in a 1:1 panning; that is moving the cursor *n* pixels resulted in the map also moving *n* pixels.

In our navigation technique NT4 (gaze; pan+zoom) we used "edge-scroll" for panning. In our implementation of edge scroll, there were two horizontal scrolling areas on each side (1/4 of the screen width) and similarly two vertical scrolling areas at the top and bottom of the screen.

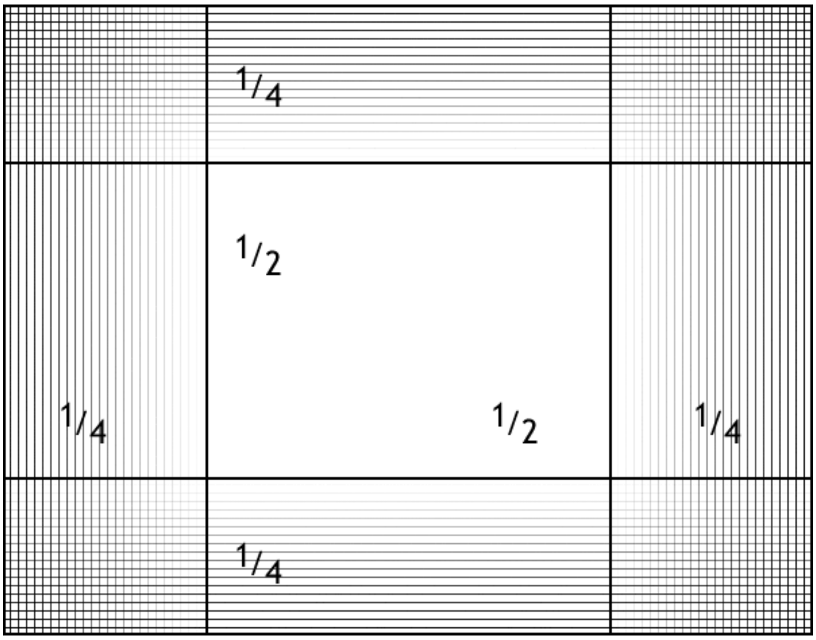


Figure 36: Scrolling areas in edge-scrolling; moving the cursor into the scrolling areas results in a panning motion with the speed related to the cursor's distance from the centre.

When the (invisible) cursor enters the scrolling areas the map shifts horizontally and/or vertically. The panning speed was a square function of the distance from the centre, for horizontal scrolling ranging from 180 pixels/sec to a maximum speed of 740 pixels/sec. and for vertical scrolling the range was from 140 pixels/sec to 560 pixels/sec. (The difference lies in the 5:4 format of the screen, although partially compensated for in our formulas.) Horizontal and vertical panning were independent of each other, thereby enabling diagonal panning.

Edge scroll is well suited for gaze interaction, as gazing at an object at the right of the screen shifts the object towards the centre where the scrolling then stops and the object can be studied as long as it remains the area of interest and the gaze is held within the centre area. This behaviour is noticeably effective while zooming, but the downside is that glancing at multiple objects may cause unintended movements when objects are near the edges of the screen. Another downside is that any additional information, such as an added overview, would be problematic to place near the edges of the screen because the process of orientation and the action of navigation would coincide.

In our implementation there were no visual indicators of where the scrolling areas were located; subjects were simply told that there were invisible scrolling areas near the edges.

NT1 (mouse; toolbar):

Our first mouse-controlled navigation technique was based on the well known pan/zoom navigation (PZN) approach of using a magnifying-glass-tool for zooming and a hand-tool for mouse dragging. In our implementation there were separate tools for zooming out and in, with no keyboard shortcuts available to the subjects for switching between them.

The toolbar buttons were located in the upper left corner of the application, with each button measuring 75x75 pixels. The active mode was displayed by changing the mouse cursor to a cursor representation of the current tool (20x20 pixels).

Figure 37: The initial view using NT1. The toolbar is located in the upper left corner.

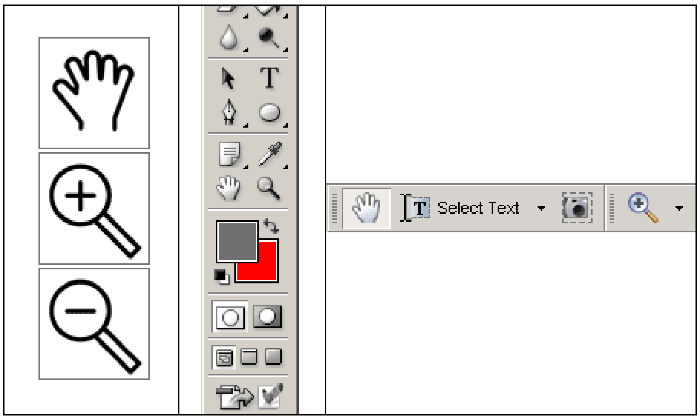


Figure 38: The toolbar used in NT1 compared for size with the toolbars from Adobe Photoshop (middle) and Adobe Acrobat (left).

Each click with the magnifying tool enlarged the underlying map by a factor of 170% when zooming in and by a factor of 60% when zooming out. Dragging with the hand tool resulted in a 1:1 panning.

NT2 (mouse; pan+zoom):

Our other mouse navigation technique was based on the code we received from Aza Raskin. Here zooming in and out was performed by pressing one of two keys on the keyboard (W or S). While zooming the map can be shifted by moving the mouse as previously described. Grabbing the map was simply done by using the mouse button, as no changing of mode was necessary, resulting in a 1:1 panning.

NT3 (gaze; zoom):

The zooming in both our gaze-controlled navigation techniques was exactly the same as in NT2 (mouse; pan+zoom), the only difference being that the location of the “mouse cursor” was controlled by gaze. In NT3 subjects had no panning option, so planar movement was only achieved by zooming out to gain perspective and then zooming back in on the intended target of interest.

After some experimenting we removed all visual representations of the cursor in our gaze-controlled implementations as it distracted our pilot subjects from the underlying elements.

NT4 (gaze; pan+zoom):
This navigation technique was the same as NT3 except for the added panning option in the form of edge-scroll, as previously described.

The starting view in NT2, NT3 and NT4 was practically the same as shown in Figure 37, with the toolbar removed. In the gaze-controlled navigation techniques the pointer symbol was hidden, while NT2 used a standard arrow pointer.

Hypotheses

Based on our previous discussion about the drawbacks of modes (Chapter 2.4) and the time and effort needed to select a new tool (changing the interaction mode) we assumed that the navigation technique NT1 (mouse; toolbar) would perform worse than the more fluid NT2 (mouse; pan+zoom).

When comparing NT3 (gaze; zoom) and NT4 (gaze; pan+zoom) we expected the performance to suffer when removing the panning option. A preliminary guess was that the difference would be in the order of 20-40%.

The most interesting comparison would be between mouse-controlled PZN implementation (NT2) and our adaptation for a gaze PZN (NT4). Based on the familiarity all our test subjects had with mouse usage we expected the *mouse-champignon* to perform somewhat better than *gaze-challenger*.

Therefore our hypotheses were as follows:

> **H1:** The NT2 (mouse; pan+zoom) interaction technique will perform markedly better than NT1 (mouse; toolbar).
>
> **H2:** Removing the panning will result in NT3 (gaze; zoom) performing worse than NT4 (gaze; pan+zoom).
>
> **H3:** Familiarity with mouse control will result in NT2 (mouse; pan+zoom) performing better than NT4 (gaze; pan+zoom).

Apparatus

We conducted the study on a system running Windows XP with a Pentium 4, 3.6 GHz processor, 1GB RAM, a 256 MB Raedon X600 video card and a 17 inch Tobii 1750 eye-tracking monitor configured at a resolution of 1280x1024 pixels. The Tobii 1750 eye-tracker has a 50Hz sample rate. The application MyTobii 2.0.0 was used to allow controlling the system pointer with gaze, it was set at mouse speed 95% and at fixation sensitivity 95%. The test itself was developed using Macromedia Flash 8 and ran in a Flash Player 8 within an Internet Explorer browser.

Design

The test took approximately 50 minutes to complete and consisted of a general introduction and 4 blocks of tasks. In each block the subject was introduced to one navigation technique and after a few practice tasks solved 6 find-tasks and 2 target-tasks using this navigation technique.

Each subject was allocated one of four test sequences (with four subjects completing each sequence). The test sequences determined in which order the subjects experienced each NT and was used to lessen any carry-over (or learning) effect. For instance; out of the 16 subjects, 4 had NT1 in their first block, 4 had NT1 in their second block, etc.

For efficiency the gaze-controlled navigation techniques were grouped together either in the first two or last two blocks, so that the eye-tracker only had to be calibrated once for each subject.

Procedure

16 subjects (5 females) between 23 and 47 years of age voluntarily participated in the study. All subjects were right handed and all males were tested for colour blindness.[78] Three of our subjects had some experience of gaze interaction, others none.

Our male subjects took a short colour blindness test, a classic dotted number representation test(see appendix F), at the start of the experiment, which they all passed. After a brief introduction to the test, subjects were introduced to their first navigation technique.

[78] The percentage of women suffering from colour blindness is less than 0.5%.
http://www.healthatoz.com/healthatoz/Atoz/ency/color_blindness.jsp

In each block there was an introductory session before conducting any real tasks. In this session subjects got a presentation of the navigation technique, followed by two practice tasks to familiarize themselves. At the start of the test the moderator also used this introductory session to explain the layout of the map.

Subjects then got 6 find-tasks; first three tasks of finding 7 two pegged bricks of a certain colour, followed by three tasks of finding 7 misplaced bricks of a certain colour (the colours being rotated between navigation techniques).

After reading the instructions for each task, the subjects started the task by dwelling over an activation-button for two seconds and upon completion of the task stopped the time counter by pressing the spacebar.

Subsequent to completing the find-tasks there were two target-tasks. In the first block, subjects got two practice tasks, introducing them to this type of tasks. In subsequent blocks subjects went directly to the real target-tasks.

Upon completing all 8 tasks in each block, subjects were asked to evaluate the navigation technique on three parameters; general comfort, perceived speed and how confident they felt in their given answers. At the end of the test, after experiencing all navigation techniques, subjects compared their evaluations and modified if needed. (For full script, see appendix E)

Measures

For all tasks performance measures were based on automatically logged timing. As previously discussed we did not collect accurate error data, nor did we make any attempt at logging how subjects navigated the maps (besides our own test notes).

The reason for not collecting more data was primarily our emphasis on navigation rather than our subjects' precision and accuracy.

4 Empirical results

As described in Chapter 3.3, the subjects were asked to perform two types of tasks; find-tasks of the types “find 7 bricks of a specific colour with 2 pegs” or “find 7 misplaced bricks of a specific colour” and target-tasks involving finding a answer by zooming in on two known locations.

In the find-tasks a certain element of chance was involved, as subjects in some cases “struck gold” and found the requested bricks early but in other cases had to spend a long time moving around on the map. In all there were 6 find-tasks for each navigation technique and in order to limit the variations due to chance, we used the combined times for the six tasks for each subject/technique combination in our calculations.

There were 2 target-tasks which were mainly aimed at measuring direct navigational speeds; in theory eliminating any element of chance. We used the combined times for the two tasks for each subject/technique combination when dealing with target-tasks.

As previously described we compared four different navigation techniques; NT1 (mouse; toolbar), NT2 (mouse; pan+zoom), NT3 (gaze; zoom) and NT4 (gaze; pan+zoom). In the following chapters we will first present the general results and then proceed to comparing the pairs of navigation techniques discussed in Chapter 3.3.

4.1 General results

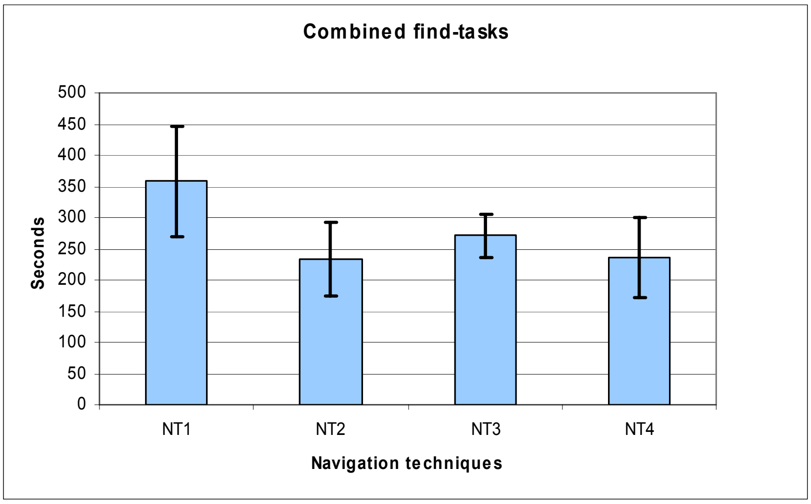
All our time measurements were automatically logged in our Flash application and then imported into Microsoft Excel for analysis. The testing environment registered the time passed between the subject starting and stopping each task, with a precision of 1/10th of a second.

Unless otherwise stated we use standard deviation as our indicator of variance. In our pair wise comparison of navigation techniques we use a paired t-test for evaluating the relations between the data collected for the two techniques.

All navigation techniques: Find-tasks

When the performances of all our 16 test subjects and all the 4 navigation techniques are compared, the time taken to solve the 3 “find 7 bricks with 2 pegs…” tasks in each NT is on average 145 seconds, and similarly solving the 3 “find 7 misplaced bricks…” takes on average 129 seconds. This results in an average of 274 seconds for solving all find-tasks in each NT, what we call a *"combined find-tasks time"* – or in other words the combined time taken to solve the 6 find-tasks for each navigation technique.

When we plot the combined find-task times for each of the four navigation techniques the results are as follow:



	NT1	NT2	NT3	NT4
Average find-tasks time	358 (88)	233 (59)	272 (35)	235 (65)

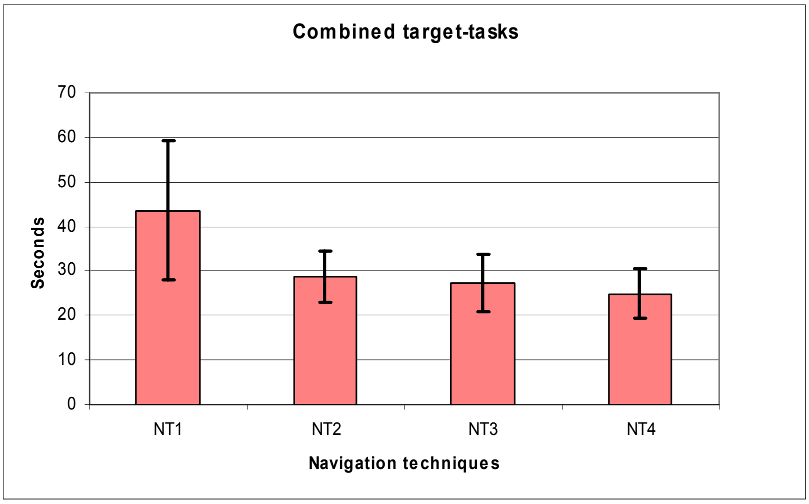
Figure 39: The average times taken to solve all the find-tasks in each of the four navigation techniques. Times are given in seconds, with standard deviation in parentheses.

As evident from these results, the NT1 (mouse; toolbar) results in the longest task times, with NT2 (mouse; pan+zoom) and NT4 (gaze; pan+zoom) practically neck-to-neck.

The results for each navigation technique will be discussed in more detail later in this paper.

All navigation techniques: Target-tasks

In a similar fashion we compared the times taken to solve the two target-tasks for each navigation technique. These tasks were created to test zooming in and out between two points of interest and give an indication of the speeds attainable with each navigation technique, eliminating the effect of strategy and chance.



	NT1	NT2	NT3	NT4
Average target-tasks time	44 (16)	29 (6)	27 (7)	25 (6)

Figure 40: The average times taken to solve the target-tasks for each of the four navigation techniques. Times are given in seconds, with standard deviation in parentheses.

Here NT1 (mouse; toolbar) again proves to be the worst performer, and the gaze-controlled NT3 (gaze; zoom) and NT4 (gaze; pan+zoom) appear to be slightly faster than NT2 (mouse; pan+zoom).

The results for the target-tasks will be discussed further when comparing the navigation techniques.

A possible carry-over effect

Since the find-tasks all take place in the same map it is to be expected that the subjects perform slightly better as they get a better feel for the layout of the map and are able to carry over their experiences from the previous tasks.

As described in Chapter 3.4, we tried to counteract any carry-over effect by using four different test sequences, rotating the order in which of the four test-blocks the subjects experienced the different navigation techniques. That is, out of the 16 subjects, 4 had NT1 in their first block, 4 had NT1 in their second block, etc.

If any carry-over effect is present it should be visible in results for each navigation technique improving when it is encountered late in the test, since when starting the fourth block subjects have already solved at least 18 similar tasks in the same map.

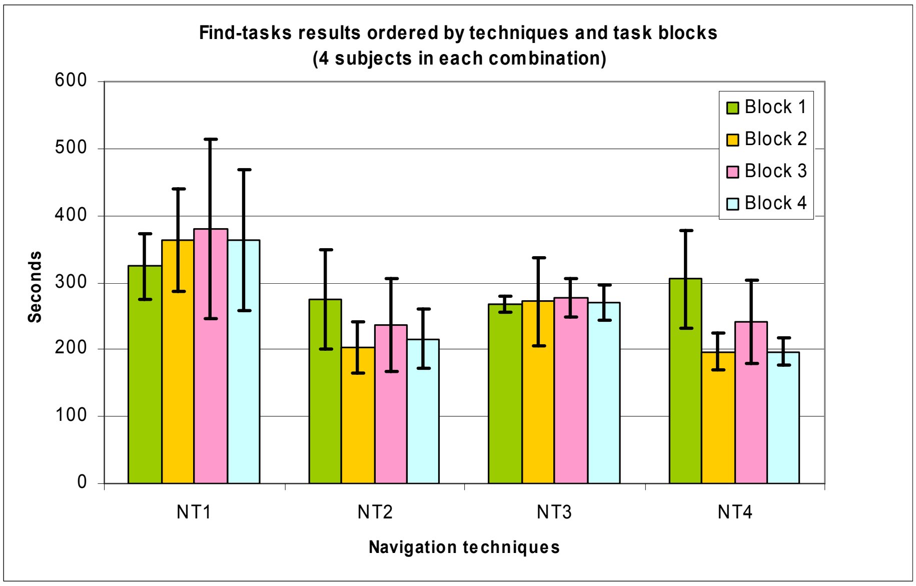


Figure 41: Looking for a carry-over effect in the find-tasks. The results are ordered by navigation techniques and task blocks, with 4 subjects in each combination.

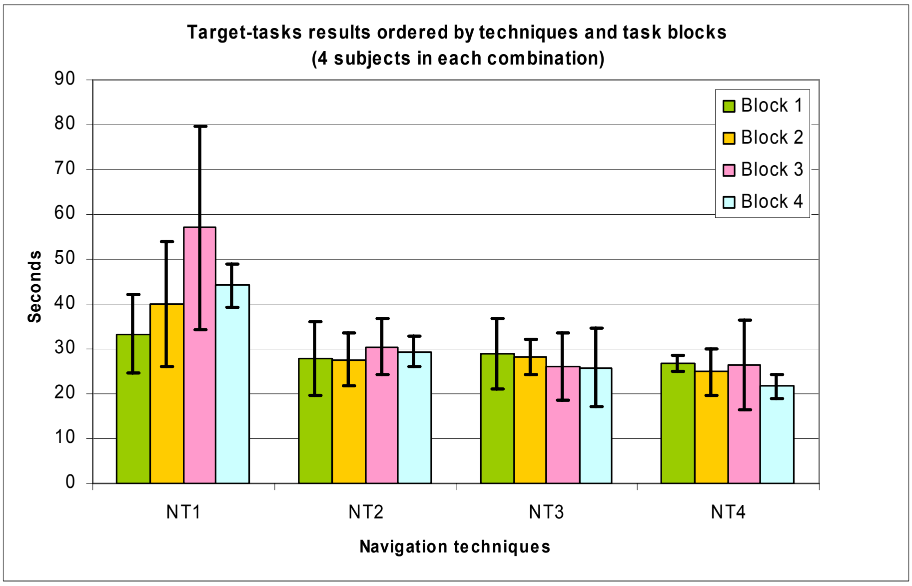


Figure 42: Looking for a carry-over effect in the target-tasks. The results are ordered by navigation techniques and task blocks, with 4 subjects in each combination.

In Figures 41 and 42 the combined find-task times for each navigation technique are divided between the 4 blocks in which they could be experienced. For instance the four columns furthest left show the task times for NT1 when encountered as the 1st-4th navigation technique in the test. A carry-over effect should present itself in a downward slope from the left to the right within each group of columns.

There are possibly hints of a downward slope for NT2 and NT4 in the find-tasks (Figure 41) and for NT3 and NT4 in the target-tasks (Figure 42). However, when taking the standard deviations into account it is difficult to find conclusive indications of a carry-over effect. As the data in each column only comes from 4 subjects, statistical fluctuations are more prominent than in a larger sample and the standard deviations are relatively large.

It is therefore our conclusion that any carry-over effect should have a negligible effect on the accumulated test results.

Error rates

As previously discussed we did not attempt any systematic logging of errors, but relied on test notes instead. Therefore the following observations should be taken with a grain of salt.

In the find-tasks the mistakes that can be classified as errors primarily consist of counting the same bricks twice. We can roughly estimate that this may have happened in approximately 5-8% of the find-tasks and these errors were not more likely to occur in the gaze-controlled techniques.

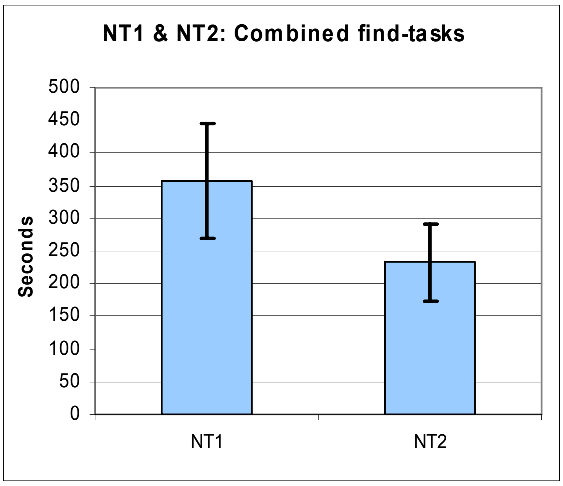
In the target-tasks, counting the bricks incorrectly, according to the requested criteria, would constitute an error, but we did not notice any instances of this.

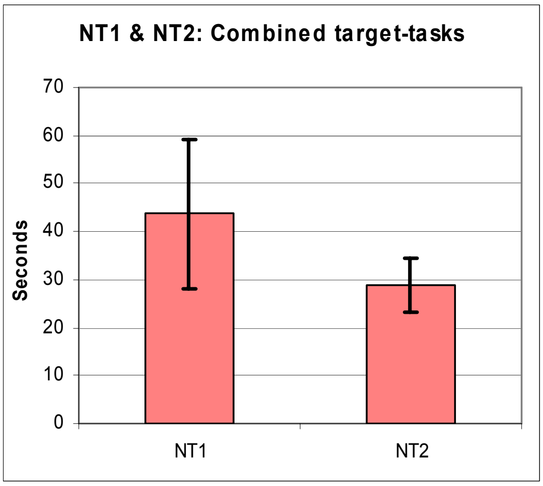
4.2 Comparisons

Comparison: Mouse-control (NT1 & NT2)

In our test we tested two mouse-controlled navigation techniques; NT1 (mouse; toolbar) and NT2 (mouse; pan+zoom). As described in Chapter 3.4 in NT1 subjects chose tools from a small toolbar containing a zoom-in-tool, zoom-out-tool and a hand-tool for panning. In NT2 the subjects used the W and S keys to zoom in and out at the location of the mouse, and could at any time “grab” the map for panning.

As we had expected, the tool-based NT1 turned out to be slow and inconvenient. Our subjects complained over the time it took them to switch between tools and frequently forgot what “mode” they were in and accidentally zoomed when they wanted to pan (or vice versa), even if the active tool was clearly indicated in the form of the cursor. This implementation is clearly problematic; requiring much cognitive reflection as a consequence of the explicit modes (Hick’s law) and the extensive amount of mouse movements for choosing a new tool (Fitts’ law).





	NT1 (mouse; toolbar)	**NT2** (mouse; pan+zoom)	**Δ**
Find-tasks	358 (88)	233 (59)	+54%
Target-tasks	44 (16)	29 (6)	+51%

Figure 43: Comparison of NT1 and NT2: The average times taken to solve all the tasks for each navigation technique. Times are given in seconds, with standard deviation in parentheses. The difference is calculated as percentage of NT2.

As the chart and table show, solving the tasks using NT1 took on average between 51 and 54% longer than when using NT2.

To verify that any difference in means between two navigation techniques is statistically valid (and not a result of random variations within the test data) we used a t-test (paired two sample for means) performed using Microsoft Excel. The key finding from our t-test analysis of the results for these two navigation techniques is that there is less than 0.001% chance that the observed difference in the find-tasks can be explained by chance alone ($p < 1 \cdot 10^{-5}$) and for the target-tasks $p < 5 \cdot 10^{-4}$.

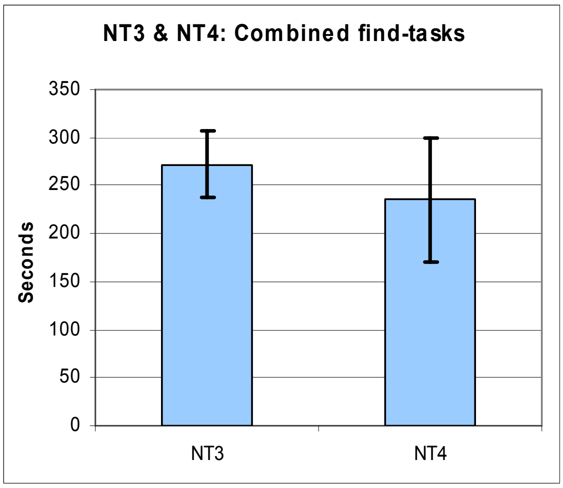
Using the data from the t-test analysis we can calculate the confidence interval for the true difference between the two techniques. According to our t-test analysis we can be 95% certain that the true difference between the two navigation techniques is (53.7 ± 18.6) % for the find-tasks, and (51.2 ± 24.3) % for the target-tasks.

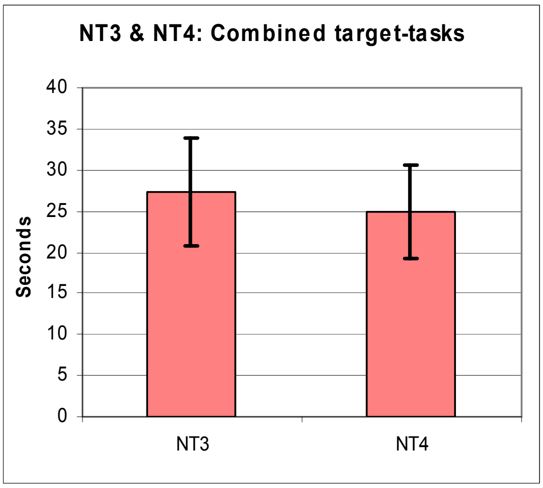
There is more data to be had from the t-test analysis, but we chose to focus on *p* (assuming a one tailed-test) and the confidence levels as our main indicators for the purposes of this paper. (The results of all our t-test comparisons can be seen in full in Appendix C.)

To summarize; in our test solving tasks using NT1 (mouse; toolbar) took about 51-54% longer than using NT2 (mouse; pan+zoom). This confirms our hypothesis H1.

Comparison: Gaze-control techniques (NT3 & NT4)

As previously discussed we chose to test two gaze-controlled navigation techniques; NT3 (gaze; zoom) and NT4 (gaze; pan+zoom), where the only difference is that NT4 has the option of panning by looking towards the edges of the monitor. In our first test round we had found indications that a zooming interface without the option for panning seemed troublesome, and our idea here was to try and measure how important the panning is for a gaze controlled pan/zoom navigation (PZN).





	NT3 (gaze; zoom)	**NT4** (gaze; pan+zoom)	**Δ**
Find-tasks	272 (35)	235 (65)	+15%
Target-tasks	27 (7)	25 (6)	+9.2%

Figure 44: Comparison of NT3 and NT4: The average times taken to solve all the tasks for each navigation technique. Times are given in seconds, with standard deviation in parentheses. The difference is calculated as percentage of NT4.

According to the t-test, $p = 0.02$ for the find-tasks and $p = 0.03$ for the target-tasks, so we can conclude that there is a distinct difference in the performances of the two navigation techniques.

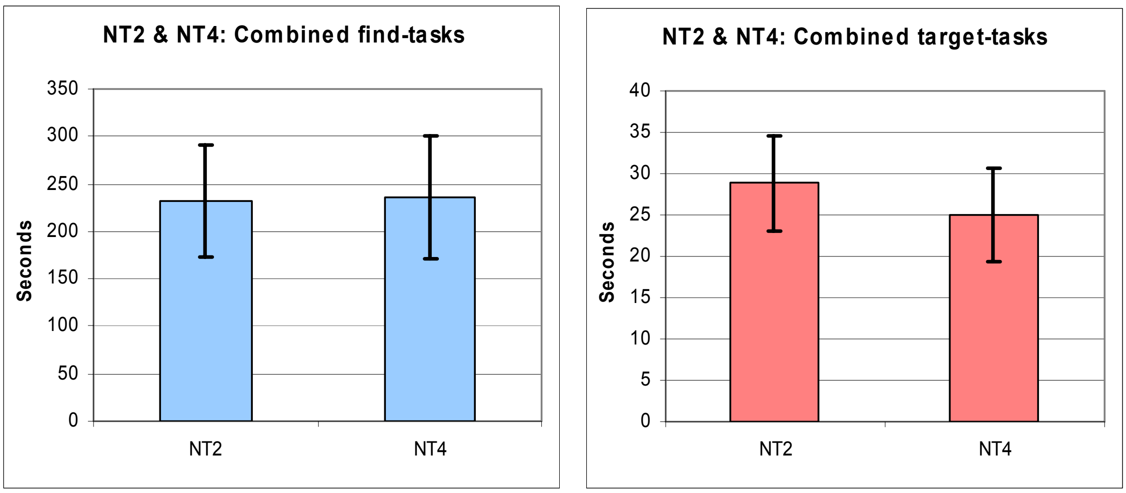
The confidence interval calculations give (15.4 ± 14.6) % for the find-tasks and (9.2 ± 10.0) % for the target-tasks.

Since the differences between the navigation techniques are more apparent in the find-tasks, there is an indication that the lack of panning has a larger effect in the find-tasks. This was to be expected as the find-tasks involve a lot more planar (vertical and horizontal) movements than the target-tasks where the most important factor is precise zooming. As the subjects in general had less use for panning in the target-tasks the effect of removing the pan is smaller.

The conclusion for these two gaze-controlled navigation techniques is that without the panning it took our subjects 9-15% longer to solve the tasks, with the effect more apparent in tasks requiring more complex navigation. So “removing the P” from a gaze-controlled PZN in our case results in task times being increased, possibly with as much as 30% (taking the confidence interval into account). These findings support our hypothesis H2. Although, it should be noted, in the lower extreme of the confidence interval there would be no noticeable difference, disproving the hypothesis.

Comparison: Mouse and gaze (NT2 & NT4)

As explained in chapter 3.3 our NT4 (gaze; pan+zoom) was an attempt at transforming the strengths of the mouse controlled NT2 (mouse; pan+zoom) to a gaze-controlled environment. The zooming is directly comparable between the two navigation techniques, but panning is implemented differently.



	NT2 (mouse; pan+zoom)	**NT4** (gaze; pan+zoom)	**Δ**
Find-tasks	233 (59)	235 (66)	-1.1%
Target-tasks	29 (6)	25 (6)	+16%

Figure 45: Comparison of NT2 and NT4: The average times taken to solve all the tasks for each navigation technique. Times are given in seconds, with standard deviation in parentheses. The difference is calculated as percentage of NT4.

Somewhat surprisingly (given how accustomed users are to a mouse) in the find-task the results are indistinguishable, and in the target-tasks the mouse-controlled NT2 is 16% slower than gaze-controlled NT4.

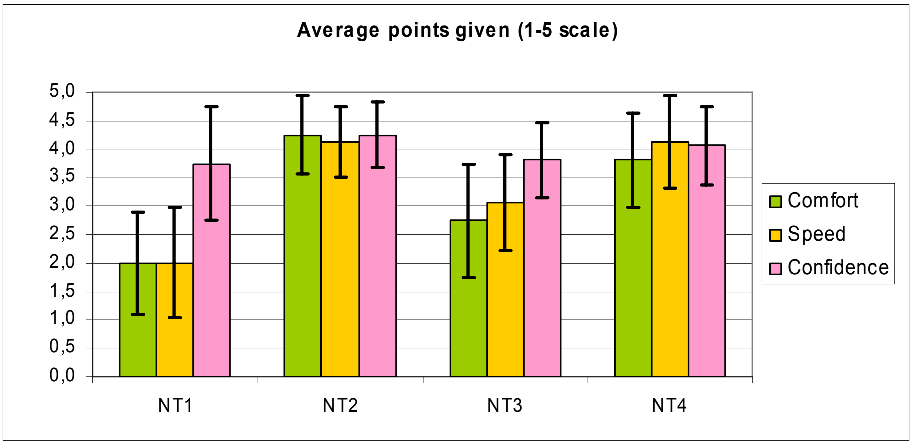
The t-test confirms that the results from the find-tasks are impossible to distinguish with $p = 0.45$ and the confidence interval (-1.1 ± 19.8) %. But for the target-tasks the t-test gives $p < 5 \cdot 10^{-3}$, confirming that the gaze-controlled navigation is in fact faster than the mouse-controlled (15.7 ± 11.2) %, the true difference ranging from 4.5% to 26.9%.

Although it is not our intention to compare NT2 (mouse; pan+zoom) and NT3 (gaze; zoom) in any detail, results show that in the target-tasks NT2 is about 5.9% slower than NT3. However the t-test gives $p > 0.1$ so this seems inconclusive (5.9 ± 9.7) %.

These results indicate that the gaze-controlled NT4 (gaze; pan+zoom) is comparable to the mouse-controlled NT2 (mouse; pan+zoom) and for tasks that put more emphasis on direct zoom than planar movements the gaze-control seems more effective than the mouse. These findings are counter to our hypothesis H3 as the mouse and gaze appear comparable, and in certain tasks the gaze control seems to be faster.

4.3 Qualitative evaluations of the navigation techniques

After completing all tasks in each block, we asked the subject to grade the navigation technique on three parameters; general comfort, perceived speed and how confident they felt in their given answers. We used a point scale from 1 to 5, with 5 points being “very comfortable”, “very fast” or “very confident”. At the end of the test, when the subjects had experienced all the techniques, they were asked to compare their evaluations and modify if needed.



	Comfort	Speed	Confidence
NT1 (mouse; toolbar)	2.0 (0.9)	2.0 (1.0)	3.8 (1.0)
NT2 (mouse; pan+zoom)	4.3 (0.7)	4.1 (0.6)	4.3 (0.6)
NT3 (gaze; zoom)	2.8 (1.0)	3.1 (0.9)	3.8 (0.7)
NT4 (gaze; pan+zoom)	3.8 (0.8)	4.1 (0.8)	4.1 (0.7)

Figure 46: The average points given for each navigation technique on a 1-5 scale for three parameters. Standard deviation is shown in parentheses.

As evident from these results NT1 (mouse; toolbar) scores low on both comfort and speed, whereas NT2 (mouse; pan+zoom) and NT4 (gaze; pan+zoom) receive very similar scores. Normalizing the grades to counteract the inevitable variation resulting from each subject having his or her own point of reference gives an almost identical result as that in Figure 46 and is not shown here.

The grades can also be viewed as indicative of the relative ranking of techniques; that is which technique(s) a subject views as being the best, which second, etc.

Comfort	**1st**	**2nd**	**3rd**	**4th**
NT1 (mouse; toolbar)	0	2	10	4
NT2 (mouse; pan+zoom)	12	3	1	0
NT3 (gaze; zoom)	1	6	7	1
NT4 (gaze; pan+zoom)	7	8	1	0

Speed	1st	2nd	3rd	4th
NT1 (mouse; toolbar)	1	3	8	4
NT2 (mouse; pan+zoom)	11	4	1	0
NT3 (gaze; zoom)	2	8	5	1
NT4 (gaze; pan+zoom)	11	5	0	0

Confidence	1st	2nd	3rd	4th
NT1 (mouse; toolbar)	10	3	3	0
NT2 (mouse; pan+zoom)	14	2	0	0
NT3 (gaze; zoom)	8	8	0	0
NT4 (gaze; pan+zoom)	12	3	1	0

Figure 47: The number of subjects (N=16) ranking a specific navigation technique as the best, second best, etc., within each of the three parameters.

When it comes to comfort NT2 (mouse; pan+zoom) is clearly ranked highest, but in speed NT2 and NT4 (gaze; pan+zoom) are indistinguishable. This shows that although subjects find the well known mouse more comfortable than gaze-control, they have a sense that NT4 is equally fast (as the performance results confirm).

One of our subjects summarizes this quite clearly:

> "I felt more comfortable using the mouse, but I think this [NT4] was a bit faster, although I'm not quite sure why".

4.4 Summary of results

According to our experimental hypotheses our results can be summarized:

R1: The NT2 (mouse; pan+zoom) interface did in fact perform much better than NT1 (mouse; toolbar) and was greatly preferred by our subjects.

R2: In gaze navigation, removing the panning did result in worse performance. The difference could be as much as 30%.

R3: Contrary to our hypothesis, in our test environment the gaze-controlled NT4 (pan+zoom) proved to be just as fast as the mouse-controlled NT2 (pan+zoom) and in some cases faster, even though the vast majority of our subjects had never experienced any form of gaze interaction before.

4.5 Discussion of results

Clearly the most interesting result of our testing is the fact that our gaze-controlled navigation technique NT4 did not only equal the performance of the mouse-champion navigation technique NT2, but did in fact surpass it in the target-tasks.

As was to be expected the navigation technique of selecting zoom and pan tools from a toolbar (NT1) turned out to be the slowest of the four navigation techniques tested. This technique is arguably the most common implementation of a pan/zoom navigation, and was included in our test as a benchmark reference of sorts. Although it is well known and does only require the use of a mouse, its limitations in strict modality and the cumbersomeness of selecting a new mode from the toolbar severely limits its effectiveness in the sorts of tasks our test focussed on.

The navigation technique we chose as our *mouse-champion*; NT2 (mouse; pan+zoom) solves many of the problems of the toolbar technique. By using explicit quasi-modes it eliminates the need for choosing a new tool and automatically reverts to a neutral state as the user exits the quasi-modes. As an example the zoom-in quasi-mode is entered by pressing the zoom-in key, and as soon as it is released the zooming stops. Similarly the familiar quasi-mode of grabbing is entered by holding down the mouse button and exited by releasing it. In our implementation the subjects had to use both the mouse and keyboard due to the limitations of the Flash Player, but a version only using the various buttons on the mouse, and thereby only one hand, is quite conceivable.

Our navigation technique NT4 (gaze; pan+zoom) was an attempt at transferring the strengths of NT2 (mouse; pan+zoom) to a gaze-controlled environment, with the main challenge being the implementation of panning.

In our initial testing of various mouse-controlled techniques we found indications that only allowing for zoom (ignoring pan) severely limited the usefulness of the technique, or in other words that zoom and pan complement each other nicely. This goes both ways as panning is also more flexible when combined with a zoom option.

In our email correspondence with Aza Raskin he remarked on this:

> “The problem with most 2D navigation is that pan is never at the right speed; it is either too fast or too slow. [...] But zooming solves that problem by allowing continuous control of panning speed (by narrowing or widening the field of view via zooming in and out).”

The best known panning techniques are probably the use of scrollbars or grabbing, neither of which lends itself well for gaze-control. Our implementation of an edge-scroll described in Chapter 3.3 was an attempt at coming up with a useful gaze-controlled panning (while it is perhaps not as appropriate in mouse-controlled situations). Leaving the edge-scroll out in NT3 (gaze; zoom) was intended for quantitatively estimating its usefulness in NT4 (gaze; pan+zoom).

Comparing the performances of NT3 and NT4 in the find-tasks, where the option of planar movements was a valuable asset, shows that removing the panning option increases the task-times by (15.4 ± 14.6) % – that is anywhere between 0.8 and 30%. In the target-tasks, where the main emphasis is on zooming, leaving out the panning also increases the task-times although not as dramatically (9.2 ± 10.0) %.

As it could be argued that any panning option is better than none, comparing NT3 and NT4 is not conclusive about the actual applicability of edge-scrolling. However our main findings that NT4 (gaze; pan+zoom) is comparable to NT2 (mouse; pan+zoom) can be seen as an indication that edge-scrolling appears to be at least an acceptable substitute for the mouse-grab panning. There is undoubtedly room for improvement in the implementation of edge-scrolling, but the potential seems to be there.

Is the eye in fact faster than the mouse?

What makes our result that NT4 (gaze; pan+zoom) is comparable to NT2 (mouse; pan+zoom) even more interesting is the fact that many aspects appear to point to a more favourable outcome for the mouse:

1. All our subjects had extensive experience of using a mouse, but only 3 had any previous experience of gaze interaction.
2. Unfamiliarity with gaze interaction caused about half of our subjects to mention experiencing some discomfort.

3. The edge-scroll sometimes resulted in unintended movements, whereas grabbing with the mouse got no negative comments.
4. The eye-tracking technology could be improved both spatially and temporally, moving from experimental to established mainstream eye trackers.

But, at least in theory, gaze-controlled navigation of this sort should have its benefits. In a typical task of zooming in on an area of interest using NT2 (mouse; pan+zoom) the procedure could be described as follows:

1. Locate the area of interest
2. Point mouse to the area of interest
3. Press the zoom-in key on the keyboard (assuming a finger is already there)

In NT4 (gaze; pan+zoom) the second step can be skipped since the zooming is automatically centred on the subject's gaze and no movement of the mouse is needed. When repeatedly zooming in and out the milliseconds saved each time start to add up.

In the case of a long continuous zooming, such as required in the target-tasks (zooming from 100% to 20,000% magnification) it is often necessary to adjust the centre of the zooming motion as more details emerge. In NT2, while continuing to hold down the zoom-in key, the mouse can be moved to compensate when necessary. In NT4 this adjustment is automated as long as the subject keeps the gaze on the intended target, thereby eliminating the need for any conscious correction and giving the appearance of an almost mind-reading experience (or perhaps more accurately *intention-reading*).

When both of the above mentioned actions (choosing where to zoom and adjusting the centre of the zooming motion) come into play in the target-tasks, as well as any limitations of the edge-scroll being less important, it is perhaps not surprising that the gaze-controlled navigation technique turns out to have an advantage in these types of tasks.

Overly strenuous gaze

In theory gaze-control should not be any more strenuous on the eye than when using a mouse, as the eye-tracking is totally unnoticeable and in fact the gaze-control involves nothing more than looking at the monitor. However, in our testing approximately half of our subjects mentioned experiencing various degrees of discomfort or fatigue when using the gaze-controlled navigation techniques, while the other half experienced no problems. This is reflected in the lower grades for comfort given to NT4 (gaze; pan+zoom) than NT2 (mouse; pan+zoom), even though they were perceived to be similarly fast.

We are used to some sort of muscular exercise being necessary to translate our intentions into action, and for some of our subjects the amount of effort they were (unnecessarily) putting into their eye-movements was clearly visible; straining their heads and tensioning neck muscles. Unfamiliarity with the technique and possibly a stressful test situation may both have had an effect. Those subjects that were able to relax had no problems.

A few subjects mentioned slight motion-disorientation (usually described as the beginnings of motion sickness). This is an effect well known from people playing fast-moving computer games from a first person perspective, either for the first time or after a long break and usually disappears quickly. Initial fatigue is of course well known for most types of input devices, for instance when unaccustomed users have to intensively use a mouse or play computer games using a console input.

It is therefore our belief that repeating our test, giving the subjects more time to become accustomed to the gaze-control, would probably result in the comfort-ratings being higher and quite possibly showing an even better performance for the gaze-controlled navigation.

Our results in a larger context

Our results show that, even for subjects inexperienced in gaze-interaction, our first implementation of a gaze-controlled pan/zoom navigation technique is just as efficient as the corresponding mouse technique in solving tasks requiring extensive zooming and panning motions. Not only that, but in tasks that consist of zooming in on specific targets, rather than investigating the whole dataset, the gaze-controlled version is significantly faster.

These findings support the previously mentioned claims of Bolt in 1981 that "eyes-as-output has a promising future in interactive graphics",[79] and indicate that further studies in gaze-controlled navigation clearly is an area of interest.

Several studies have showed that in selection or pointing tasks gaze is faster than mouse, such as Ware in 1987 [80] and Ovekova et al.: "Recent work showed that the eye is faster than the mouse as a

[79] Bolt, Richard A. (1981) p. 10
[80] Ware, C. et al. (1987)

source of visual input in a target image identification task." [81] This is also the case in the previously described study by Zhai et al.; where combining gaze and manual selection gave the best results.[82]

But even though we have extensively searched for texts dealing with gaze-control for navigation tasks, and found several discussing its apparent applicability, we have not found any study that shows the eye to be as fast or even faster than the mouse for navigation tasks as our does.

Immediately an interesting application could be in using gaze not only to select a target, but to navigate towards it by zooming (such as in our target-tasks) as this appears to be the sort of tasks where gaze-control yields significantly better results by focusing directly on the interaction between the eye and the monitor.

This seems rational as the user already uses gaze to decide which areas are of interest, so when the decision is made to research an object or area further, the gaze is already directed at the target of interest:

> "The control-to-display relationship for [the eye] is already established in the brain." [83]

---

[81] Oyekoya, O. et al. (2006) p. 1
[82] Zhai, S. et al. (1999)
[83] Jacob, R. J. K. (1991) p. 12

5 Conclusions

Our findings have shown that gaze interaction can be successfully implemented for searching large multiscaled environments by employing pan/zoom navigation.

Zooming can be seen as an analogy for our natural focus-targeting. Even though objects on a computer monitor may all be in the same focus distance, zooming emulates the way we target and perceive a specific point of interest in the real world.

Zooming also translates easily to gaze interaction, as it extends our already established process of using perception to locate areas of interest. Zoomable gaze interaction allows a direct response to these perceptual stimuli, bypassing the need for redirecting the attention of the user to moving a mouse.

Two experiments were conducted by us in order to shed light on how navigation could best be implemented using gaze interaction.

In the initial experiment 4 mouse navigation techniques were tried; three different pan options and one semantic zoom navigation technique. The information space was in a single scale tabular layout with different levels of structure: structured, semi- structured and random. The tasks required the subjects either to find a single target or to browse and find a sequence of targets.

The overall conclusions were, first of all, that for two of the panning techniques; grabbing the map with the mouse and an implemented edge-scrolling technique, the results were almost indistinguishable. These findings proved valuable in designing the panning options in our second experiment.

Secondly, zoom implemented without the full benefit of a perspective and without panning proved very inadequate when solving most of the tasks. Both these findings also influenced the design of our second experiment.

In the subsequent experiment we compared zoomable interfaces controlled by both mouse and gaze for different search tasks in multiscaled environments.

There were two information spaces. One large multiscaled environment with 2000 nodes placed in four different scale layers on the map, and a smaller multiscaled map designed for specific target exploration.

Tasks in the map with 2000 nodes consisted of visual inspection tasks requiring the subjects to use extensive pan and zoom to find several bricks according to different criteria. Tasks in the smaller map required subjects to zoom in on targets in already known areas.

Four navigation techniques (NT) were tested against each other, two mouse-controlled and two gaze-controlled. The mouse-controlled NT1 functioned by selecting different modes of pan/zoom from an onscreen toolbar. In the mouse-controlled NT2 zooming was performed by using keyboard keys and panning by grabbing the map with the mouse. The gaze-controlled NT3 only had zooming using keyboard keys implemented and finally gaze-controlled NT4 allowed zooming by using keyboard keys and panning with an onscreen edge scroll.

The findings from our experiments show that the familiar toolbar pan/zoom navigation (NT1) was by far the most inefficient way of navigating; resulting not only in the longest task times but also the highest level of frustration. Also, zoom navigation without the ability to pan (NT3) expectedly proved to be less efficient than the pan/zoom gaze solution (NT4).

The most relevant of our findings concerned the comparison of the mouse pan/zoom technique using keyboard keys and mouse (NT2) and the gaze pan/zoom technique using keyboard keys and an eye-tracker (NT4).

Our findings show that even for first time users of gaze interaction, the gaze technique rivals that of mouse interaction, when searching in large multiscaled information spaces; to the point where they are statistically indistinguishable.

These results were contrary to our initial estimates regarding the efficiency of the navigation techniques, as we had assumed the familiarity of mouse control would lead to the mouse-controlled pan/zoom technique (NT2) performing better than the gaze-controlled implementation (NT4).

In tasks involving inspecting known areas of an information space, our findings show that navigating with gaze surpasses mouse navigation. Tasks took 16% longer to solve with mouse-controlled NT2 than the gaze-controlled NT4.

Although this was also contrary to our expectations, theoretically gaze target acquisition should be faster than mouse selection because the targets have already been perceptually processed, before we begin the action of moving the mouse.[84]

Even though this experiment did not involve the selecting of targets, the results indicate that navigating to a specific target can in some cases be done faster with gaze interaction in zoomable interfaces than when using a mouse. This efficiency implies that gaze interaction could beneficially supplement the mouse in tasks involving navigation and selection of targets in large multiscaled information spaces.

In regard to whether or not gaze interaction can substitute the mouse, our results are inconclusive as we only tested a small subset of the actions typically controlled by a mouse. However, navigating with gaze in zoomable interfaces seems a viable substitute to navigating with a mouse, especially in situations or circumstances where gaze is the only alternative.

Subjects, who started the test sequence with gaze, remarked that they had to re-familiarize themselves with the mouse; having to point with it before zooming in on a specific point. This indicates how quickly we adapt to the concept of the interface responding to our point of interest. Zooming allows the interfaces to respond in a familiar way and also minimizes a potential information overload, by allowing the user to limit the detailed view as much as is necessary or wanted.

An advantage of gaze in zoomable interfaces is that it allows us to skip the cognitive process of moving the mouse to correspond with our focus of interest. Gaze interaction in zoomable interfaces allows us to react and respond to an object of interest without being penalized for instinct.

Recapping the three general hypotheses presented in Chapter 2.5 on gaze-controlled pan/zoom navigation, we can now try to shed some light on their validity:

---

[84] Zhai, S. et al. (1999)

1.1. Eye-tracking can effectively supplement existing input devices.

We still believe this to be true, and although our experiments have not been intended to directly validate this claim, we discuss some ideas in the following chapter about future research that might support this hypothesis.

1.2. Zooming interfaces have traits that correspond well with eye-based interaction.
1.3. A gaze-controlled navigation technique should be able to successfully entail many familiar aspects of sight and perception.

The results of our second experiment show that despite our subjects' unfamiliarity with gaze-interaction the gaze-controlled NT4 proved to be at least as efficient as our *mouse-champion* NT2. This seems to indicate that our *gaze-challenger* is easy to learn and observing our subjects, the interaction appeared to be self-explanatory. It is therefore our firm belief that zoomable interfaces are very well suited implementations for gaze-interaction.

6 Future research

Although we have attempted to test many elements of our proposed gaze-controlled pan/zoom navigation technique (NT4), there are still questions that remain unanswered and design decisions that remain to be properly evaluated for the navigation technique to be as robust and usable as possible.

In order to accommodate gaze-controlled panning we have proposed an onscreen edge-scrolling solution. Several aspects of this navigation technique could be researched further and improved upon. Pan speed is an aspect which can be studied in order to improve performance:

> "The problem with most 2D navigation is that pan is never at the right speed; it is either too fast or too slow." [85]

Studying what the average preferred pan speed is, and whether the user should be able to calibrate it to suit individual preference, could be projects for the future. As well as examining how large the scroll areas should be and whether or not it would improve the usability of the system to have them visualized in some way.

Similarly, the same kind of investigation could be applied to the zoom function. Exploring how fast the zoom could realistically be and still maintain the useful aspect of allowing the user to perceive content while zooming. The corrective behaviour implemented in Aza Raskin's version could also be examined, in regard to further fine-tuning for gaze interaction.

When studying the speeds of pan and zoom, the sampling rate of the eye-tracker and the refresh rate of the visualizations are important factors. In our gaze experiments, the eye-tracker had a sampling rate of 50 Hz and the Flash application ran at 45 fps, both numbers that could be improved.

Zoom appears to lessen the importance of eye-tracker precision and accurate calibration. In regard to the implementation we have designed, it would be interesting to find out what the lower limits in resolution are. A self-adjusting calibration is even conceivable by analyzing the objects zoomed in on and comparing them to the preceding gaze measurements.

[85] Excerpt from correspondence with Aza Raskin.

We chose to use a static multiscaled information space in our experiments, but a study done on gaze-control in semantic zooming, e.g. of a dynamically presented tree structure, could expand the field of use.

In situations where gaze is the sole form of available interaction further investigation into zoomable interfaces should prove valuable, such as investigating what non-keyboard controls could be useful for zooming in and out. Depending on the context of the usage and the tasks needing to be solved, either alternative input devices, such as voice activation, or some kind of onscreen zoom controls could be efficiently used for zooming in or out.

Most of our test subjects were first time users, but a study into the pan/zoom navigation patterns of long time users could perhaps allow for observations that would further the scope of use for this navigation technique. Subsequently an experiment conducted with long time users may reveal whether or not zoomable interfaces ultimately benefit from overviews, as these seem primarily of use for novice users.

In regard to our implementation, locating an overview at the edges of the monitor would be problematic. Solutions are plausible where using a quasi-mode would temporarily bring up an overview in the middle of the screen, similar to how maps are often implemented in "first person shooter" games. In very large intricate information spaces this could be a valid option.

This study separates navigation from selection. If selection is necessary in gaze interaction, zoomable interfaces have great potential; as the importance of precision is greatly lessened with a more flexible target size. A version of selection is conceivable, by using parallels to semantic-zoom, where an object is not made available for dwell-activation until it has been zoomed in on, thereby lessening Midas' Touch problems.

An interesting project could be to design an escape from the WIMP model in zooming. Rather than clicking on an icon to open a window; semantic zooming could be used to zoom in on an icon, which in turn would reveal the "window" it represents.

In a foreseeable future a commercial brand of eye trackers could very well become a part of the standard PC setup. Already all Apple monitors, both stationary and on portable computers, have built-in cameras; adding the option for eye tracking is (at least in theory) just a matter of adjusting a camera to detect infrared frequencies and adding an infrared light source.

It therefore requires only a small leap of imagination to envision a world where gaze navigation is an ordinary part of human computer interaction. We conclude this thesis by presenting three possible scenarios involving gaze-interaction, demonstrating some of the possibilities we can envision.

Future scenario 1:

In an alarm central an emergency operator receives a phone call from a witness to a car accident. On a large monitor the operator uses gaze-control to zoom in on a map of the area where the accident has happened and by asking the witness to describe its location in relation to visible landmarks, the operator quickly pinpoints the precise location. As emergency vehicles approach the site of the accident the operator zooms in and out on the semantic map, which shows the GPS coordinates of the vehicles, keeping an eye on the progress of ambulances and police cars approaching the site.

Since the zooming does not require the use of a mouse, the operator is free to use both hands on the keyboard; zooming in/out and switching between audio-communication with the witness (getting a more detailed description of the accident) and the emergency vehicles (giving them route information and informing them of the situation at hand). Aided by the dynamic overview the operator helps the police position their cars to secure the site, and then collaborates with a colleague controlling traffic lights to secure an open route for ambulances to the nearest hospital.

Future scenario 2:

A medical doctor specializing in X-ray analysis uses a gaze-interactive monitor to zoom in on details of a patient's X-ray photos, looking for evidence of bone cancer. The gaze-contingent display is capable of showing particularly high resolution at the area of interest and the doctor can use both hands to take notes. After coming to a conclusion she initiates a dynamic on-screen conversation with the patient's attending doctor at another hospital. The same image is visible on the monitors of both doctors, and while verbally describing her findings the X-ray expert ensures that the attending doctor is looking at the right places by watching an on-screen indication of his gaze.

Future scenario 3:

> Average Jane is in her home office. A small group of emails are lying side by side in the middle of the screen. A few unopened letters are flashing to the right; by looking at them the view scrolls and the group is centred. She zooms in on a letter from her children's school, and as it gets bigger the content is revealed. Jane zooms out again and looks at the letter to her right; it is an invitation for dinner on Saturday; by gazing on the return address and zooming in, an answer sheet opens.
>
> The ringing of an incoming call interrupts her; she accepts the call by pressing a button on the keyboard and the screen changes; the separate video images of two friends are shown. As they talk and she looks from one to the other their voices subtly fade in and out. When they are done the screen returns to her cluster of emails.

As these scenarios show, there are many possibilities for gaze-interaction, in particular coupled with pan/zoom navigation. They might at first seem somewhat implausible, but the necessary technologies already exist, and it is primarily a matter of applying our imagination to the possibilities at hand. If gaze-interaction starts to become commercially viable, even if it is only for very specialized tasks at first, it is likely to lower the cost of eye-tracking equipment, which could eventually lead to gaze-interaction becoming a part of our everyday lives.

Literature

1. Bates, Richard (1999): Multimodal Eye-Based Interaction for Zoomed Target Selection on a Standard Graphical User Interface
http://www.cse.dmu.ac.uk/~rbates/Bates1.pdf

2. Bates, Richard; Istance, Howell (2002): Zooming Interfaces! Enhancing the Performance of Eye Controlled Pointing Devices
http://doi.acm.org/10.1145/638249.638272

3. Baudisch, P.; Good, N.; Belotti, V.; Schraedley, P. (2002): Keeping Things in Context: A Comparative Evaluation of Focus Plus Context Screens, Overviews, and Zooming.
http://doi.acm.org/10.1145/503376.503423

4. Baudisch, P., DeCarlo, D., Duchowski, A., and Geisler. W. (2003): Focusing on the Essential: Considering Attention in Display Design. In Special Issue on Attentive User Interfaces
http://doi.acm.org/10.1145/636772.636799

5. Bederson, B. B.(2001): PhotoMesa: A zoomable image browser using quantum treemaps and bubblemaps.
http://doi.acm.org/10.1145/502348.502359

6. Bederson, Benjamin B.; McAlister, Britt (1999): Jazz: An Extensible Zoomable User Interface Graphics Toolkit in Java
http://elib.cs.berkeley.edu/docs/data/pl/Digital_Documents/29/PDF/00000001.pdf

7. Bederson, B. B.; Hollan, J. D.; Perlin, K.; Meyer, J.; Bacon, D.; Furnas, G..(1996) Pad++: A zoomable graphical sketchpad for exploring alternate interface physics.
http://www.cs.umd.edu/hcil/jazz//learn/papers/jvlc-96-pad.pdf

8. Bly, S.; Rosenberg, J.K. (1986): A Comparison of Tiled and Overlapping Windows.
http://doi.acm.org/10.1145/22627.22356

9. Bolt, Richard A. (1981): Gaze-Orchestrated Dynamic Windows
http://doi.acm.org/10.1145/800224.806796

10. Bolt, Richard A. (1982): Eyes at the Interface
http://doi.acm.org/10.1145/800049.801811

11. Butler, Jill; Holden, Katrina; Lidwell, William (2003): Universal principles of design, Rockport Publishers, Inc.

12. Heer, Jeffrey; Card, Stuart K.; Landry, James A. (2005): prefuse: A Toolkit for Interactive Information Visualization
http://guir.berkeley.edu/pubs/chi2005/prefuse.pdf

13. Fjeld, Morten; Ironmonger, Nicholas; Voorhorst, Fred; Bichsel, Martin; Rauterberg, Matthias (1999): Camera Control in a Planar, Graspable Interface
http://www.idemployee.id.tue.nl/g.w.m.rauterberg//publications/IAESTED99paper.pdf

14. Fono, David; Vertegaal, Roel(2005): EyeWindows: Evaluation of Eye-Controlled Zooming windows for Focus Selection.
http://www.dgp.toronto.edu/papers/dfono_CHI2005.pdf

15. Franconeri, Steven L.; Simons, Daniel J. (2003): Moving and looming stimuli capture attention
http://www.psych.ubc.ca/~steve/myPapers/FranconeriSimons_Motion.pdf

16. Frøkjær, Erik; Hertzum, Morten; Hornbæk, Kasper (2000): Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?
http://doi.acm.org/10.1145/332040.332455

17. Frøkjær, Erik; Hornbæk, Kasper (2001): Reading of Electronic Documents: The Usability of Linear, Fisheye, and Overview + Detail Interfaces
http://doi.acm.org/10.1145/365024.365118

18. Funkhouser, T; Kai Li (2000): Large-Format Displays
http://ieeexplore.ieee.org/xpls/abs_all.jsp?isNumber=18499&arNumber=851745

19. Furnas, George W.; Bederson, Benjamin .B (1995): Space-scale diagrams: understanding multiscale interfaces
http://doi.acm.org/10.1145/223904.223934

20. Furnas, G.W. (1986): Generalized Fisheye Views.
http://doi.acm.org/10.1145/22627.22342

21. Greeno, James G. (1994): Gibson's affordances
http://ecologylab.cs.tamu.edu/courses/physicalInterfaces/hostedMaterials/gibsonAffordances.pdf

22. Gregory, Richard L. (1997): Eye and Brain: The Psychology of Seeing, Fifth Edition, Oxford University Press

23. Gutwin, C. (2002): Improving Focus Targeting in Interactive Fisheye Views.
http://doi.acm.org/10.1145/503376.503424

24. Hansen, John Paulin; Johansen, Anders Sewerin (2002): Augmentative and Alternative Communication: The Future of Text on the Move
http://www.it-c.dk/research/EyeGazeInteraction/Papers/Johansen_Hansen_2002.pdf

25. Hansen, John Paulin; Hansen, Dan Witzner; Johansen, Anders Sewerin; Itoh, Kenji; Mashino,Satoru (2003): Command Without a Click: Dwell Time Typing by Mouse and Gaze Selections
http://www.itu.dk/research/EyeGazeInteraction/Papers/Interact%202003-submit_final.pdf

26. Hansen, John Paulin; Hansen, Dan Witzner; Mackay, David; Nielsen, Mads (2004): Eye Tracking off the Shelf
http://doi.acm.org/10.1145/968363.968375

27. Hornbæk, Kasper; Benderson, Benjamin B.; Plaisant, Catherine(2002): Navigation Patterns and Usability of Zoomable User Interfaces with and without an Overview
http://doi.acm.org/10.1145/586081.586086

28. Hyrskykari, Aulikki; Päivi Majaranta; Kari-Jouko Räihä (2005): From gaze control to attentive interfaces.
http://www.cs.uta.fi/research/hci/gaze/docs/Hyrskykari-HCII2005.pdf

29. Itoh, Kenji; Aoki,Hirotaka; Hansen, John Paulin(2006): A Comparative Usability Study of Two Japanese Gaze Typing Systems
http://doi.acm.org/10.1145/1117309.1117344

30. Istance, Howell (2006): Communication Through Eye-Gaze: Where We Have Been, Where We Are Now and Where We Can Go From Here
http://doi.acm.org/10.1145/1117309.1117311

31. Jacob, Robert J. K.(1991): The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get
http://doi.acm.org/10.1145/123078.128728

32. Jacob, Robert J. K.(1996): Human-Computer Interaction: Input Devices
http://doi.acm.org/10.1145/234313.234387

33. Jacob, Robert J. K; Karn, Keith S (2003): Eye-tracking in human-computer interaction and usability research: Ready to deliver promises
http://www.cs.ucl.ac.uk/staff/J.McCarthy/pdf/library/eyetrack/eye_tracking_in_HCI.pdf

34. Lamping, John; Rao, Ramana (1995): The Hyperbolic Browser : A Focus 1 Context Technique for Visualizing Large Hierarchies
http://www.cs.kent.edu/~jmaletic/cs63903/papers/Lamping96.pdf

35. Lankford, Chris(2000): Gazetracker: Software Designed To Facilitate Eye Movement Analysis
http://doi.acm.org/10.1145/355017.355025

36. Ma, Kwan-Liu (2000): Visualizing Visualizations: User Interfaces for Managing and Exploring Scientific Visualization Data.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?isNumber=18744&arNumber=865874

37. Ma, Kwan-Liu (2001): Large-Scale Data Visualization.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=933520&isnumber=20200

38. Mackinlay, Jock D.; Robertson, George G.; Card, Stuart K. (1991): The Perspective Wall: Detail and Context Smoothly Integrated.
http://doi.acm.org/10.1145/108844.108870

39. Majaranta, Päiva; Räihä, Kari-Jouko (2002): Twenty Years of Eye Typing: Systems and Design Issues
http://doi.acm.org/10.1145/507072.507076

40. Nekrasovski, Dmitry; Bodnar, Adam; Guimbretière, Francois; McGrenere, Joanna; Munzner, Tamara (2006): An Evaluation of Pan&Zoom and Rubber Sheet Navigation with and without an Overview
http://www.cs.ubc.ca/~dmitry/papers/chi2006.pdf

41. Noik, Emanual G. (1993): Exploring Large Hyperdocuments :Fisheye Views of Nested Networks
http://portal.acm.org/citation.cfm?id=962367.962372

42. Norman, Donald A. (1999) Affordance, Convention, and Design
http://doi.acm.org/10.1145/301153.301168

43. Oyekoya, Oyewole; Stentiford, Fred(2006): An Eye Tracking Interface for Image Search
http://doi.acm.org/10.1145/1117309.1117324

44. Perlin, Ken; Fox, David(1993): Pad: An Alternative Approach to the Computer Interface
http://doi.acm.org/10.1145/166117.166125

45. Plaisant, Catherine; Carr, David; Shneiderman, Ben (1995): Image Browsers: Taxonomy, Guidelines, and Informal Specification
http://portal.acm.org/citation.cfm?id=624606.625443

46. Pomplun, Marc; Ivanovic; Nada; Reingold, Eyal M.; Shen, Jiye (2001): Empirical Evaluation of a Novel Gaze-Controlled Zooming Interface
http://www.cs.umb.edu/~marc/pubs/pomplun_et_al_HCI2001.pdf

47. Pook, S.; Vaysseix G.; Barillot E.(1998) Zomit: biological data
http://www.infres.enst.fr/net/zomit/bioinformatics1998/bioinformatics1998.pdf

48. Pook, Stuart; Barillot, Emmanuel; Lecolinet, Eric; Vaysseix, Guy (2000): Context and Interaction in Zoomable User Interfaces
http://doi.acm.org/10.1145/345513.345323

49. Rao, Rajesh P. N.; Zelinsky, Gregory J. ; Hayhoe, Mary M.; Ballard, Dana H. (1997): Eye Movements in Visual Cognition: A Computational Study
http://www.cs.rochester.edu/research/brain/vrlab/papers/tr97.1.pdf

50. Raskin, Jef (2000): The Humane interface, ACM Press

51. Reingold, Eyal M.; Loschkky, Lester C.; McConkie, George W.; Stampe, Dave M.(2003) Gaze-contingent Multiresolutional Displays: An Integrative Review.
Human Factors: The Journal of the Human Factors and Ergonomics Society; Volume 45, number 2, summer 2003.

52. Sarkar, Manojit; Snibbe, Scott S; Tversky, Oren J; Reiss, Steven P (1993): Stretching the Rubber Sheet: A Metaphor for Viewing Large Layouts on Small Screens.
http://doi.acm.org/10.1145/168642.168650

53. Sarkar, M.; Brown, M.(1992): Graphical Fisheye Views of Graphs.
http://doi.acm.org/10.1145/198366.198384

54. Tanriverdi, Vildan; Jacob, Robert J.K. (2000): INTERACTING WITH EYE MOVEMENTS IN VIRTUAL ENVIRONMENTS
http://www.eecs.tufts.edu/~jacob/papers/chi00.tanriverdi.pdf

55. Velichkovsky, Boris M.; Hansen, John Paulin (1996): New Technological Windows into Mind: There is More in Eyesand Brains for Human-Computer Interaction
http://doi.acm.org/10.1145/238386.238619

56. Vertegaal, R. (2003) Attentive User Interfaces.
http://interruptions.net/literature/Vertegaal-CACM03-p30-vertegaal.pdf

57. Voss, Jonas C. (2005): Where the hand meets the eye, Rethinking computer games for visual control

58. Ware, Colin; Mikaelian, Harutune H. (1987): An evaluation of an Eye Tracker as a Device for Computer Input
http://doi.acm.org/10.1145/29933.275627

59. Weiser, M. (1993): Some computer science issues in ubiquitous computing.
http://doi.acm.org/10.1145/329124.329127

60. Zeki, Semir (1993): A vision of the brain, Blackwell Scientific Publications

61. Zhai, Shumin; Morimoto, Carlos; Ihde, Steven (1999) Manual And Gaze Input Cascaded (MAGIC) Pointing
http://doi.acm.org/10.1145/302979.303053

Links

62. *http://www.cs.umd.edu/hcil/jazz/learn/patterns.shtml#Semantic%20Zooming*

63. *http://www.cs.umd.edu/hcil/jazz/*

64. *http://www.krak.dk*

65. *http://rchi.raskincenter.org/index.php?title=Core_Principles*

Appendixes

Appendix A: Key results from Experiment 1

Appendix B: Results from Experiment 2

Appendix C: t-Test results from Experiment 2

Appendix D: The script for Experiment 1 (in Danish)

Appendix E: The script for Experiment 2 (in Danish)

Appendix F: Examples from the colour-blindness test

Appendix A: Key results from Experiment 1

All results are given in seconds. Obvious errors in task solving are omitted.

Average times for task types, ordered by the NT being used:

	Edge Scroll	Cylinder	Grab	Zoom
How many of a suit?	14,2	36,3	21,1	12,7
s (standard deviation)	3,8	22,0	10,3	4,3
What values are they?	13,7	28,1	16,2	22,0
s	3,1	19,9	6,3	6,7
How many of a number?	19,0	40,6	22,1	49,3
s	7,3	19,0	13,3	18,1
Find specific card?	9,2	12,0	10,3	14,6
s	5,1	8,6	6,1	8,2

Avg. for finding a specific card, by map and NTs:

	Fully str.	Semi Str.	Random
Edge Scroll	6,7	8,0	13,0
s (standard deviation)	3,5	2,3	6,0
Cylinder	5,9	8,8	20,9
s	1,9	2,7	8,8
Grab	7,1	10,4	13,5
s	4,6	5,1	6,8
Zoom	9,9	16,5	17,5
s	6,5	7,4	6,0

Average times for task types, ordered by the map type being used:

	Fully str.	Semi str.	Random
How many of a suit?		15,2	27,0
s (standard deviation)		6,6	19,3
What values are they?		15,9	24,1
s		5,5	15,5
How many of a number?	23,4	42,1	32,8
s	11,1	22,9	18,6
Find specific card?	7,4	11,1	16,3
s	4,7	6,0	8,3

Avg. for "how many of a number"-tasks, by map and NTs:

	Fully str.	Semi Str.	Random
Edge Scroll	16,6	23,4	17,0
s (standard deviation)	7,0	7,9	4,7
Cylinder	27,5	55,6	38,6
s	7,2	22,7	11,0
Grab	16,2	26,8	23,3
s	8,9	10,8	16,5
Zoom	33,2	62,5	52,2
s	10,0	15,1	14,6

Avg. for "how many of a suit/number"-tasks, by NTs and order:
(i.e. test persons having each NT as their 1st, 2nd, 3rd or 4th)

Order	1	2	3	4
Edge Scroll	15,1	18,0	17,1	18,1
s (standard deviation)	5,7	8,0	5,8	6,2
Cylinder	34,9	29,6	45,4	45,4
s	15,0	15,8	22,6	22,1
Grab	25,1	20,5	15,9	25,4
s	12,9	5,3	5,2	17,6
Zoom	30,9	38,1	35,0	34,7
s	18,7	25,9	21,2	24,6

Avg. for "how many of a suit"-tasks, by map and NTs:

	Fully str.	Semi Str.	Random
Edge Scroll		11,9	16,5
s (standard deviation)		2,7	3,3
Cylinder		20,1	52,4
s		8,3	19,3
Grab		16,8	25,4
s		5,7	12,0
Zoom		11,9	13,6
s		4,2	4,3

Avg. for "what values are they"-tasks, by map and NTs:

	Fully str.	Semi Str.	Random
Edge Scroll		12,0	15,2
s (standard deviation)		1,7	3,2
Cylinder		16,1	39,1
s		3,9	22,3
Grab		12,9	19,5
s		3,1	7,0
Zoom		22,9	21,2
s		4,6	5,5

Avg. for "how many of a number"-tasks, by NTs and order:
(i.e. test persons having each NT as their 1st, 2nd, 3rd or 4th)

Order	1	2	3	4
Edge Scroll	17,0	19,3	19,0	20,7
s (standard deviation)	6,0	9,2	6,7	6,5
Cylinder	36,0	29,5	46,4	50,4
s	12,6	10,8	18,3	23,8
Grab	22,7	24,4	20,6	28,9
s	10,0	5,9	6,0	21,1
Zoom	42,2	54,9	49,0	51,1
s	15,7	20,0	15,8	18,0

Combined task times, by NTs:

	Edge Scroll	Cylinder	Grab	Zoom
Total time	169,0	321,1	203,3	303,7
s (standard deviation)	39,5	66,1	71,0	66,1

Appendix B: Results from Experiment 2

All results are given in seconds.

Combined "find-tasks" times by subjects				
NTs:	NT1	NT2	NT3	NT4
Subject 1	273,8	191,8	309,0	224,7
Subject 2	382,4	163,2	286,7	185,6
Subject 3	297,4	255,6	270,5	197,9
Subject 4	342,4	204,5	242,7	178,9
Subject 5	242,9	195,9	197,6	284,2
Subject 6	491,7	307,6	351,4	389,1
Subject 7	328,3	279,6	292,2	215,7
Subject 8	391,5	161,5	244,2	331,0
Subject 9	338,6	175,9	290,3	335,2
Subject 10	269,8	290,5	296,6	211,8
Subject 11	404,1	280,8	245,2	220,7
Subject 12	443,0	353,5	249,2	201,1
Subject 13	364,1	219,8	277,2	173,0
Subject 14	367,6	224,2	276,1	206,1
Subject 15	555,3	262,8	261,5	230,7
Subject 16	231,3	156,9	254,0	178,7

Combined "target-tasks" times by subjects				
NTs:	NT1	NT2	NT3	NT4
Subject 1	33,1	27,7	20,6	21,0
Subject 2	22,9	19,5	19,3	18,0
Subject 3	44,4	33,3	29,7	23,6
Subject 4	32,8	29,9	34,7	24,0
Subject 5	45,2	25,1	23,5	26,7
Subject 6	37,8	26,7	33,0	28,5
Subject 7	43,8	31,0	27,1	24,3
Subject 8	49,8	39,3	28,8	27,1
Subject 9	20,5	20,9	14,1	12,4
Subject 10	53,2	35,4	34,5	31,9
Subject 11	41,2	20,3	25,2	27,1
Subject 12	45,2	34,5	29,5	34,9
Subject 13	78,0	34,2	39,4	27,2
Subject 14	39,3	27,0	24,2	26,2
Subject 15	75,3	29,4	29,9	28,6
Subject 16	35,3	27,2	22,1	17,4

Average results: All subjects				
NTs:	NT1	NT2	NT3	NT4
Find 7 bricks with 2 knobs	188,0	126,4	142,1	124,6
s (standard deviation)	59,3	39,9	35,3	48,7
Find 7 wrongly placed…	169,8	106,4	129,5	110,7
s	42,0	32,4	29,0	25,7
Combined find-tasks times	357,8	232,8	271,5	235,3
s	88,4	59,1	34,8	64,6
Combined target-tasks times	43,6	28,8	27,2	24,9
s	15,6	5,7	6,6	5,6

Average results: Test progress 1 (NT1, NT2, NT3, NT4)				
NTs:	NT1	NT2	NT3	NT4
Find 7 bricks with 2 knobs	158,7	116,7	142,3	87,7
s (standard deviation)	20,9	20,3	43,0	7,3
Find 7 wrongly placed…	165,4	87,1	134,9	109,1
s	55,1	34,2	48,2	19,3
Combined find-tasks times	324,0	203,8	277,2	196,8
s	48,2	38,6	27,9	20,2
Combined target-tasks times	33,3	27,6	26,1	21,7
s	8,8	5,9	7,4	2,8

Average results: Test progress 2 (NT3, NT4, NT1, NT2)				
NTs:	NT1	NT2	NT3	NT4
Find 7 bricks with 2 knobs	192,7	112,3	136,4	104,1
s (standard deviation)	84,6	33,2	18,9	29,6
Find 7 wrongly placed…	186,9	103,6	130,8	93,1
s	54,7	13,7	8,2	14,4
Combined find-tasks times	379,6	215,9	267,2	197,1
s	133,2	43,8	11,3	26,6
Combined target-tasks times	57,0	29,5	28,9	24,9
s	22,8	3,3	7,7	5,1

Average results: Test progress 3 (NT4, NT3, NT2, NT1)				
NTs:	NT1	NT2	NT3	NT4
Find 7 bricks with 2 knobs	198,6	120,2	132,3	172,6
s (standard deviation)	72,2	43,4	43,3	41,9
Find 7 wrongly placed…	165,1	115,9	139,1	132,4
s	40,6	49,6	32,8	39,3
Combined find-tasks times	363,6	236,2	271,4	305,0
s	104,9	68,8	65,9	73,4
Combined target-tasks times	44,2	30,5	28,1	26,7
s	4,9	6,4	3,9	1,7

Average results: Test progress 4 (NT2, NT1, NT4, NT3)				
NTs:	NT1	NT2	NT3	NT4
Find 7 bricks with 2 knobs	202,0	156,2	157,3	134,0
s (standard deviation)	56,7	54,1	40,9	60,2
Find 7 wrongly placed…	161,9	119,0	113,0	108,3
s	24,6	23,9	14,9	11,4
Combined find-tasks times	363,9	275,2	270,3	242,2
s	76,1	73,6	26,9	62,5
Combined target-tasks times	40,0	27,8	25,8	26,6
s	13,9	8,3	8,7	10,0

Appendix C: t-Test results from Experiment 2

All results are calculated using Microsoft Excel's Data Analysis Toolpack

t-Test: Paired Two Sample for Means
NT1 vs NT2 (find-tasks)

	NT1	*NT2*
Mean	357,76	232,76
Variance	7822,88	3494,64
Observations	16	16
Pearson Correlation	0,4481	
Hypothesized Mean Difference	0	
df	15	
t Stat	6,14	
P(T<=t) one-tail	9,46E-06	
t Critical one-tail	1,75	
P(T<=t) two-tail	1,89E-05	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT1 vs NT2 (target-tasks)

	NT1	*NT2*
Mean	43,61	28,84
Variance	242,69	32,56
Observations	16	16
Pearson Correlation	0,5765	
Hypothesized Mean Difference	0	
df	15	
t Stat	4,50	
P(T<=t) one-tail	2,13E-04	
t Critical one-tail	1,75	
P(T<=t) two-tail	4,26E-04	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT2 vs NT4 (find-tasks)

	NT2	*NT4*
Mean	232,76	235,28
Variance	3494,64	4174,94
Observations	16	16
Pearson Correlation	-0,0009	
Hypothesized Mean Difference	0	
df	15	
t Stat	-1,15E-01	
P(T<=t) one-tail	4,55E-01	
t Critical one-tail	1,75	
P(T<=t) two-tail	9,10E-01	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT2 vs NT4 (target-tasks)

	NT2	*NT4*
Mean	28,84	24,93
Variance	32,56	31,58
Observations	16	16
Pearson Correlation	0,5753	
Hypothesized Mean Difference	0	
df	15	
t Stat	2,99	
P(T<=t) one-tail	4,55E-03	
t Critical one-tail	1,75	
P(T<=t) two-tail	9,09E-03	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT3 vs NT4 (find-tasks)

	NT3	*NT4*
Mean	271,53	235,28
Variance	1207,94	4174,94
Observations	16	16
Pearson Correlation	0,2716	
Hypothesized Mean Difference	0	
df	15	
t Stat	2,25	
P(T<=t) one-tail	2,00E-02	
t Critical one-tail	1,75	
P(T<=t) two-tail	4,01E-02	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT3 vs NT4 (target-tasks)

	NT3	*NT4*
Mean	27,23	24,93
Variance	42,91	31,58
Observations	16	16
Pearson Correlation	0,7133	
Hypothesized Mean Difference	0	
df	15	
t Stat	1,96	
P(T<=t) one-tail	3,46E-02	
t Critical one-tail	1,75	
P(T<=t) two-tail	6,92E-02	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT2 vs NT3 (find-tasks)

	NT2	*NT3*
Mean	232,76	271,53
Variance	3494,64	1207,94
Observations	16	16
Pearson Correlation	0,2322	
Hypothesized Mean Difference	0	
df	15	
t Stat	-2,53	
P(T<=t) one-tail	1,15E-02	
t Critical one-tail	1,75	
P(T<=t) two-tail	2,30E-02	
t Critical two-tail	2,13	

t-Test: Paired Two Sample for Means
NT2 vs NT3 (target-tasks)

	NT2	*NT3*
Mean	28,84	27,23
Variance	32,56	42,91
Observations	16	16
Pearson Correlation	0,6798	
Hypothesized Mean Difference	0	
df	15	
t Stat	1,30	
P(T<=t) one-tail	1,07E-01	
t Critical one-tail	1,75	
P(T<=t) two-tail	2,14E-01	
t Critical two-tail	2,13	