

<b>1</b>	<b>INTRODUCTION .....</b>	<b>2</b>
1.1	PREVIOUS RESEARCH.....	2
1.1.1	<i>Newspaper reading</i> .....	2
1.1.2	<i>Online news reading</i> .....	4
1.2	THEORETICAL BACKGROUND AND HYPOTHESES.....	4
<b>2</b>	<b>DATA COLLECTION.....</b>	<b>5</b>
2.1	QUESTIONNAIRE.....	6
2.2	READING SESSION.....	6
2.3	RETROSPECTIVE SESSION.....	7
<b>3</b>	<b>DATA ANALYSIS.....</b>	<b>8</b>
3.1	QUESTIONNAIRE DATA.....	8
3.2	EYE MOVEMENT DATA.....	8
3.3	RETROSPECTIVE DATA.....	11
<b>4</b>	<b>RESULTS AND DISCUSSION.....</b>	<b>11</b>
4.1	PRIORITY.....	11
4.1.1	<i>Discussion</i> .....	14
4.2	AREAS OF INTEREST.....	15
4.2.1	<i>Discussion</i> .....	18
4.3	READING BEHAVIOUR.....	19
4.3.1	<i>Discussion</i> .....	22
4.4	READER'S INTENTIONS AND EXPECTATIONS.....	22
4.4.1	<i>Discussion</i> .....	23
4.5	A SUMMARY OF QUALITATIVE DATA FROM THE RETROSPECTIVE SESSION.....	24
<b>5</b>	<b>GENERAL DISCUSSION.....</b>	<b>25</b>
5.1	DISCUSSION ABOUT STUDY DESIGN.....	26
5.2	FUTURE RESEARCH.....	26
<b>6</b>	<b>ACKNOWLEDGEMENTS.....</b>	<b>26</b>
<b>7</b>	<b>REFERENCES.....</b>	<b>26</b>

Behaviour in online news reading

Maria Barthelson

Department of Cognitive Science, Lund University

## 1 Introduction

Traditional publishers of journalistic material have found a new channel through which they can reach the readers namely the Internet. The material is often the same in the online version as in the paper based version, but not enough thought has been paid to how the different media influence the way of reading. The Internet as medium imposes new criteria as to how the content should be published. This study has investigated how readers actually read news published online. How do the readers' intentions and expectations influence their reading behaviour?

Eye-tracking technique has been used to get reliable data on how the reader processes and traverses through the material. Eye movements are tightly connected to low-level cognitive processes. This means that eye movements are involuntary. The eyes may be diverted by reflex to areas with high contrasts such as strong colours, blinking objects etc. However, high-level processes, guided by the users' specific intention or interest for their visit, have a strong impact and effectively control what is actually attended to. This means that a persons intentions and interests guides their visual attention. Due to this, eye movements can tell us about what a person is thinking (Hansen, 1994). The high-level processes are in focus in this study.

This paper starts by reviewing eye tracking studies on newspaper reading and online news reading undertaken during the last decade to depict the research history in the A report of the methodological lay out for this study follow. Then the results of the eye tracking experiment take place. The results describe in which order the readers prefer to traverse the content and which areas they find most interesting. Different reading behaviour will be described and discussed.

### 1.1 *Previous research*

Research in the area of news reading has been undertaken during the nineties. The studies on online news reading have provided some insights to the aspects of online news design, but they have not been high on experimental control or statistical significance. Previous research in newspaper reading and online news reading relevant for this study will be reviewed here.

#### 1.1.1 Newspaper reading

Widman and Polansky (1990) conducted a study on advertise reading in traditional newspapers. They found that colour and photos increased the observation of advertisements. Photographs and pictures more than anything attract attention first. Placing an ad in the middle of a page increases the amount of attention more than any other placement.

This implies that content placed in the middle will be more noticed than content placed in other areas. It also implies that colours will increase the observational value of the coloured object and that photos will attract attention first.

This is verified in an eye track study on newspaper reading conducted by Garcia & Stark (1990). They found that readers do not read newspapers. Instead, readers scan newspapers searching for an interesting piece to read. The scanners enter the page wherever the most powerful element is, like a photo.

Traditionally the inverted pyramid has been used for news layout. When this design principle is used the most important news are placed in the top area and the lower on the page a reader travel, the less important the content is. Garcia & Stark (1990) found that this design technique did not facilitate scanners, who enter the page wherever the most dominant element is such as a photo or a headline. The readers follow trails that editors lay for them. Additionally, the readers look at the right-hand page first and then travel to the left.

Garcia and Stark provided new insights to how newspapers were really processed by the readers. However, They did not investigate what the users bring to the page in form of interests and familiarity with a specific topic.

This study aimed at, shedding light over the high-level processes involved in the choice of topics pursued. Moreover, it was looked at if the inverted pyramid was a wrong design template also in online news design and how the readers tended to traverse the material. The readers' order of traversing the material should mirror the prioritisation implied by the editor's layout trail. Hansen (1994) have further developed this idea in his newspaper reading study. According to Hansen, a good layout will ensure that the reader in a minimum of time can assimilate the most interesting information at hand. The reader will be able to prioritise the information quickly. A badly thought out layout would result in the reader overseeing relevant information, tending to loose himself/herself in details or let the gaze randomly shift.

The implications for this study was that the prioritisation implied by the layout would be the actual prioritisation perform by the readers when traversing the material. Eye movements should align to the designer intentions. If the layout is not satisfying as a guide for attention, the fixations are randomly placed on the different content areas of the sites.

Hansen (1994) also found that readers do not read full articles. In his study only the shorter articles were fully read. He suggests that the articles are kept short.

The above mentioned is what has been done regarding research of traditional newspaper reading. These results are used as a base for this study.

### 1.1.2 Online news reading

If little research has been done in news reading, even less research effort has been made investigating reading of online news. Lewenstein, Edwards, Tatar and DeVigal (2000) conducted a study on online news reading. They had a similar focus as Garcia & Stark. The study has been subjected to some criticism due to the lack of experimental control (Jacobson, 2000). From a scientific point of view, their findings do not qualify as being predictive in other areas, merely to provide a descriptive base for further research. However, the study is strong in ecological validity.

Lewenstein et al's (2000) major conclusion was that text is the preferred entry point over graphics among online newsreaders. In contradiction to the Garcia and Stark study, the online newsreaders here fixated on briefs and captions first and not the most dominant element in the page, such as photos and other graphics.

Regarding the reading behaviour, Lewenstein et al (2000) found that online readers read shallow but wide while pursuing selected topics in depth. This resembles to the Garcia & Stark findings that readers do not read newspapers, they scan them in their hunt for something interesting to investigate further.

Lewenstein et al(2000) as Garcia & Stark (1990) raises a few questions for future research to answer regarding the high-level cognitive processes influencing the reader behaviour. Both research teams raise the question how the readers' background knowledge and personal interests influence the readers' choice of topics to pursue. Lewenstein et al (2000) wonder whether the readers, when they hyperlink to a related story, do return to the original first page.

In this section, the research efforts made in reading news published in both traditional papers and on the Internet have been summarised. The findings described above serve as a base for the hypotheses for this study presented in the following part.

## 1.2 *Theoretical background and hypotheses*

In this part, the hypotheses will be presented. The main points of this study to explore were which areas, such as textual or graphic areas, attract attention first and which receive most attention. How do the readers traverse the material? What role does high-level processes, like intentions, expectations and previous knowledge and interests, play?

Hansen (1994) suggests that a well thought out layout will help the user prioritise the material. If this is the case, a clear priority order will appear in the data analysis. This result in the following hypothesis for this study:

**Hypothesis 1:** The readers will follow the priority implicated by the layout. If there is a clear order in when different areas get attention the layout has facilitated the readers' prioritisation of the content.

Lewenstein, Edwards, Tatar and DeVigal's (2000) study of online newsreaders suggested that the text was the most attracting element in a news site. This is contradicted by the Garcia & Stark news paper reading study, which states that photos and headlines get the attention first. Widman & Polansky (1990) showed that the middle region of a page gets most attention. These results predict that the majority of the attention will be given to the areas in the middle of the page where the headlines and photos reside together with the text.

**Hypothesis 2:** The readers will give most attention to the content in the middle area. Headlines, captions and photos will receive attention early.

The explorative part of the study investigated the occurrence of different reading strategies. As pointed out by Garcia & Stark (1990), readers do not read newspapers. They search until they find an interesting piece of news to read. In online published news the reader does not have the possibility to flip through the entire news site as in a traditional newspaper. This implies that the readers will use different strategies to browse through the online material in search of something interesting to read. Some readers may compare headlines in link lists on the first page before deciding on an article to read and consequently return to the first page. Another possible type is the reader who plunges in to the text and continues reading associated articles without returning to the start page. The "text readers" might directly dive in to the textual areas while the "link viewers" invest much time in investigating link lists of headlines.

**Hypothesis 3:** Different reading strategies will emerge. Some readers prefer to compare headlines on the first page before deciding on an article to read and return to the start in-between article selections. Another reader type prefers to plunge in to an article and continue reading associated articles without returning to the start page. The reading strategy will stay consistent over different stimuli.

The influence of the readers' personal interests and background knowledge in article selection has so far not been thoroughly investigated. It may be presumed that the readers' personal interests influence their choice of articles to read. According to Hansen (1994) the readers' expectations and intentions influence their scanning behaviour.

**Hypothesis 4:** The readers' intentions and expectations influence their scanning behaviour and their personal interests influence their choice of articles to read.

## 2 Data collection

This section will describe the methodological structure for the study.

The experiment was conducted at the eye-tracking laboratory at the University of Lund. It was undertaken during one session in November 2001 with 12 participating subjects. The subjects had an average age of 28,4 years. Some were students and others were employed. They had various levels of Internet and computer skills. All subjects were familiar with online news.

The study consisted of three phases. (1) Filling in a questionnaire (see part 2.1), (2) the reading session (see part 2.2) and (3) the retrospection and follow up interview (see part 2.3). The total experiment time per subject was approximately 50 minutes to one hour.

## 2.1 Questionnaire

The initial part consisted of letting the subjects fill in a questionnaire about their age, gender, computer and Internet skills and conversance with online news. It took about 10 minutes to fill in the questionnaire and the main purpose was to gain some background data to provide additional insights to user behaviours. The questionnaire can be viewed in full in appendix xxx.

An overview over the subjects is presented in the table below.

Subject no.	Age	Gender	Occupation	Computer experience	Internet experience	News online exp.	Reading interval
1	26	Female	Student	Very Good	Very Good	1-2 years	> 1 time/ week
2	27	Male	Student	Very Good	Very Good	> 2 years	> 1 time/ week
3	24	Female	Student	Quite Good	Quite Good	0-6 months	< 1 time/ week
4	26	Female	Working	Very Good	Very Good	> 2 years	> 1 time/ week
5	26	Male	Student	Quite Good	Quite Good	6-12 months	< 1 time/ week
6	28	Male	Working	Very Good	Very Good	> 2 years	> 1 time/ day
7	26	Male	Working	Very Good	Quite Good	1-2 years	1 time/ day
8	21	Female	Student	Quite Good	Quite Good	6-12 months	< 1 time/ week
9	62	Male	Working	Very Good	Very Good	> 2 years	> 1 time/ day
10	29	Male	Working	Very Good	Very Good	> 2 years	> 1 time/ day
11	24	Female	Working	Very Good	Very Good	> 2 years	1 time/ day
12	22	Male	Student	Quite Good	Quite Good	6-12 months	> 1 time/ week

Table 1 Subject overview

## 2.2 Reading session

The second part incorporated the reading session where the subjects' interaction and eye movements were recorded using eye-tracking technique. The eye tracking data is the main data source for this study. The eye tracking equipment used was an SMI iView 50 Hz pupil-corneal reflex video imaging system. It records the eye movements of subjects by filming their eyes with a camera located underneath the computer screen. This made it possible to get precise information about where on the screen subjects are looking. The gaze point is superimposed onto an image of the browser window and recorded both onto file and onto video. The latter to be used in the retrospective session.

The supporting equipment used in the study was a 17" computer screen and a 120 MHz Pentium II PC with a 10 Mbps-Internet connection. A browser that was specially built to compensate for scrolling during the session was used. In every other aspect, the browser allowed the subjects to interact as they would with an ordinary browser. The browser took a screenshot each time the subject scrolled or clicked on a link in the page. The screenshots were saved to a file and used later as background scenes in the fixation data analysis.

Each subject was instructed to find a comfortable position in front of the screen and to avoid moving the head during the reading session. The subject was told that he/she was going to read news on two different sites for five minutes on each while his/her eye movements were recorded. The experimenter let the subject know that he was keeping track of the time. The subject was then guided through a calibration procedure by the eye-movement technician. In the next step, the subject was asked to follow either the link to [www.dn.se](http://www.dn.se) or to [sydsvenskan.se](http://sydsvenskan.se).

The subject was allowed to freely indulge in whatever type of news he/she found interesting during the five minutes. When the first five minutes of reading has passed the subject was asked to switch to the other news site by returning to the start page and follow the next link<sup>1</sup>.

Different stimuli were used in the study to explore design implications on reader behaviour. Two different online news providers were selected. Both of them are traditional newspapers with large reading groups. Sydsvenska Dagbladet was one of the selected stimuli. Dagens Nyheter was the other news provider selected. Sydsvenska Dagbladet is the biggest daily paper in the Malmö region and Dagens Nyheter is a big daily paper in the Stockholm region. The two newspaper sites were selected because of their different publishing strategies on the web. The design and layout on the web-based versions differ from each other. (SDS will hence refer to the online version of Sydsvenska Dagbladet. DN will be used henceforward when referring to the online version of Dagens nyheter.). The total reading session lasted for about 20 minutes including instructions, calibration procedure and the reading of the news sites for five minutes each.

When the second reading session was over the subjects was given the video tape recording and continued on to the third test phase, the retrospective session, conducted in another room.

### **2.3 Retrospective session**

The purpose of this session was to validate the eye tracking data. The data collected in this session is to regard as a complement to the eye movement data, which is the main data source in this study. As Hansen et al (1991) points out, the eye movements can not be guaranteed to mirror the cognitive processing that occurs in the interaction. The retrospection gives information on the users interpretation of the interaction. This information combined with the eye movement data gives us information on what the subject was thinking, when he/she was looking at a certain point, and why.

The method used in this study was the retrospective protocol advocated by Hansen et al (1991) over the more common simultaneous verbal protocol. The latter interrupts the subject in the currently performed task, which is a draw back. In the retrospective protocol the eye gaze point superimposed over the stimuli is used as a memory cue, helping the subject remember the interaction (Hansen et al, 1991).

The subject together with a test leader watched the video recording of the subject's interaction on a 28" television, with the eye movements superimposed onto the video of the computer screen. Each subject was instructed to continuously comment on his/her actions and intentions while the video recording was playing.

Finally, a semi-structured interview was undertaken to investigate the subjects' opinions about the experiment and their attitudes towards different news media, and DN and SDS in particular. The interviews also provided data about the user experience from their natural occurring interaction with online news. The questions can be found in appendix xx.

---

<sup>1</sup> The time constraint of five minutes was sufficient according to the subjects' statements about the length of their naturally occurring online news reading sessions. The subjects claimed that they usually spent approximately 6 minutes in a news browsing session.

During the session, the test leader took notes on the subjects' comments. The total retrospective session lasted about 20 minutes.

This part of the paper described the structure of the study and the purpose for the different phases in the study.

### **3 Data analysis**

In this section, the processing and analysis of the collected data for the different test phases will be described.

#### **3.1 Questionnaire data**

The questionnaire was used for receiving quantitative data on the subjects' backgrounds such as age, gender, Internet and computer skills and conversance with online news. This data was as a complement to the eye movement data recorded during the reading session and has been used in correlation analysis on how these factors influence reading behaviour.

#### **3.2 Eye movement data**

To read a text or process a picture in any detail, focus or *fixation* is needed. It is only during a fixation that detailed information can be acquired. A fixation lasts for about 200-600 ms (Glenstrup & Egnell-Nielsen, 1995). Pre-attentive and attentive cognitive processes determine where to fixate in a visual scene. The movement of the eyes to a different area of the visual stimuli is called a *saccade*. Saccades are sudden rapid movements of the eyes, which takes about 30-120 ms to complete. No visual information is acquired during a saccade. Visual attention in the peripheral area of our vision selects the next location to be further attended to and a saccade is made to enable the eye to fixate on the new item.

To study eye movements is to study cognitive processes. Eye movements provide information about what the subject potentially processes. Different eye movement patterns can tell us which type of processing is occurring. Reading for example, has an easily recognisable pattern. To read a text, one has to fixate more or less on every word in it. This makes reading easy to distinguish from other eye movements.

When using eye movements to analyse what people are processing it is necessary to keep in mind the fact that people can look without seeing and see without looking. Looking without seeing occurs when you find you have read a whole page but not remembering what it was about. Seeing without looking occurs when you have registered information in the periphery without moving the eyes or the head.

Usually we look and see. However, eye-tracking data can not assure to a 100% that this is the case, but we presume that this is true (Hansen, 1994). To validate the data in this study, a verbal retrospective protocol was used to ensure that the subjects really saw what they were looking at.

To be able to analyse which type of information got the most fixations and in which order the readers preferred to scan through the material an object set for each paper has been defined.



The objects were defined based on the type of information in that particular area. The stimulus is used as a background to the different areas defined.

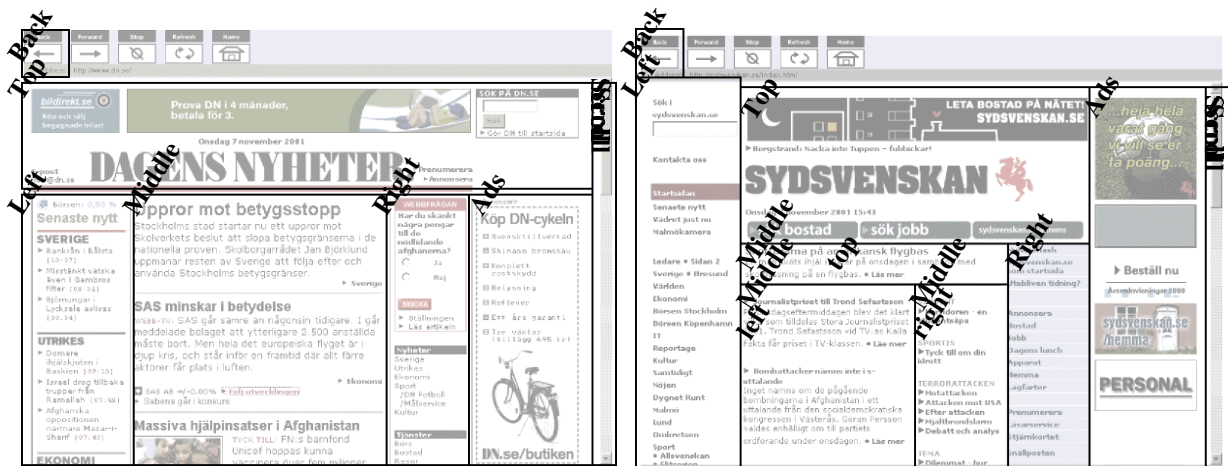


Figure 2: First page of DN and SDS respectively 2001-11-07

Area name	DN content	SDS content
Back	Browser back button	Browser back button
Top	Header, ads and search box	Header, ads and topic links
Left	Headline links	Topic category links
Middle	Top stories with headlines, briefs with a link to the full story.	
Right	Topic category links and diff. Services available	Links to the weekly polls and diff. Services available
Middle top		Top story brief and link to the full story
Middle left		Other top story briefs with links to the full story
Middle right		Headline links
Ads	Advertisements	Advertisements
Scrollbar	Scrollbar	Scrollbar

Table 2: The different areas in DN and SDS and their content

These areas were then used for calculations in the following analysis.



Figure 3: Example of fixation patterns

Figure 3 shows the fixation patterns for one subject overlaid onto one stimulus. This data was then used for analysis of the total amount of fixations in a particular area, which will be referred to as Areas of Interest in the result section. Areas of Interest illustrate the areas that attracted the most amount of attention.

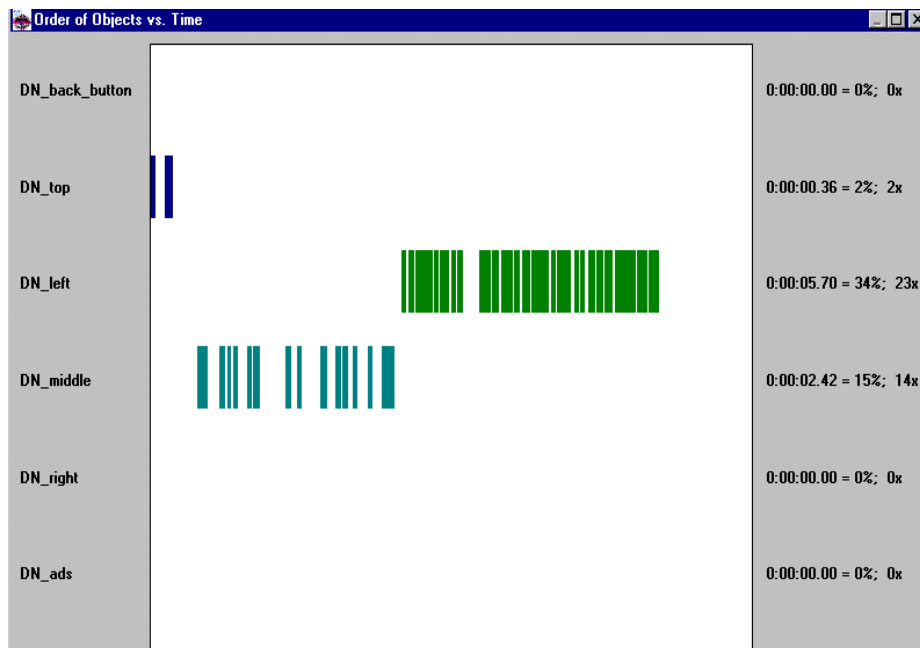


Figure 4: Example of time vs. object and fixations

The areas defined in previously described analysis were also used for counting transitions between different areas. Figure 4 shows the fixation density on each object and which area that was attended to at a certain point in time. The transitions were analysed for each subject and each site. The total number of transitions was counted for each site. The transitions were also summarised for certain areas for all subjects. This data representation was used to analyse attention transitions between objects and was then used as a basis for statistical calculation of the results.

The eye movement data was recorded onto both file and video. The latter to be used in the retrospective session.

### **3.3 Retrospective data**

Further, a retrospective session has been undertaken. The data from this procedure was used as validation and an aid for data interpretation. It provided insights to the subjects' naturally occurring interaction with online news and their opinions of online news, the traditional news media and their expectations of the content structure. The data is qualitative and was collected by note taking. Subjects' comments will be included in the result discussion when relevant.

## **4 Results and discussion**

### **4.1 Priority**

**Hypothesis 1:** The readers will follow the priority implicated by the layout. If there is a clear order in when different areas get attention the layout has facilitated the readers' prioritisation of the content.

Hansen (1994) pointed out that the layout implies a certain priority to the news. The most important piece of news is placed in the upper part of the page and added a large headline. Garcia & Stark (1990) referred to this conduct as the "Inverted pyramid". This means that important news are placed in the top and less important pieces is placed further down on the page and given less space, reflecting a well thought out prioritisation order. This design principle is strongly condemned by Garcia & Stark for newspaper layout. Due to it being unable to facilitate scanners.

To investigate the matter of scanning priority in our study the all ready defined an object set for each paper's first page was used. The data shows that design and layout influences the readers' way to scan the material on the first page in their first visit. The order, in which the subjects attended the defined areas (5 for DN and 8 for SDS), was calculated by giving the first area that was fixated the value 5 and 8 respectively. The second area fixated was given a 4 for DN and a 7 for SDS and so on until all fixated areas was given values. The areas that were not fixated at all were given 0 as value. Then an average value for each area was calculated over all subjects.

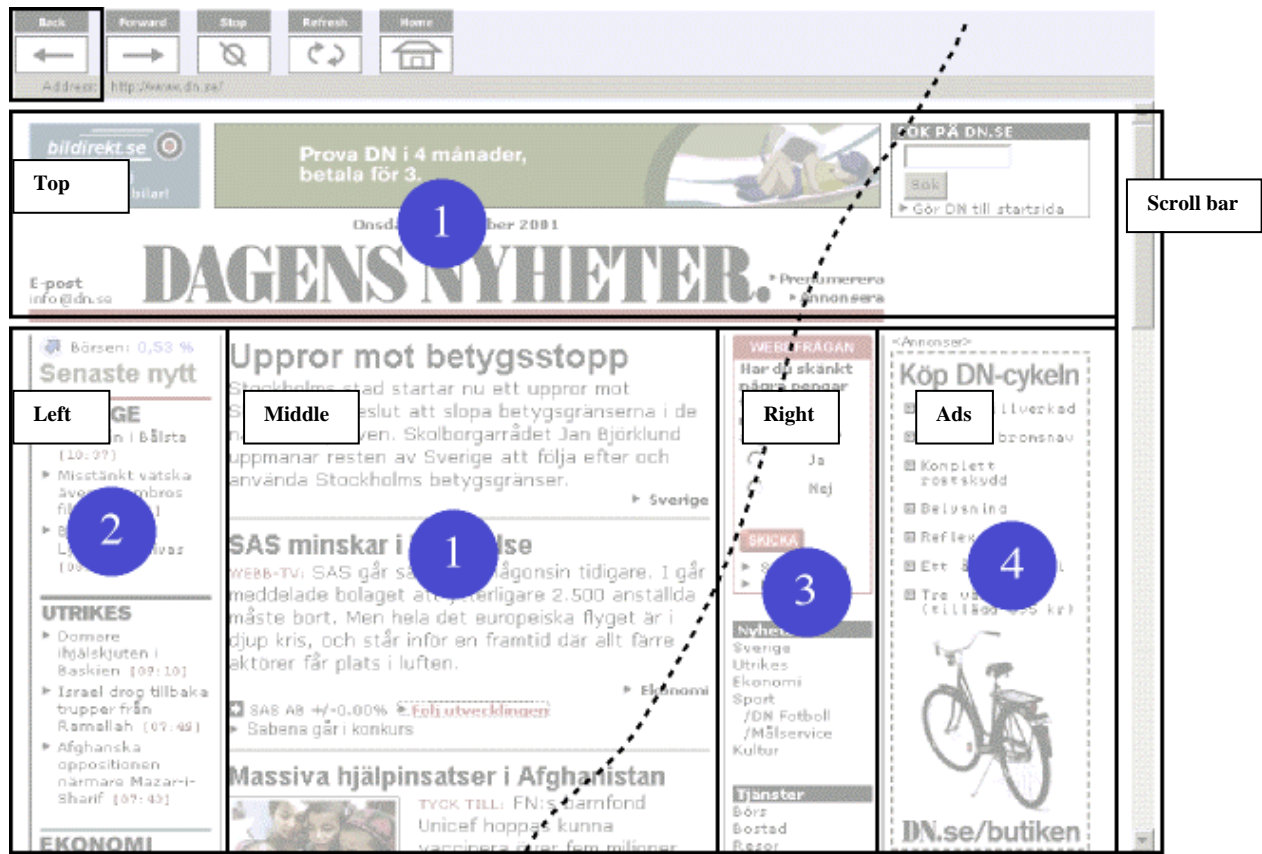


Figure 5: Priority and significance value for DN

	Order	Left	Middle	Right	Ads
Top	5,00	ns	ns	p 0,01	p 0,01
Left	3,08		ns	Ns	p 0,05
Middle	4,67			p 0,01	p 0,01
Right	1,58				Ns
Ads	0,83				

Table 3: Order and ANOVA, Tukey HSD p-values for priority for DN. ns = non significant

The findings for DN show a clear order of scanning priority between the different areas defined. There is a significant difference in priority between the upper left part of the page and the lower right area of the page. The readers paid attention earlier to the upper left part than, if at all to the lower right area. The dashed line in figure XX approximately illustrates between which areas the priority difference is statistically significant.

The top and the middle area competes about the first fixations, a part from that, there is a clear order in which the readers prefer to scan the material. They start at the top or directly in the middle and then travels left to the headline links. This supports Garcia & Stark (1990) findings that readers travel from the right to the left when reading news.

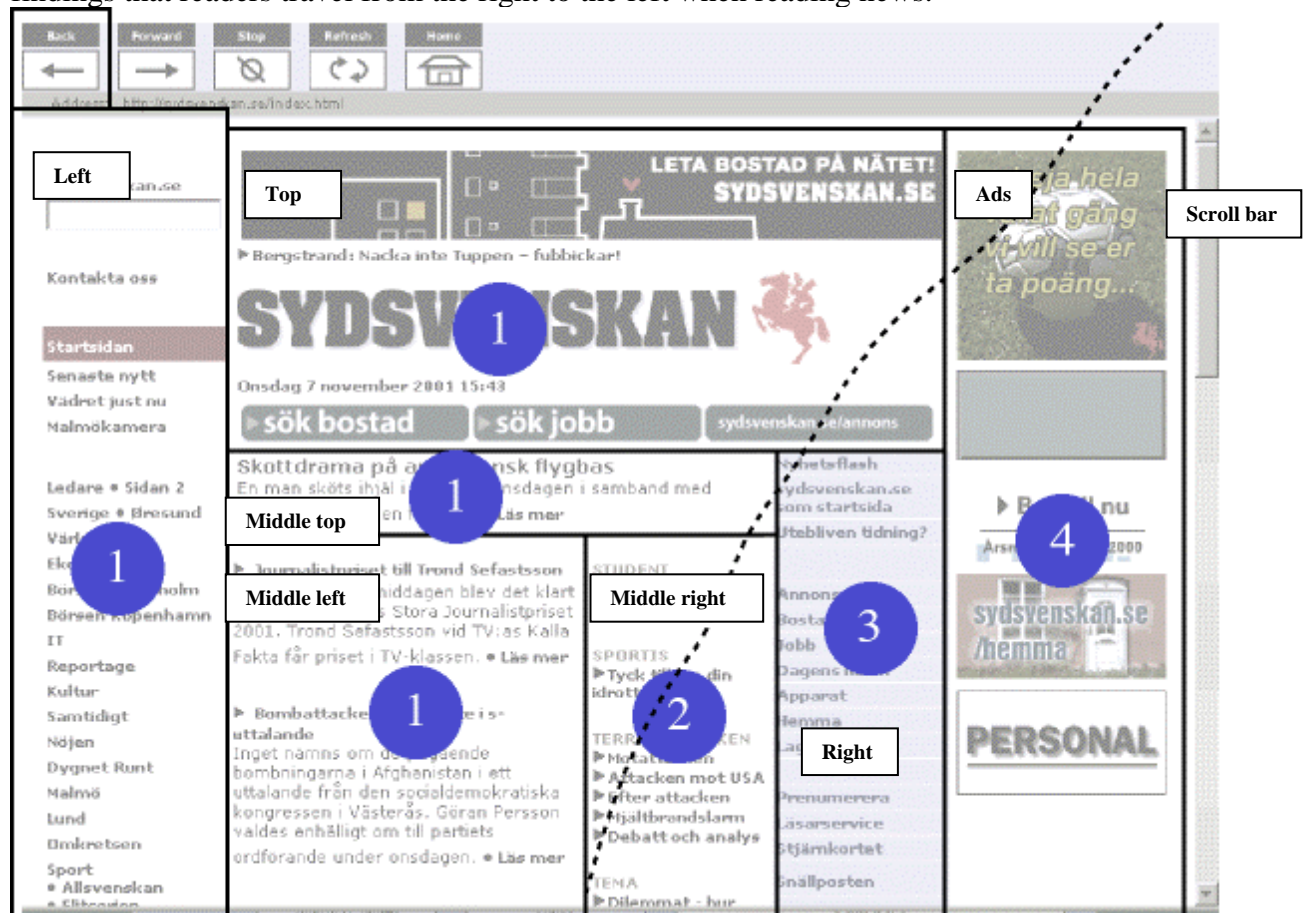


Figure 6: Priority and significance for SDS

	Order	Top	Ads	Right	Middle-top	Middle-left	Middle-right
Left	5,58	ns	P < 0,01	P < 0,05	ns	ns	ns
Top	5,83		P < 0,01	P < 0,01	ns	ns	P < 0,05
Ads	0,92			Ns	P < 0,01	p < 0,05	ns
Right	2,17				P < 0,01	P < 0,01	ns
Middle-top	5,75					ns	P < 0,05
Middle-left	5,83						P < 0,05
Middle-right	2,83						

Table 4: Order and ANOVA Tukey HSD significance values for priority for SDS

In the SDS case the scanning order was not as clear between the different areas as in DN. Some readers spent their first fixations in the top area and others in the middle area while others still looked at the left link area first. However, there was a significant difference in priority between the upper left area and the lower right part of the screen. The readers fixated significantly earlier in the left parts than, if at all, in the right area. The dashed line in Figure 6 illustrates where the significant difference is.

Summarised, the upper left part got attention significantly earlier than the lower right part of the page in both DN and SDS, but the order in which the subjects scanned the material is only apparent in DN.

#### 4.1.1 Discussion

According to Hansen (1994) the layout in ordinary newspapers guides the readers attention to the most important news first. The results from our study presented above verify the first hypothesis stating that the readers will follow the priority implicated by the layout. If there is a clear order in how different areas get attention then the layout has facilitated the readers' prioritisation of the content.

The clear order for traversing the material in DN shows that DN has been successful in helping the readers focus on the most informative areas first. DN guides the readers through the material in a more effective manner than SDS, which does not accomplish this. The reason would lie in the design differences of the sites.

DN has four columns and the largest area on the screen is the text area with big headers and the major news in the top. SDS has chosen a different layout strategy with five columns of equal size. The different information areas compete about the reader's attention and the use of an extra column adds one extra competitor to the game. The probability of a certain area to get the reader's attention decreases with the amount of added information. When more information is added the demand for a clear design to help guide the users increase.

Comments from the subjects state that DN has a clearer layout than SDS. This was one reason for 8 of 11 subjects preferring DN to SDS. Our findings further suggest, in support of the findings of Lewenstein et al (2000) that text is the main attraction in an online news page. This contradicts Garcia & Stark's (1990) claims that photos and graphics serve as entry points for reading. Their objections to the inverted pyramid as a design template, due to its poor ability to facilitate scanning, may be true regarding traditional newspaper layout. However, when publishing news online the information space is constricted to a small area, which additionally is subjected to scrolling. Only a small part of the whole material on the first page is at display at one time which makes the inverted pyramid an appropriate model for design of the material. This is confirmed by the subjects' comments. They expect to find the biggest news in the top of the page. Additionally, they state that they prefer to have the article presented in different levels, starting with a brief and at the next level the shorter online version of the article is presented. The subjects claim that they would even like a third level with the full-length newspaper version of the article available. The readers read few articles but the ones they read they want in full length.

In this section, results regarding the readers' prioritisation when scanning through the material have been presented. The results verify the hypothesis that a good layout will guide the reader through the content. In the next part the results regarding the distribution of attention over the

areas defined will be declared and discussed. This will show which areas that were of most interest to the readers.

## **4.2 Areas of interest**

This part describes how the attention transitions between the different pre defined areas were made and it will begin with the results on which areas that were of most interest to the readers.

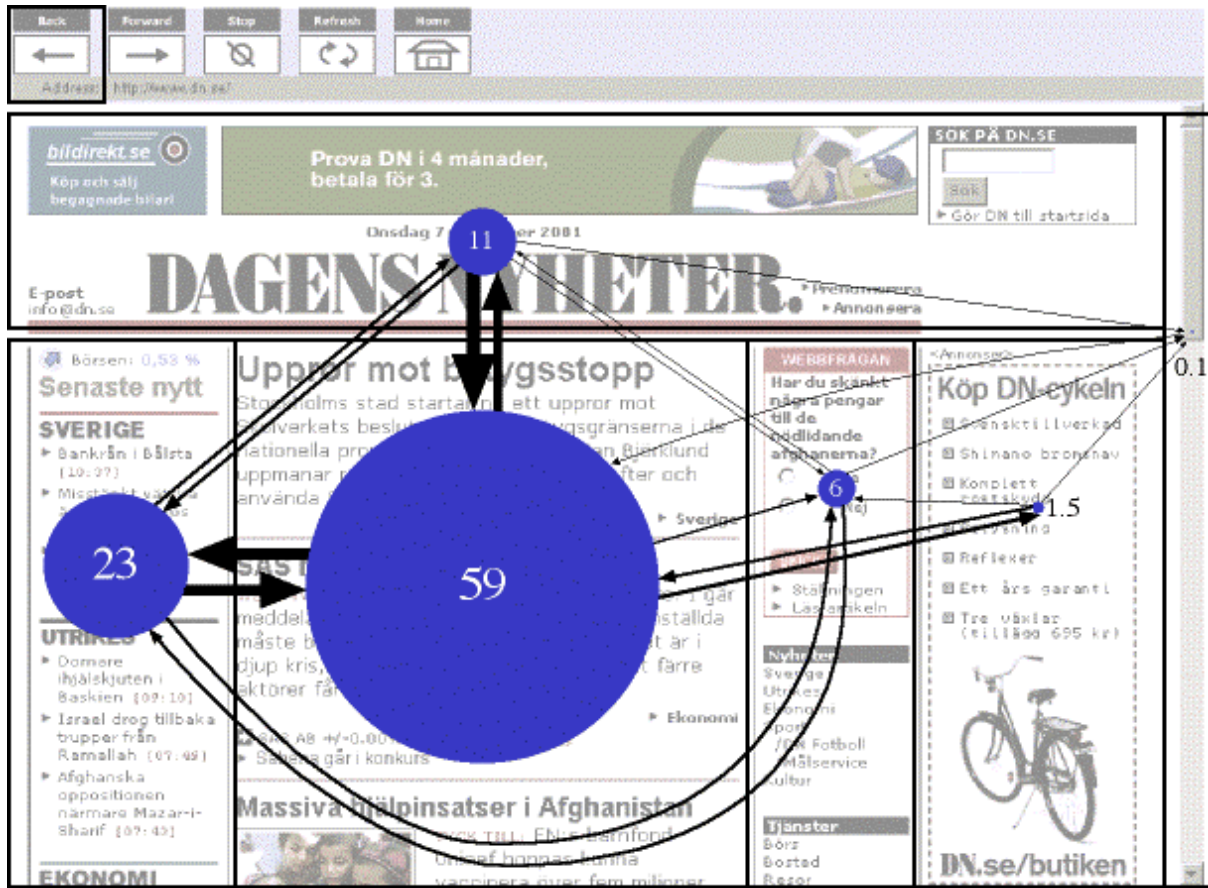
**Hypothesis 2:** The readers will give most attention to the content in the middle area. Headlines, captions and photos will receive attention early.

The same object definitions as in the priority analysis for calculation which areas received the most attention was used also in this analysis. The figures are calculated from the fixation data of the first page of the site. The first page was of most interest regarding layout strategies in that the first pages are the most different when comparing the two sites.

A summarisation of the total fixation times for the defined areas shows that the areas received different amounts of attention. In our study 59 % of the total fixation time is spent on the middle areas, which includes headlines, briefs and photos. This excludes all link list areas, ads and graphic areas such as the top. This percentage is the same for both sites tested.

To illustrate the different percentages for the areas of interest defined, the notation that Itoh, Hansen and Nielsen used in their ship navigator study (1987) was appropriate to illustrate the transitions. The size of the circular area is relative the total amount of fixation time received by the underlying area. The figure in or near the circles represents the percentage.





See Figure 7: AoI and transitions for DN

	Top	Left	Middle	Right	Ads
Top	0,0	4,4	17,5	0,9	0,0
Left	3,5	0,0	14,0	3,5	0,0
Middle	11,4	14,9	0,0	2,6	5,3
Right	0,9	3,5	2,6	0,0	1,8
Ads	0,0	0,0	4,4	1,8	0,0

Table 5: Transition percentages for each transition in DN

The ANOVA, Tukey HSD shows a statistically significant difference regarding the amount of attention between the textual content against all the other areas with  $p < 0.01$ .

The illustrations also show the transitions between information areas. This is done with arrows where the width of the lines represent the percentage of transitions made from one objects to another relative to the total number of transitions during the reading of the first page. The transitions are interesting because they give further insights to how the readers traverse the content of the first page. The priority order was established in part 4.1. This section will describe how the attention travels between the areas. The illustrations will show which transitions that are most common, between which areas they occur, and where they do not. This also connects to which way the reader chooses through the content and how the transitions differ over different sites.



Regarding the link areas, there is a large difference between how much attention the link lists receive. The left-hand links receive almost four times as much fixations as the right-hand links. When asked about this in the retrospective session the subjects stated that they expected to find the most important and interesting information in the middle or to the left. The right-hand links was their last choice when they did not find anything interesting in the text material in the middle or amongst the headline links to the left.

As the illustration shows, the ads get very little attention only 1,5 % of the total fixations. Only two individuals made the fixations on the ads. It was the same subjects fixating on ads in both papers.

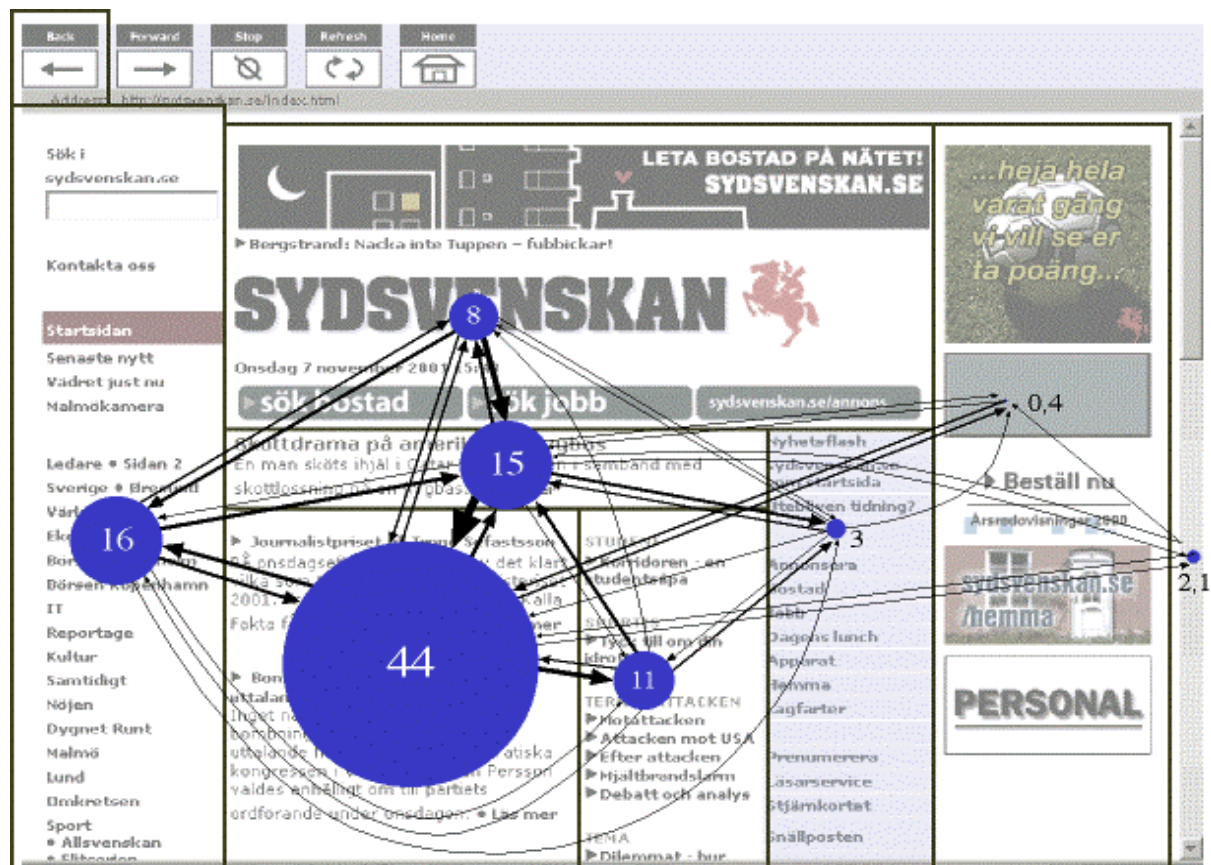


Figure 8: AoI and transitions for SDS

	Left	Top	Ads	Right	Middle-top	Middle-left	Middle-right	
Left		0,0	3,0	0,0	0,6	4,8	4,2	0,6
Top	3,6		0,0	0,0	0,6	7,8	2,4	0,0
Ads	0,0	0,0		0,0	0,0	0,6	3,0	0,0
Right	0,0	1,2	0,6		0,0	3,0	0,6	1,8
Middle-top	1,8	4,2	0,6	4,2		0,0	11,4	2,4
Middle-left	5,4	3,0	2,4	0,0	4,8		0,0	6,0
Middle-right	1,2	1,2	0,0	3,0	3,6	2,4		0,0

Table 6: Transition percentages for each transition in SDS

There is a statistically significant difference regarding the amount of attention between the textual content in the middle against all the other areas (ANOVA, Tukey HSD  $p < 0.01$ ) also for SDS.

In the SDS case, the total amount of attention has to be divided between more areas due to the layout strategy. The majority of fixation time is spent in the middle text area also in SDS. The fixation time percentage is equal to DN (59%) when the top middle part (15%) is added to the text area in the middle (44%). These areas contain the same type of information.

The left-hand link area receives more attention than the links in the area called middle-right. This difference is however smaller than in DN. The most interesting finding however is that the right-hand link area only gets 3% of the total amount of attention.

Regarding the transitions, the findings show a difference between the two news sites in the experiment. DN has fewer transitions (an average of 9,5 over 7 objects) compared to SDS (an average of 13,8 over 9 areas). The overall transition pattern in DN is coherent between the defined areas. The main transition paths are between top-middle and top-left. This indicates that the attention pattern is governed by the layout, which additionally supports the first hypothesis.

In SDS the transitions are scattered over the defined areas. The most frequent transitions take place between top and middle top and from middle top to middle left. However, these trails are not the highways of attention they are in DN. The readers appear to have a hard time deciding where the most relevant content resides on the screen.

There was a strong complementary relation between the reading area and the left link area (-0.91). This indicates that the more time a reader spends looking in the left-hand link area the less attention is paid to the reading area and vice versa. This relation is stronger for reading in DN than in SDS. In SDS this correlation (-0.77), is not as strong as in DN. This correlation cannot be found between the reading area and the right-hand link area.

#### 4.2.1 Discussion

The results show that the middle area of the page receives the most attention. This verifies the second hypothesis. The text attracts most attention in the page as suggested by Lewenstein, Edwards, Tatar and DeVigal (2000).

The reason for the large amount of fixations in the top may partly be due to page loading differences. When DN's first page is loaded, it happened during the session that only the top area was present for a few seconds. When this occurred, the readers of course will fixate on this area because this was the only element present. Due to this, the percentage got high. This suggests that ads placed in the top are more likely to be seen than ads placed in the right-hand side of the page.

The attention pattern that the illustration describes shows that transitions between the middle area and the right-hand link areas never are made. This together with the low percentage of the total fixation time and the subjects' statements, that the right links are seen as the last

possibilities when looking for interesting information, further verifies that important information should not be placed in the right-hand side of the page.

The complementary relation between the left-hand links and the reading area indicates that a reader pays attention trying to find interesting content by reading the article briefs in the reading area. When the readers continue searching for another article to read they look for it in the left-hand link area. The fact that the right-hand links are not a part of this equation further strengthens the findings that the right-hand link list is not a prioritised area amongst the readers. It is also interesting to consider how this pattern is the same over the sites despite the contrary placed link elements. DN has the latest headlines to the left while these are placed in the right-hand links in SDS. The stronger correlation in DN web may indicate that the readers prefer to have headlines in the left-hand link list to content categories as in SDS.

When comparing the actual content and layout in DN and SDS (see table xx in part 3.2), one sees that DN has chosen to put headline links to the left, while SDS has chosen to put category links to the left. The category links is placed to the right in DN, while this area in SDS contains the news headline links. This may be one explanation to the lower percentages in SDS's left-hand link area. The corresponding content in DN is placed in the right-hand link list.

It seems that the readers do not find the category links appealing. The subjects hunt for news headlines. This is confirmed by comments in the retrospective interview session. The subjects were asked about their usual situation and purpose for reading online news. They unanimously claimed that the purpose is to quickly find out if anything has happened. This results in a ruthless hunt for interesting headlines and briefs.

In this part, results on which areas and type of content the readers found most interesting and how they searched the site for a story to read have been presented.

### **4.3 Reading behaviour**

In this section the results from the explorative part of the study will be presented. The results will be based on the quantitative data collected in the questionnaires and the eye tracking session combined with the qualitative data from the retrospective session.

**Hypothesis 3:** Different reading strategies will emerge. Some readers prefer to compare headlines on the first page before deciding on an article to read and return to the start in-between article selections. Another reader type prefers to plunge in to an article and continue reading associated articles without returning to the start page. The reading strategy will stay consistent over different stimuli.

When the reading patterns were compared over the subjects, tendencies regarding subjects' reading behaviour appeared that was constant over the sites tested. The total amount of fixations made in the reading area in the middle was calculated for the start pages. The same was done for the left-hand area of the page containing links. The subjects were then categorised depending on their reading behaviour for each site. The result shows that some subjects have a strategy that is constant over the different sites. The strategy can be either to give most attention in the left-hand link list or to focus mostly on the text in the middle hence forward referred to as the reading area. Some subjects repeated the same behaviour for the

first page and through the whole content of the site. This strategy can be constant over the both sites but in some cases it shifts so that a “text viewer” in DN becomes a “link viewer” in SDS and vice versa.

To analyse the data in search for evidence of this hypothesis, the total amount of time spent in the middle area, containing the text, and the total amount of time spent in the left link list for the first page, was calculated

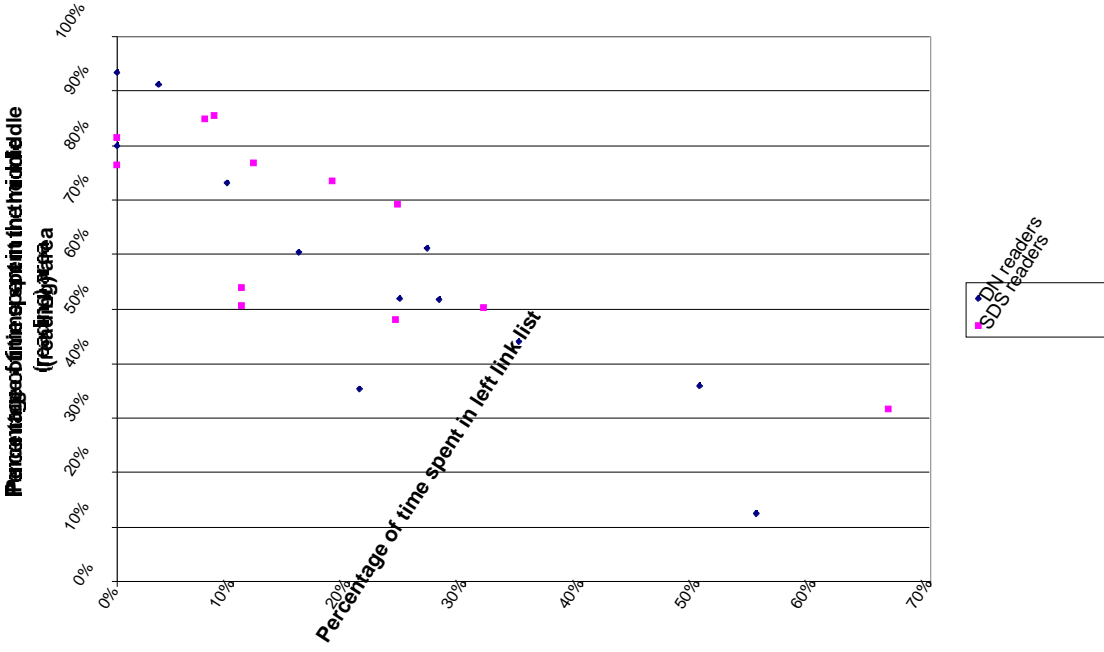


Diagram 1

Diagram 1 suggests that it might be possible to divide the readers in to two groups. One that resembles to our predefined definition of a text viewer who prefers to search for interesting articles in the middle reading area. The other type spends most time systematically searching the left-hand link lists.

However, the data on the first page shows strong individual differences. Due to this the total amount of time spent in the areas during the whole reading session for each paper was summarised.

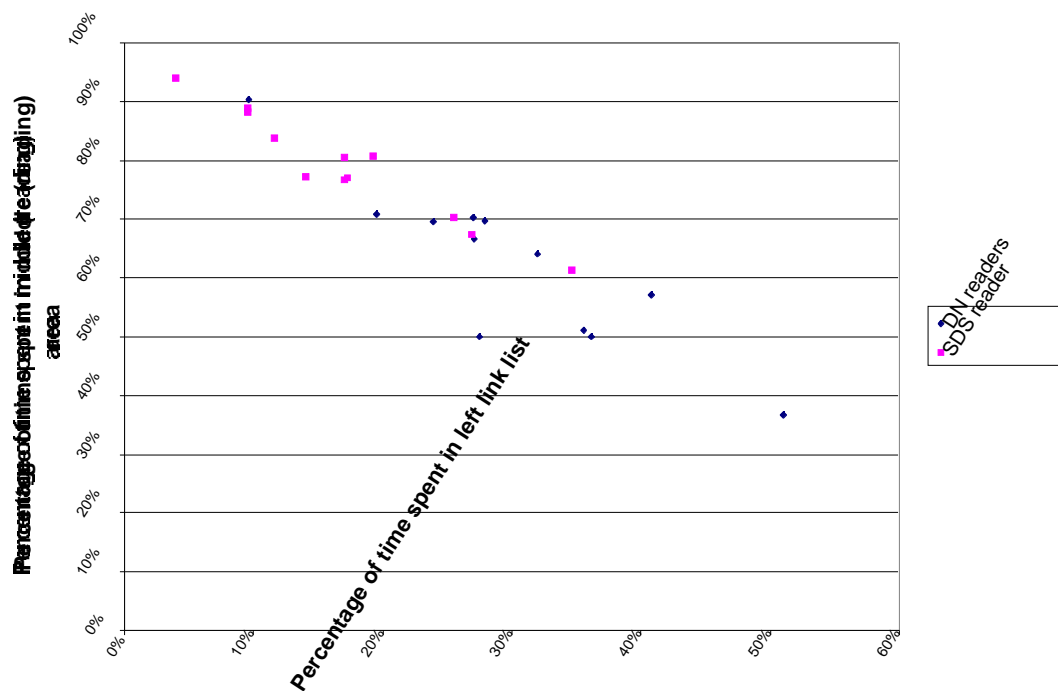


Diagram 2

Diagram 2 shows an even clearer line and thereby clearer support for the two reading strategies. It is also apparent that the links play a more important role in DN than in SDS.

The data shows that two fourths of the subjects are clear “text viewers” or alternatively clear “link viewers” and the remaining half consists of strategy shifters. Much effort was made to find a correlation between reading behaviour and variables such as age, gender, computer skills, Internet experience and customisation to online news, but it was in vain. No strong correlation to the factors tested could be found. An additional regression analysis only supported the non-existent correlation.

These results suggest that different strategy and behaviour exists, as posed in the third hypothesis. However, the reading strategy is not related to variables like age, gender, computer skill, Internet experience or online news experience. If any correlation exists, it would be based on other factors than those tested in this study.

To compare reading behaviour for online news to ordinary news publications data from a previously undertaken eye tracking experiment on news reading conducted at LUCS eye tracking laboratory in 2000 was used. The data was run through a filter that recognise reading patterns in eye movements. The result showed that actual reading is undertaken 55% of the time. The net reading data was run through the same filter and the figure for net reading was 44%, which was significantly less (independent, two-tailed t-test,  $p = 0.002$ ). This implies that more time is spent scanning in net reading compared to traditional newspaper reading.

As an answer to the question regarding the readers browsing strategy, posed by Lewenstein et al: It is noteworthy that 9 out of 12 subjects preferred to return to the first page after exploring the details of an article. When asked about this behaviour the subjects stated that they wanted

to return to the first page to assure that they did not miss anything. When they probed deeper into an article, they felt that they had left the main trail and in order to have total overview of the content they felt a need to use the first page as a starting point.

#### 4.3.1 Discussion

The results presented here partly verify the third hypothesis. It appears that different reading strategies exist among the readers. Some readers prefer to search the textual content for an interesting article to read, while others tend to focus their search efforts to the left-hand link lists. However, the reading strategy does not always stay consistent over different stimuli or over the total news site content. There are readers that consequently use the same strategy over different stimuli, but there are equally as many that shift strategy over the to sites. This indicates that some readers are more easily influenced by the layout strategy. If a comparison is to be made, DN to a higher degree encourages searching headline links.

An analysis was made to evaluate if any background factors such as age, computer skill etc. influenced the reading style. The subject pool participating in our experiment was small and had a quite homogenous constitution. However, I did not manage to find any correlation between reading strategy and factors like, age, gender, computer skill, Internet or online news experience. Other factors have to be tested to find the underlying causes for the different reading strategies.

The difference between reading a traditional newspaper from online news is apparent. The findings from this study show that when the reader is not reading the textual content he/she is scanning through the left-hand link list.

The majority of the subjects do return to the first page after clicking in to a full article. This suggests that the readers make a mental representation of the whole content and that the first page is the point from where all roads take a start. The preference for returning to the first page in-between articles should be facilitated in the design with clear links back to the first page. This was not the case in DN where no one of the subjects got that the paper logo in the upper left part of the page was the “home”-link. The logo’s link function is not obvious by only looking at it. Instead, the subjects used the back-button repeatedly, which caused some irritation.

#### **4.4 Reader’s intentions and expectations**

In this section focus will be on how the reader’s intentions, expectations and interests influence their reading behaviour. These results are based on the quantitative data collected in the questionnaires and the qualitative data from the retrospective session.

**Hypothesis 4:** The readers’ intentions and expectations will influence their scanning behaviour and their personal interests will influence their choice of articles to read.

According to the subjects’ statements the main purpose for a visit to a news site is to quickly get updated on anything that might have happened during the day. The usual situation for online news reading, is when a few minutes break turns up in their work tasks. The subjects state that they usually spend about six minutes on a news browsing session, then they usually

were interrupted and returned to their work. If you only have a short time, you tend to focus on finding the most interesting piece of news. This might also explain the large amount of time spent scanning headline link lists. The average news browsing time stated by our subjects is much shorter than the average 34 minutes spent on the online news reading session in the Poynter study (Lewenstein et al 2000).

The fixation patterns show that the readers give the right parts of the screen significantly less and later attention than the upper left parts. This is true for both sites tested. When asked about this in the retrospective interview the subjects stated that they did not expect to find anything interesting in the right-hand area. In this area they expected to find less relevant information such as ads and unimportant links.

Regarding the influence of the readers' personal interests in their choice of articles to read, the results from this study show that the readers do explore topics based on their own interests and knowledge. The law student read articles about legislation propositions and court cases, the woman interested in women's rights read articles about women oppression in Islam, the teacher assistant read about the grade system and the IT-consultants read about IT-related news. Additionally this study supports the findings of Lewenstein et al (2000), that all subjects read disaster and scandal coverage. The topics that generally were shown interest by all subjects were the ones about nature catastrophes, accidents and scandals. Such as people killed by tsunami, a car accident and a dialysis scandal.

#### 4.4.1 Discussion

The results presented in previous section verify the fourth hypothesis. The readers' expectations and personal interests play an important role in guiding their attention.

The readers do not find it necessary to give attention to information in the areas they expect to be of less relevance. They expect to find the most informative element in the top, middle and to the left. This is the normal way to distribute information on the web. It is standard space disposition, to which the readers seem to have adapted. This presumption lies behind their attention distribution on the site. It seems as a continuous use of the Internet have taught the users how the prototypic site is designed information wise. The important links are placed to the left and the textual content is placed in the middle. The regular users have adopted this content structure and focus their attention accordingly on the screen.

This support Widman & Polansky (1991) findings that the middle and slightly left area are the best areas to place information to be sure it gets attention.

The results showed that the readers' personal interests do influence their choice of articles to read, nonetheless some topics are common selection by most readers. These topics are crises, scandals and catastrophes as claimed by Lewenstein et al (2000).

These findings suggest that high-level cognitive processes have a major influence on the reader behaviour. Low-level processes stimulated by blinking objects in the periphery do not seem to interfere with the readers' goals and intentions. With this in mind, it appears doomed to place important information elements to the right on the screen. This content is likely to be missed by the readers.

#### **4.5 A summary of qualitative data from the retrospective session.**

Online news reading is very different from the situation in which the subjects read ordinary newspapers. The subjects find traditional newspaper reading relaxing and like to have plenty of time to sit down with a cup of tea or coffee and read the paper from cover to cover. The average time spent on an online news reading session in the Poynter study was stated to be 34 minutes. This is not supported by this study where the subjects stated that they usually spent about six minutes on a news browsing session when a few minutes break turned up in their work. Then they usually were interrupted and had to return to their work tasks. The short time does not encourage long reading sessions. If you only have a few minutes, you tend to focus on finding the most interesting piece of news.

The screen reading is another factor that influences their reading behaviour. The majority of the subjects claimed that they find reading from a screen cumbersome and therefore prefer to keep their reading sessions short. Additionally, the media it self seems to influence the reading behaviour. One subject stated that she generally read restlessly on the Internet. She felt that the Internet implicated a quick run-through.

In the subjects' opinion, online news does not replace more traditional news channels such as TV, radio and printed media. The online news is seen as a complement to these channels. They all fill different needs. The online news is used to quickly get updated on what has happened in the world. Its main advantage is the quick access and that it is constantly updated. Regarding disadvantages, the subjects mention the fact that it is hard to read from the screen and that the usually short articles does not provide enough depth on a topic.

Regarding the readers preference in article length they stated that when they do click in to an article the like to have it all. The scenario is as follows: The reader scans a list of the latest news where each piece is presented with a headline and the brief. The reader finds an interesting piece and clicks in to it. The article version provided in this stage is the often-shorter online version. If the reader finds this article fascinating, he/she wants to know even more and then would like a link in to the full-length version as printed in the ordinary newspaper.

The readers also like the associated article links to the one they are currently reading. This is one of the biggest advantages with news online as the subjects saw it. The total topic history is available at their fingertips. They can search for old articles on the subject and additionally see what different providers have published on the matter and thereby have the possibility to easily compare different sources.

When asked how much they would be willing to pay for online news services the more experienced computer and Internet users laughed and said that they would not pay anything, they would just go to another provider. The less experienced users showed some hesitation and said that they might consider paying for the service as long as it did not cost more than the traditional paper version.

Regarding preference for one of the sites tested 8 of 11 subjects stated that they preferred DN to SDS. The subjects' favouritism of DN is partly based on the layout. The subjects find DN easier to overview with a clearer strategy. SDS is perceived as jumbled. The subjects ask for more images and colour. Some readers state that they would like to see more colours in DN



and that the current version is too serious and they like the use of more colours in SDS. Others think this makes SDS feel less serious than DN.

## 5 General discussion

The four hypotheses formulated for the study has with one exception been verified. The readers do follow the priority implicated by the layout. And a good layout facilitates the readers' prioritisation of the content. Additionally, the readers do give most attention to the content in the middle area, as suggested by the second hypothesis.

Regarding the third hypothesis this has been partly verified. Different reading strategies did appear. Some readers preferred to scan link list in their search for interesting information and others preferred scanning the textual areas directly. However, there was no correlation between the different strategies and factors such as age, gender, computer skills, Internet experience or conversance with online news.

Another finding was that the readers' personal interests and intentions do influence their choice of articles to read and their scanning behaviour, in support of hypothesis four.

Additionally the readers seem to have adapted to how information is usually placed on web pages and they show a strong tendency of following their presumption about the information structure in their search for interesting information. This study tells us that it is a good idea to follow this standardised information design to assure that the users quickly find the information they want where they expect it to be.

The ruthless search for interesting information gives the editors little time to catch the reader's attention. It suggests that the most important aspect of news design is to provide interesting headlines, as Lewenstein et al (2000) also claimed. The scanning possibilities in online news reading can be equalised to searching lists of headlines. Some readers prefer this behaviour. And providing headlines that well mirror the content of the article should facilitate this. This is also suggested by Lewenstein et al (2000) in the Poynter study. However, the subjects in present study stated that the headline alone often did not give them enough information to decide if the article was worth reading and then discarded it and continued the hunt for the ultimate piece. They prefer to get a brief in support for Garcia & Stark (1990) who claimed that briefs were big also with their readers.

The subjects participating in the study did not mind scrolling long pages. In fact, they almost immediately started scrolling the page. Some subjects never read below the horizontal middle on the page. When they reached the middle they started scrolling.

According to Garcia and Stark (1990) the inverted pyramid is not useful as a design principle in traditional news design, because it does not facilitate the scanners. In online news design the implications of scrolling must be considered. The readers have to scroll in order to view the content and only a little piece of it can be viewed at a time. This supports the use of the inverted pyramid in online news design. The subjects also expect the content to be ordered in this manner.

## **5.1 Discussion about study design**

One dilemma caused by using a table mounted eye movement camera is that the subject at all times has to face the screen and not makes any big head movements or the camera loses track of the eye. This fact combined with the inhibiting laboratory environment compromises the ecological validity of the data. None the less, most subjects stated that the test situation was quite unobtrusive, some mentioned that they felt a bit strained by the need to keep their heads still. A few subjects stated that they felt a bit stressed due to the time limit, which they were made aware of in the beginning of the session. None the less, the subjects claimed that they on average spent about six minutes on each news browsing session. It might have been better not to inform the subjects of the time constraint to avoid the stressing the subjects' interaction.

## **5.2 Future research**

As described, there was a clustering of individual reading behaviour for either text viewers or link viewers. However, this is not statistically significant. Future research may further explore on the underlying causes for this. One interesting point of view to explore is the implications of personality traits on the reading behaviour.

The high-level cognitive processes guided by the readers' intentions and expectations seem to have large influence on the reading behaviour. This study has shown that the readers' intentions and expectations highly influence their reading behaviour. They manage to stay focused on finding something interesting to read in spite of the frequent flicker of advertisement animations in the periphery, provoking the low-level cognitive processes influencing the eye gaze. The high-level process control in visual search tasks in online reading should be interesting to additionally investigate.

## **6 Acknowledgements**

I would like to thank Kenneth Holmqvist and Jana Holsánova, Department of Cognitive Science, Lund University, for their theoretical and practical help. I would also like to thank all subjects for taking time to participate in the study and also my fellow students who reviewed this report and provided useful comments.

## **7 References**

Clark, M et al (1995): *Effektiv kommunikation med multimedia*. Docendo Läromedel, Göteborg

Garcia, M. R. & Stark, P. (1991): *Eyes on the news*. The Poynter Institute for Media Studies, St. Petersburg, Florida

Lewenstein, M., Edwards, G., Tatar, D. and DeVigal, A. (2000): Poynter Eyetrack Study, <http://www.poynter.org/eyetrack2000>

Hansen, J. P. (1991): The use of eye mark recordings to support verbal retrospection in software testing. *Acta Psychologica* 76, p. 31-49

- Hansen, J. P. (1994): *Analyse af læsernes informationsprioritering*, Forskningscenter Risø, Roskilde
- Itoh, K., Hansen, J. P., Nielsen, F. R. (1998): Cognitive modelling of a ship navigator based on protocol and eye-movement analysis, *Le Travail Humain*, no 2/1998 p. 99-127
- Jacobson, A. (2000): Projects Too Much From Too Little, <http://www.poynter.org/centerpiece/071200alan.htm>
- Widman, L. & Polansky, S.H. (1990): *Annonsläsning. En ögonrörelseundersökning av DN-läsare*. Unpublished report. Stockholm: Dagens Nyheter