Chapter 30

Reading or Scanning? A Study of Newspaper and Net Paper Reading

Kenneth Holmqvist, Jana Holsanova, Maria Barthelson, and Daniel Lundqvist

Net paper readers have been shown to read deeper into articles than newspaper readers. It has also been claimed that newspaper readers rather scan than read newspapers. Do these findings mean that net paper readers read proportionally more than newspaper readers? This paper presents results showing that in fact net paper readers scan more and read less than newspaper readers. We furthermore investigate whether this result can be explained by the differences in layout, navigation structure and purpose of reading between the two media.

Introduction

Today, we use the Internet for several purposes: to get information, entertainment and to do errands. During February 2002, 5,080,000 Swedes were connected to the Internet, which is 58.2% of the whole Swedish population above 2 years of age (80% of the age group 18–49 years).¹ Sixty-two percent of the Swedish Internet users state that they read news and net papers online. This activity is second to e-mailing (84%).² In 1994, Aftonbladet.se opened the first net paper site in Sweden. In February 2002, this site had 1,836,000 unique visitors.

News providers try to adjust to the trend and are present on the Internet. Due to the competition to catch the readers’ attention, news providers are anxious to understand the behaviour of net paper users. Some years ago, they could only get feedback from logs and online surveys. Nowadays, they conduct eye-tracking studies.

Traditional printed media exist in parallel to the new media. Reading traditional newspapers implies looking for headlines, briefs, photos and drop quotes. Does the new medium influence our way of searching for news? Does net paper reading differ from newspaper reading?

¹ The Mind’s Eye: Cognitive and Applied Aspects of Eye Movement Research
Copyright © 2003 by Elsevier Science Ltd.
All rights of reproduction in any form reserved.
On the one hand, the new media differ considerably from the old: Online readers read from a computer screen and move around by clicking on links and menu buttons. They navigate through a virtual space and may run into problems orienting themselves in the complex net structure (see De Léon & Holsanova, 1997).

On the other hand, readers themselves report on important differences in the use of the traditional and new media. In Barthelsson (2002), experienced online news readers describe the medium differences as follows: Reading a newspaper is something they do with pleasure and if possible in a situation that allows few distraction (along with breakfast, in a coffee break after lunch, on the train or in the subway). It is a relaxing activity to traverse through the folds and it usually takes quite a long time. In contrast, reading a net paper is something you do in much shorter breaks, perhaps between two e-mails, usually in your office in the early morning, or during lunch. The purpose is to become updated on a few particular questions. The users make brief visits to the news sites several times a day with the expectation of obtaining a quick overview over the latest events.

A net paper consists of a front page (also called the first page), topic pages (sometimes called section pages) and article pages. A page is a two-dimensional surface containing text and images that can be accessed by scrolling up and down in a web browser window. The information for each page is stored at a different URL.

All these two-dimensional surfaces are then linked together in a complex multi-dimensional structure. By hyper-linking, hitting the back button, choosing a different

Figure 30.1: The front page of the DN net paper (www.dn.se).
bookmark, or typing in a new web location, the readers can move around in this virtual space. Net papers usually organise their pages in a hierarchical tree structure, with the front page as the root and the topics and stories branching out from it.

**Previous Research**

So far, only a small number of explorative and experimental eye tracking studies that concern newspaper and net paper reading have been conducted. One of the main questions has been what item on a page first catches the readers’ eye. This issue of the so-called priority order on a page has often been studied (see Barthelsson, 2002; Garcia & Stark, 1991; Hansen, 1994; Lewenstein et al., 2000). The distribution of fixations in different regions has been focused on in several papers (see Barthelsson, 2002; Hansen, 1994, Widman & Polansky, 1990). Sometimes the transitions between regions and between pages have also been analysed (see Lewenstein et al., 2000) to find out how readers navigate in the virtual space. Last but not least, the processing and reading of newspaper advertisements has been investigated (see Lundqvist & Holmqvist, forthcoming, Widman & Polansky, 1990). These studies are all concerned with layout questions, such as the position of various graphical elements on a page, the role of photographs versus text, or the role of colour in the layout. Apart from the explorative studies by Garcia and Stark (1991) and Lewenstein et al. (2000), the overall reading behaviour has often been investigated as a by-product of these results. The focus was on results that had an immediate relevance for designers and journalists.

Widman and Polansky (1990) tested 129 readers of the Stockholm newspaper *Dagens Nyheter* for advertisement reading. The unpublished internal report shows that 39% of all ads are seen, and the bigger the ad, the more likely that it is seen and remembered. Widman and Polansky also investigates the importance of different positions and contents of newspaper ads, noting that a position slightly to the left of the middle is optimal for the smaller ads. Photographs, pictures and colour in ads are reported to increase fixation frequency on the ad.

Lundqvist and Holmqvist (forthcoming) tested 14 readers of *Dagens Nyheter* in a follow-up study to Widman and Polansky, but focusing entirely on the effect of size on perception, memory and attitude towards ads. It is concluded that these variables strongly correlate, and that size and attitude data explain around 50% of the variance in the perception and memory data. Ad content and design were not taken into account as a variable.

Garcia and Stark (1991) tested 90 readers of three newspapers at three different sites in the US. The editions (prototype A and B) given to the readers were manipulated with respect to colour so as to show whether colour is an attractor that can direct the reader. Participants could read as long as they wished and their eye movements were tracked during the reading session. A videotape with newspaper pages and the tracks of the eye movements across the pages were used to find out whether elements on the page were processed, read or read in depth. The material was considered “read” if the reader’s eyes moved across one or more lines of print from left to right. When at least one half of any text was read, it was considered “read in depth.” “Processing”
was rather vaguely defined: the reader’s attention stops long enough at an individual element for information to be acquired. These results give an immediate feedback to journalists and news designers, but the lack of precise definitions and measures make them difficult for researchers to interpret.

A major finding in Garcia and Stark’s study was that readers do not really read but rather scan newspapers. At certain entry points they stop scanning and start reading the story that the entry point belongs to. Pictures and graphics were identified as the main entry points, followed by front-page promotion boxes. Readers usually enter the page through the dominant photo and then move to a prominent headline or another dominant photo. But virtually all elements, anywhere, even editorial text, can serve as entry points into reading. This, the authors argue, is strong evidence against the classical inverted information pyramid. The inverted information pyramid tells us that important information should be given most space, and placed at the top, while gradually less important information should be given less space further down. Instead they speak in favour of a creative design using more graphical elements on the folds.

Garcia and Stark also report about processing text and photos. Headlines, cut lines and briefs are processed often and in depth. Only 25% of the articles are processed. Text processing is highest in the news section, lowest in the sports. Readers devote more time to photo groupings when they are in colour. Size increases the attraction to a photo. Garcia and Stark’s discussion ends by defining newspaper design as the task “to give readers material that is worthy of their scan, that makes them stop scanning and start reading” (ibid., 1991: 67).

Hansen (1994) studied 12 readers of the Copenhagen newspaper Det Fri Aktuelt. He investigates the order in which objects on folds were scanned and the readers’ priorities with respect to text. He found that pictures are first seen, then icons and graphics followed by headlines of different sizes and text, with form items at the bottom of order hierarchy. Hansen’s hierarchical order reflects that of Garcia and Stark’s entry points.

When investigating readers’ priorities, Hansen considered the length of articles, their placement and their genre (news, features, debate, sport, etc.). He measured how many centimetres of an article were read by a subject and calculated a depth index (the amount of text that has been actually read as a percentage of the whole text length) and a total response index (depth index as a percentage times the number of readers). His results show that only short articles are fully read. The longer the article, the smaller the proportion of it will be read. In this respect, Hansen’s results resemble the result in Garcia and Stark (1991) showing that a mere 25% of all articles are seen, and only 12% are read deeper than half of their length. Concerning article placement, subjects were most engaged in reading on pages 2–6 and after that, their interest in reading decreased with the exception of the last page of the newspaper. Stories on the left-hand side of the fold were seen significantly earlier than stories on the right hand side. Articles with cut lines had significantly better total response values than neutral articles. News, features and debates had approximately the same total response whereas sport articles had significantly less. Hansen stresses the importance of designing the newspaper layout so that it quickly leads the reader to the information that the designer wants to emphasise.
Lewenstein, Edwards, Tatar and DeVigal (2000) conducted an explorative eye-tracking study investigating naturally occurring reading behaviour of 67 users at different US news sites. Like its newspaper predecessor (see Garcia & Stark, 1991) study, the net paper follow-up has a broad scope and merely descriptive statistics. The subjects (experienced users of online newspaper sites) were asked to read in the manner they usually read, for as long a session as was typical for them. Their eye movements were recorded by the SMI EyeLink eye tracker. After the reading session, subjects were interviewed about their media habits. The screens viewed were coded by format (headlines, articles, briefs, photos, etc.), by topics (science, medicine, national, international, etc.), by action (keyboard events) and by visit and time spent. Eye fixations were overlaid on screen dumps of the pages. Using this kind of data, they studied front page entries, fixation order on the pages, returns to front page and story reading order.

The study of Lewenstein et al. has been subjected to some criticism due to the lack of experimental control (e.g. 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.

A major conclusion drawn by Lewenstein et al. (2000) is that text is the preferred entry point among online newsreaders. First fixations on the front page of a net paper go to text (78%) rather than to photos or graphics. The online newsreaders fixate on briefs and captions first and not on the dominant pictorial elements in the page. This result very much contrasts with the previous findings from Garcia & Stark and Hansen that pictures and images are the foremost entry points in newspapers (see Jacobson, 2000 for criticism of this result).

Scanning Versus Reading

Several researchers focused on the question whether newspapers and net papers are read or scanned. Results in Hansen (1994) show that only short articles in a newspaper are fully read. As pointed out by Garcia and Stark (1991), readers do not really read newspapers. They search until they find an interesting piece of news to read. Only 25% of all articles are seen, and only 12% are read deeper than half of their length. In contrast, net paper stories are read to a depth of 75% on average (see Lewenstein et al., 2000). Net paper readers, as Lewenstein et al. conclude, go deeper into articles than newspaper readers.

Does the deeper reading found in net papers also mean that net paper readers read proportionally more than newspaper readers? Do newspaper readers have to scan more to find the stories they want to read because of the differences in how information is laid out in the two media? In the first experiment, we want to test the first hypothesis that readers of newspapers scan more than do readers of net papers.

Hypothesis 1: More reading and less scanning in net papers.
Experiment 1

We made two recordings of eye movement data from readers of two net papers and two newspapers.

Design of Net Paper Reading Recording

Twelve subjects read the net papers Dagens Industri (www.di.se) and Aftonbladet (www.aftonbladet.se), which are net versions of a financial and an evening paper. The readers were experienced with these net papers. Recordings were made during one day in November 2000, using the SMI iView remote set at 50 Hz. Subjects read from a regular 120 MHz Pentium computer with a 100 Mbit Ethernet connection and a 17 inch screen. Each subject read both net papers in a regular browser (Explorer). Eye movement data was recorded in the stimulus screen co-ordinate system (1024*768).

Subjects were instructed to read each net paper for as long as they liked. They were asked to read in the same way as they normally read these net papers. They had to stay within the pages belonging to the paper, but could read these pages in any order they liked. We planned to let our subjects switch to the other net paper after 15 minutes, but virtually all were bored and stopped reading after 5–8 minutes. In post questionnaires, subjects stated that the presence of the eye tracker did not influence their reading.

Design of Newspaper Reading Recording

Fifteen subjects read the European newspaper Metro (Stockholm edition) and 14 readers the Stockholm newspaper Svenska Dagbladet, which are both morning papers. Each subject read one newspaper. Recordings were made during two days in March 2001. Metro is distributed free. Svenska Dagbladet is a regular morning paper. Both newspapers are in the tabloid format. The newspapers were placed on a board for support, which allowed recording co-ordinate data using the SMI iView remote set. Subjects turned pages themselves. Data were recorded as 1024*768 coordinates in the reference system of the newspaper.

Subjects were instructed to read the newspaper for as long as they liked. They were asked to read the newspaper as they would normally read it. They could turn pages back and forth as they pleased. Each subject read for 15 minutes or more. After one hour, if they had not stopped, we stopped them. In post questionnaires, subjects stated that the presence of the eye tracker did not influence their reading.

Data Treatment

From the raw data of both the net and the newspaper recording, we calculated fixations using SMI’s algorithm with a minimum fixation time of 100 ms. These fixation data were then run through a custom-made reading filter.
The reading filter identifies fixations that appear before, between or after at least two successive forward-going saccades. Also, fixations that appear before and after return sweeps are identified. Correction saccades following a return sweep are not recognised, nor backward reading saccades.

The sizes of reading saccades and return sweeps were calibrated from samples of data in each recording. The filter therefore only identifies forward going saccades of a size characteristic to reading, excluding the longer scanning saccades in the same direction. Also, the text and columns in newspapers are much smaller than in net papers, yielding shorter reading saccades in the newspaper data. We used the criterion that 95% of all the reading saccades that we saw in the scan path plots of four selected pages per data set should have the start and stop fixation recognised by our filter.

The fixations that are not part of reading patterns are assumed to be part of scan- ning processes.

Results

In newspaper reading, 55.1% of all fixations by all subjects were recognised by our filter as being part of a reading pattern. For the net paper reading, 44.4% of all fixations are part of a reading pattern, significantly less (independent, two-tailed t-test, $p < 0.001$).

For newspapers, reading proportion strongly correlates with the time spent on a fold ($0.75$, $p < 0.0001$).

Intermediate Discussion

Hypothesis 1, inspired by Lewenstein et al. (2000), that there is less scanning and more reading activity in net paper reading than in newspaper reading, must be rejected.

In order to explain the difference, we need to study how scanning versus reading is distributed across the different layouts in net papers and newspapers. In our newspaper data, reading proportion varies across the different folds.

Figure 30.2 shows the average proportion of reading through all the folds of the Metro newspaper during the average 25 minutes of reading. The first ten folds with more editorial material have a reading percentage of about 60%, and then reading proportion drops sharply for folds 11–19, which are dominated by ads.

How does reading proportions vary in net papers? Specifically, does it increase on article pages and decrease on front pages? Exactly which layout elements in net papers are scanned and which are read?

Hypothesis 2: More reading on article areas than on other areas.

Hypothesis 3: Scanning mainly takes place over link lists.
Experiment 2

In the second experiment, we wanted to confirm that reading percentages increase during article reading and decrease between articles. If this were the case, it would yield a zig-zag pattern in the plot of reading percentage over pages, with a low reading value on the pages (presumably front pages) being scanned for entry points and high reading values on the articles.

Also, we wanted to test whether link lists in net papers have taken over the role of entry points in newspapers. If this were the case, people who scan a lot would spend more fixation time in the link lists, while people who read more would spend more time in the reading area.

Data Recording

In November 2001, we recorded 12 readers of the two net papers Dagens Nyheter (www.dn.se, with a Stockholm newspaper owner) and Sydsvenska Dagbladet Snällposten (www.sydsvenskan.se, with a Malmö newspaper owner. Recordings were made during one day in November 2001, using the SMI iView remote set at 50 Hz. Subjects read from a regular 120 MHz Pentium computer with a 100 Mbit Ethernet connection and a 17 inch (43 cm) screen. Data was recorded as screen co-ordinates in a custom-made browser. The browser compensates eye-tracking data for scrolling, saves screen dumps of the pages visited, and allows all normal browser activities. Also,
importantly, the browser signalises page changes (link clicks, back buttons etc.) to the eye-tracker data file. We also recorded data on video.

Each reader read both net papers. They were asked to read the net paper as they would normally read it. We imposed a reading time restriction of 5 minutes for each net paper.

Directly after the reading session, each subject was shown the video of his/her own reading session, showing the computer screen and the gaze marker (cf. Hansen, 1991). In this retrospective phase, subjects were asked to comment on what they read, how they found their way through the net paper structure etc. in a semi-structured interview.

**Data Treatment**

From the raw data, we calculated fixations using SMI’s algorithm with a minimum fixation time of 100 ms. These fixation data were then run through the same custom-made reading filter as in Experiment 1.

Retrospective comments were recorded and thematically sorted. Full results from the retrospective phase can be read in Barthelsson (2002).

All pages of both net papers were divided into areas of interest, following the design of the page (see Figure 30.3 and Table 30.1). The proportion of fixation time spent in each area and transitions between them were calculated.

---

**Reading or Scanning?**

---

**Figure 30.3:** Areas on the DN front page coded according to their content and function.
A zig-zag pattern emerges for all readers (example in Figure 30.4), showing an alternation in the amount of scanning and reading. This result supports the hypothesis that reading takes place on the pages with specific stories.

The reading rates for the two net papers were on average 44.8%, very close to the value in Experiment 1 for net papers and again significantly less than the value for newspapers.

The correlation between proportion of reading and time spent on a page is only 0.25.

Figure 30.5 exemplifies proportion of dwell time in areas, and transitions between them on the first page of the DN net paper. The middle reading area attracts the most fixation time, here 59%.

Of the different links on the front page, readers preferred the left link list. They looked significantly less at the right link lists (ANOVA, Tukey HSD, $p < 0.01$), even if both lists linked to thematically similar material.

The correlation between reading pattern frequency against dwell time in the reading area is 0.55, $p < 0.01$ over all readers and the entire reading session.

**Intermediate Discussion**

Hypothesis 2 is supported by the general zig-zag pattern for reading proportion found in the data. Reading takes place on article pages, and decreases when the reader returns to the front page.

Table 30.1: The different areas on the DN page coded according to their content and function.

<table>
<thead>
<tr>
<th>Area name</th>
<th>Content and function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back</td>
<td>Browser back button in the top navigation area</td>
</tr>
<tr>
<td>Top</td>
<td>Promotion box, name plate, advertisements and search box</td>
</tr>
<tr>
<td>Left</td>
<td>Headline links ordered by topics (national, international, economy, sports, culture) including the exact hour, linked to full articles</td>
</tr>
<tr>
<td>Middle</td>
<td>Lead editorials: Top stories with dominant visual elements such as headlines, photos and graphics, briefs and links to topic pages and to full articles</td>
</tr>
<tr>
<td>Right</td>
<td>Links to topic pages and different services available (travels, accommodation, stock market, etc.)</td>
</tr>
<tr>
<td>Ads</td>
<td>Advertisements</td>
</tr>
<tr>
<td>Scrollbar</td>
<td>Scrollbar</td>
</tr>
</tbody>
</table>
Figure 30.4: Proportion reading over pages in net papers.

Figure 30.5: Proportional dwell time and transitions between areas on the DN front page.
The zig-zag pattern is particularly clear with the DN net paper. DN has a design that allows for alternating between scanning for links on the front page (odd page numbers) and reading the text that appears when clicking the link (even page numbers). The more complex structure of the SDS net paper does not give such a clear alternation.

The correlation in Experiment 1 between reading activity and time spent on a fold reflects the fact that scanning a newspaper is made in search of entry points. When no interesting entry points are found, the reader does not continue to scan the fold but turns the page. For the folds with the lowest reading rates (below 15%), this happens after 3–5 seconds.

The non-existent correlation with net papers shows that if you do not find an interesting entry point in a net paper, you cannot turn the page. You have to keep scanning. Net paper readers choose their own path through the paper, and the majority of net paper pages are never seen.

Hypothesis 3 is supported by the detected correlation between reading pattern frequency against dwell time in the reading area. The correlation indicates that scanning in net papers mainly takes place outside the reading area; that is, in and across the link lists. This result can be interpreted as indicating that link lists in net papers taken over the role of entry points in newspapers.

The lengths of our recordings differ greatly between net and newspaper data. But the scanning and reading proportions largely remain constant through the recordings (see Figures 30.2 and 30.4). The variation we can see in the data reflects the content on the folds. For instance, reading proportion drops after fold 10 in the newspaper data, because the section starting there displays many full size ads.

General Discussion

Lewenstein et al. (2000) conclude that we read deeper into the texts we find in a net paper than we do in the newspaper. Since readers can be very selective in what they read in a net paper — by clicking only the links that interest them — it appears to mean that net papers offer more efficient reading than do newspapers. In this view, turning page after page in the order decided by the designers of the newspaper forces you to read, or at least scan, many texts that you would not have chosen to read in a net paper.

Still, as we have shown here, net paper readers scan more than newspaper readers. If efficiency in a media means that you read more and spend less time searching for what to read, then newspapers are more efficient.

In the semi-structured interviews after the reading session, a majority of our readers claimed that they scan more in net papers than in newspapers. They claim to scan more in order to find the two or three stories they will read in the net paper.

Concerning article topics, the net paper readers read stories thematically close to their own specific profession or interests. All net paper readers were reading the news category that could be termed “scandals and catastrophes”.

The newspaper readers in our study, like Hansen’s (1994) and Garcia and Stark’s (1991) readers, were much less selective. They read (parts) of text on all the different pages containing editorial material, including a wide variety of genres and topics.
Not only do net paper readers go deeper into articles than newspaper readers, as noted by Lewenstein et al. (2000). They also visit much fewer and more specialised texts. And in looking for these texts, they have to scan more than newspaper readers.

The linear structure of the newspaper invites linear browsing. News on each page therefore has at least the possibility of being seen. In net papers, most texts are never open for the reader to look at them. Instead of being opened in a browsing process, the net paper texts must compete for reader attention by means of links from the front page. Using a link to catch a reader to a story means feeding the reader with much poorer information on the story content than when the story presents itself spread out on the page of a newspaper.

The poorer chances of links catching reader interest may be the explanation why net paper readers have to scan more, why they read only certain types of texts and why they are bored so soon.

Experiment 2 showed a general zig-zag pattern in the reading proportion with higher scanning rates on the front page. Nine out of 12 subjects in Experiment 2 preferred to return to the front page after exploring the details of an article. In the retrospective phase we asked them about this behaviour, and our subjects stated that they wanted to return to the front page to ensure that they did not miss anything. When they probed deeper into an article they felt that they had left the main trail. In other words, only the front page is conceived of as a reliable provider of entry points.

The preference of the left link list detected in Experiment 2 could be a result of reader expectations of the design of net papers and web pages in general. But there are other studies of media perception that also show a preference for the left hand side of media layouts. For instance, Widman and Polansky (1990) show that of all small ads, those placed slightly to the left of the middle on a fold get the most fixations. Hansen (1994) showed that articles on the left page on a fold are seen significantly earlier than articles on the right page.

In a study of the use of Danish yellow pages, Hansen (1999) gave strong evidence for the dominance of the left hand side. Not only did Hansen’s subjects look more at the ads slightly to the left, they also phoned these companies more frequently, and they never once looked at or phoned a company advertising in the rightmost column.

Garcia and Stark (1991: 33) claim to have found the opposite, however. According to them, entry points on the right hand side of a fold are seen first. Newspaper designers that we have talked to often say they think the right hand side to be more looked at in their products.

Also, in perceptual psychology (see Arneheim, 1974; Gombrich, 1977), and within the so-called sociosemiotic approach (see Kress & van Leeuwen, 1990, 1996), the right hand side is regarded as more informative: “any pictorial object looks heavier at the right side of the picture” (Arneheim, 1974: 34). Kress and van Leeuwen assume that visual layouts form a grammar where the dimensions left-right, up-down and centre-margins are associated with different information values. The left-hand side is associated with known information, whereas new information is on the right.

Today, the links as a substitute for entry points are taken from net papers into newspaper design. In November 2001, the SDS newspaper presented a new design with a
list of links in the right column of the front page, motivating it by writing that they want to make it easier for readers to quickly find the story they are interested in.

Notes

1 Jupiter MMXI.
2 Intelligence.se.
3 Compare areas on the front page, Figure 30.3.
4 In netpaper reading, all subjects look at different pages in different orders. We therefore only show one typical subject reading the two netpapers SDS (6 pages) and DN (8 pages).

References


