User Preference Gathering through the Web

Y. Rho
School of Computer Science & Engineering
University of New South Wales, AUSTRALIA
yrho@cse.unsw.edu.au

T.D. Gedeon
School of Information Technology
Murdoch University, AUSTRALIA
t.gedeon@murdoch.edu.au

Abstract

Two online surveys were applied to identify user preferences as requirements for Web-based academic articles in terms of reading activities and format preferences. The first survey was email-based, where answers were totally dependent on participant's imagination with no given examples. Meanwhile, in the second Web-based survey, examples were embedded in the Web-based questionnaire. The examples helped the participants to recognise the importance of the first screen and understand the questions clearly. The data shows that the selection of reading from the first screen was 0% from the first survey but 18% from the second survey.

Reading activities with academic papers on the Web are not well investigated and various formats are being used for them. Our online survey results show that readers get overview information from reading from the screen, print the article and read it. The simple two-frame format was most preferred by 47% of the respondents, but the cascaded page format was regarded as the worst by 65%. An interesting result is that 26% selected the paper-like format as the worst style, which is being widely employed.

Keywords: Web-based preference gathering; online survey; Web-based articles; Reading activities and formats;

1. Introduction

Difference leads to difference. Different genres lead to different uses. Different media provide different facilities. Based on difference, the Web has been ever expanding its application areas and genres [8]. Web users are distributed all over the world and can be anyone on the Web. The wide geographical spread of Web users makes requirement gathering difficult. Remote online methods are being considered as a solution for the problem.

Requirement gathering

Rapid prototyping is a requirement gathering method to reduce the iterative development cycle by the development of different proposed concepts through prototypes and the evaluation of them (see [5]). Through this method, user preferences can be identified at a very early stage in the development process. This method is not based on users' imagination but experience. Hence, subjects evaluating the prototypes need to be real users.

Users understand the system through user interfaces so that prototyping methods are applied well for the collection of user interface requirements. Web documents are considered as one of the most popular user interface types. They usually can be accessed by anybody on the net even though not all of the readers are active. Active users for academic articles are people in research or academic environments but not strictly limited to them. Selecting subjects from those environments for surveys is reasonable.

Online surveys

An online survey is a survey method where data collection from participants is computerised, which is applied well as a feedback or evaluation mechanism for Websites ([11]). Participants answer questions from the screen and the answers are directly processed. This minimises the effort required for data processing and management.

Slaughter et al [9] reported that online surveys are as good as paper and pencil surveys and get more comment. User experience to the system is really important [12]. Our survey has demonstrated this quantitatively.

In the case of Web-based questionnaires to ask about Web-based documents, the benefit of being online can be maximised because realistic examples can be embedded at every necessary location in a questionnaire. It was also reported [12] that the checklist style questions were effective.

Properties of Web-based questionnaires

In order to identify usage patterns, called reading activities, with Web-based articles and the proper formats for those activities, we applied the two different online surveys: an email-based survey for simple questions about usage patterns and a Web-based survey with a Web-based questionnaire (see [7]). Harper, et al [4] described the advantages of Web-based questionnaires. These are:
- High portability;
- Easy to extend by anyone with a basic knowledge of HTML;
- Automated data collection and processing;
- A large number of users in a number of locations;
- Collection of information from real users in their natural habit;
- Access to a highly specialised user population that may be difficult to query personally; and
- Minimum delay between exposure to the interface and the measure of satisfaction.

However, no homogeneous experimental environment is ensured because respondents can work in different environments such as different browsers, different window sizes, and different interaction devices. Environmental effects can be critical for questionnaires involving examples. To minimise the effects, some sections to homogenise the environments need to be involved in a questionnaire, which are not likely to be involved in paper-based questionnaires.

Properties of academic articles on the Web

We have looked at some popular websites that contain Web-based academic articles. The sites include:
- ACM CHI97 (http://www.acm.org/sigchi /chi97/proceedings/paper/plp.htm);
- ACM/SIGCHI Bulletin (http://www.acm.org/ sigchi/bulletin/1997.4/ross.html);
- ACM Transactions on Computer-Human Interaction (http://www.acm.org/pubs/contents/ journals/tochi/1998-5/);
- Alertbox (http://www.useit.com/papers/ webwriting/writing.html);
- International Journal of Human-Computer Studies (http://ijhcs.open.ac.uk);
- WWW8 (http://www8.org/fullpaper.html).

The structural properties of the article formats in the sites can be described as follows:
- Abstracts are always in the articles, but tables of contents (in short, TOC) are not always present;
- Scrolling is a common method. If a TOC exists, indexing is always provided. The sliding method is used mainly as a secondary method;

- The single window layout is most popular;
- The ACM TOCHI seems to use the Web only as a delivery medium.

Long single page formats are being widely used for Web-based articles, which are not much different from the formats of their paper-based equivalent articles.

2. First survey: Email-based gathering

Purposes of the survey

This survey was to see whether researchers find research articles from the Web. If they do, what would be their usage patterns with Web-based articles?

Method

Email-based questionnaires of four questions were distributed to the research people numbering one hundred and thirty in our School. No example was given to participants.

Figure 1 shows the questions in the questionnaire.

```
Q1: Do you find academic articles from the Web? (Yes / No)
If 'Yes' for Q1,
Q2: What describes your behaviour best when you have an article
on the Web?
   (1) You just print it out, and then read the printed article.
   (2) You read the first few lines on the first screen, print out
       the article if you are interested in it, and read the
       printed article.
   (3) You read some concise parts such as titles & abstracts,
       print out the article if you are interested in it, and read
       the printed article.
   (4) You scan through the article, print it out if you are
       interested in it, and then read the printed article.
   (5) You read the article from the screen.
   (6) Others (please describe)
Q3: What could be your second choice in Q2?
If 'No' for Q1,
Q4: Why?
   (1) The Web articles are NOT credible.
   (2) The Web articles are frequently updated.
   (3) Others (please describe)
```

Figure 1: Questions in the first survey

Results

We received twenty-three replies, 18% of reply rate. Twenty-two (96%) replied that they find articles from the Web. Table 1 shows the selections on usage patterns. Majority by 64% of the replies for Q2 in

Figure 1 selected Activity 3, (3) in Q2, for their first choice and 32% selected Activity 4, (4) in Q2. The other selections were not considerable. An interesting result is that no one selected Activity 2

for the first choice. For the second choice, Activities 2 and 3 together recorded the highest selection frequency by six, 30%.

Table 1: Usage patterns with Web articles

| \	Details	1st (Q2)		2nd(Q3)	
Activity 1	Print and read	1	5%	1	5%
Activity 2	Read from the first screen, print and then read	0	0%	6	30%
Activity 3	Read concise parts, print and then read	14	64%	6	30%
Activity 4	Scan through, print and then read	7	32%	4	20%
Activity 5	Read from the screen	0	0%	3	15%
Others		0	0%	0	0%
N		22	100%	20	100%

Discussion

The most typical usage pattern with Web-based articles is Activity 3, (3) in Q2 in

Figure 1. However, the choice (4) should not be ignored because it has been selected for the first choice by 32%. This resultant pattern is very similar to the result on usage patterns of paper-based academic journals shown in [1]. It seems because readers read articles according to the fine metastructure (see [2][3]) they have, which leads to the reading activity of Figure 2. The respondents preferring Activity 3 seem to want to view the concise parts from the first screen. In the other hand, the respondents preferring Activity 4 seem to focus much more on the content overview of an article than its interface.

Survey conclusion

Readers find articles from the Web and get some ideas by reading some from the screen, print them out and read the printed articles, but seldom read them from the screen. Then, what formats do readers prefer and dislike? The second survey answers this question.
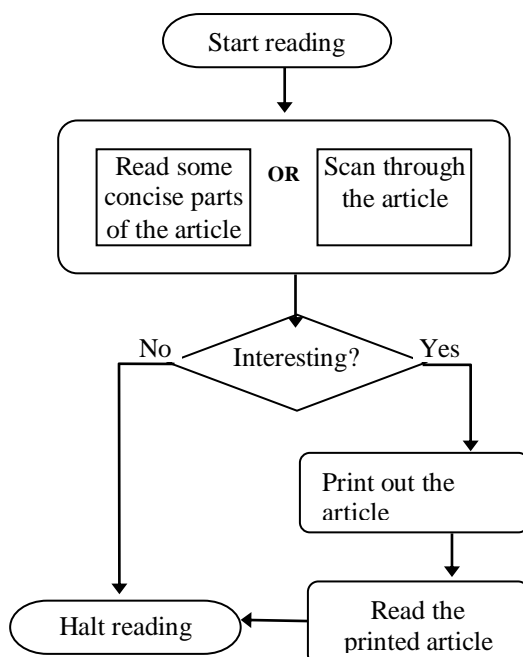
Figure 2: A typical reading activity pattern

3. Second survey: Web-based gathering

From the previous survey, we know typical reading patterns with Web-based academic articles. Based on that knowledge, we have tried to identify readers' preferences in the format structure of Web-based articles. In the second survey, real examples are provided to the respondents as embedded in the Web-based questionnaire, which consists of checklist style questions.

Figure 3 is a partial screen shot.

Purposes of the survey

The second survey was to identify the format structure which readers prefer and the effect of emamples. The survey also was to see the effects of examples to participants' reading activity selections.

Methods

We prepared a Web-based questionnaire of three parts of five question groups of eighteen questions in total (see [11]). The first part of the questionnaire is for environment set-up to control the visual properties and the window size of the browser in terms of the information amount, which is to avoid different visual volume effects to participants. The second part consists of the questions about three different layers of overview types, window layouts and manipulation methods. The last part is about usage patterns and overall preferences in format structures. Each question has at least one corresponding example link. The survey was announced via email with the URL for a questionnaire to researchers in information technology numbering one hundred and fifty. They were research students, research staff and academic staff. Neither undergraduate nor coursework student were included because they seldom use the Web to find academic articles.
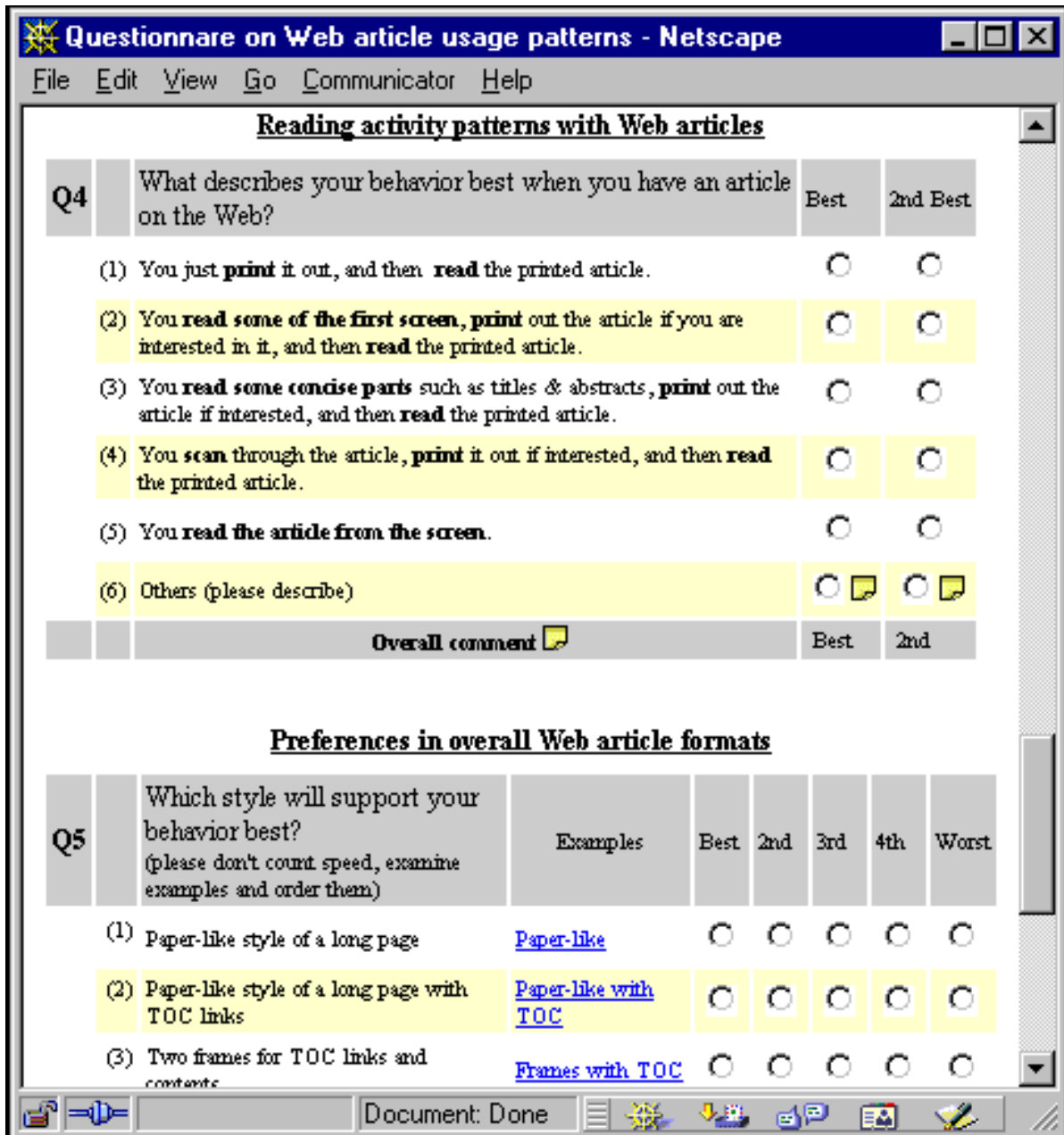
Figure 3: The Web-based questionnaire [11]

Results

We received thirty-four replies, 23% of reply rate. Most of them used 17" monitors to fill out the questionnaire and like the previous email-based survey result, Activity 3 is the most typical reading activity that 55% selected, but Activity 2 and 4 were selected by 18% each (Table 2). In the previous survey, the ratio was 0% for Activity 2 and 32% for Activity 4, which are very different from the Web-based survey result. This will be discussed in the discussion section.

Table 2: Reading activities

\	Details	1st choice		2nd choice	
Activity 1	Print and read	2	6%	2	6%
Activity 2	Read from the first screen, print and then read	6	18%	10	29%

Activity 3	Read concise parts, print and then read	19	56%	8	24%
Activity 4	Scan through, print and then read	6	18%	10	29%
Activity 5	Read from the screen	1	3%	4	12%
Others		0	0%	0	0%
N		34	100%	34	100%

Five example formats are embedded in the questionnaire (try Q5 in [11]). 47% of the respondents selected the two-frame format for the best and 35% selected the paper-like-with-TOC format (Table 3). Meanwhile, 65% selected the cascades as the worst. It is interesting that 26% of our participants selected the paper-like format as the worst. No one indicated either the paper-like-with-TOC or the two-frame format for the worst.

Table 3: Preferences in overall formats

	\	Best		2nd		3rd		4th		Worst	
1	Paper-like	2	6%	7	21	9	26	6	18	9	26
2	Paper-like with TOC	12	35%	13	38	5	15	4	12	0	0
3	Two frames	16	47%	5	15	7	21	5	15	0	0
4	Slides	2	6%	8	24	10	29	12	35	1	3
5	Cascades	2	6%	1	3	2	6	5	15	22	65
	N/A	0	0%	0	0	1	3	2	6	2	6
	N	34		34		34		34		34	

Table 4: Activity-format correlation

\	Format 1			Format 2			Format 3			Format 4			Format 5			N(activity)	
Activity 1	0			2	100%	17%	0	0%	0%	0			0			2	6%
Activity 2	1	17%	50%	1	17%	8%	3	50%	19%	0			1	17%	50%	6	18%
Activity 3	1	5%	50%	6	32%	50%	9	47%	56%	2	11%	100%	1	5%	50%	19	56%
Activity 4	0			2	33%	17%	4	67%	25%	0			0			6	18%
Activity 5	0			1	100%	8%	0	0%	0%	0			0			1	3%
N(format)	2	6%			35%		16	47%		2	6%		2	6%		34	100%

There seems a correlation between the activities and the formats (Table 4). The underscored percentages of the shaded row show the format distribution over Activity 3. The Format 3's contribution to Activity 3 is highest by 47% and the Format 2 is next by 32%. In the other view, the shaded column shown corresponds to Activity distribution over the Format 3. Activity 3 shares 56%, which is followed by Activity 4 and 2. Activity 3 matches the Format 3 and vice versa. In addition, Activity 2 and 4 match the Format 3 as well, by 50% and 67%, respectively.
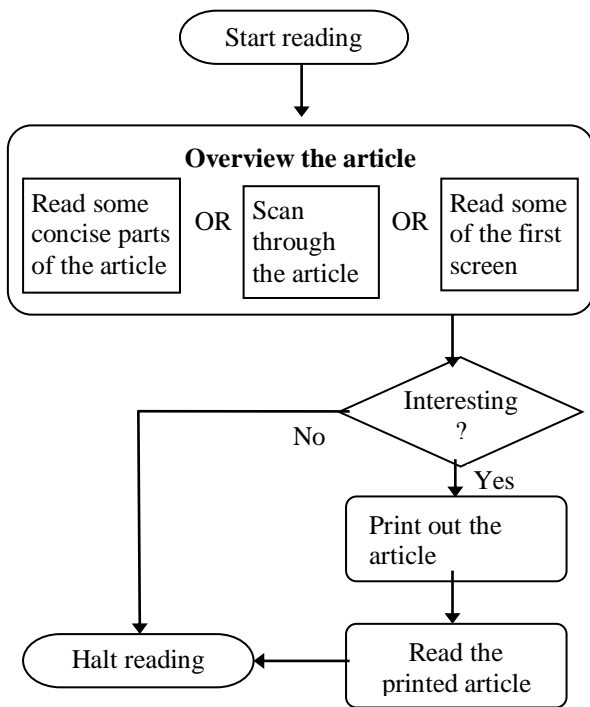
Figure 4: A typical reading activity pattern

Discussion

The most typical reading activity with a Web-based article is that the reader reads some concise parts of an article, prints out the article if they are interested in it and then reads the printed article. The second survey also produced a similar result.

A big difference between two surveys on the reading patterns happened at Activity 2 and 4. In the first survey, Activity 4 shared 32% but Activity 2 had 0%. Then, in the second survey, Activity 4 decreased by 14% to 18% but Activity 2 increased by 18%. The portions are even. Why? There was no significant difference in the distribution list and the question for this topic. The difference seems to have been from the existence of examples that the participants could experience. The first questionnaire was based on email with no examples to try. Meanwhile, the second questionnaire was based on the Web with many examples. The participants had many chances to try different formats of Web-based articles to answer the preceding questions. They are likely to be aware of the importance of the first screen. There were some additional comments on that point from participants.

The two-frame format was most preferred by the respondents, which is against the general idea. The majority (47%) selected the two-frame format for the purpose and the paper-like-with-TOC format for the next (35%). The cascade format was worst because of its complexity on the screen (65%), but no one selected either the two-frame format or the paper-like-with-TOC format for the worst. Activities and formats showed the close correlation; Activity 3, 2 and 4 most match the two-frame format (Table 4).

Survey conclusion

The early part of the reading process, which is to overview an article, happens on the screen when reading a Web-based academic article from the screen. The survey results show that its early parts,

overviewing and printing, are most likely to be supported by the two-frame format that insures the quality of the paper-based article format when printed. The examples affected users' selections.

4. Conclusion

Features revealed

- Embedded examples were necessary to provide users with chances of experience and helpful to make questions clear as well.
- To avoid environmental effects on example presentation, some instructions for the configuration set-up were needed in the questionnaire.
- Some participants commented that they didn't like scrolling the questionnaire.
- Participants are not likely to go backward to review what they answered.
- Data processing did not take a big effort as expected. However,
- Participants could make multiple replies so that filtering was required.
- How frequently the respondents examined the given examples was not quite sure.

Survey results

- Readers get overview information from the screen, print the article and read it (Figure 4).
- The two-frame format was most preferred.
- Providing two versions together or either one seems reasonable.
- An interesting discovery is that one quarter of the respondents did not like the paper-like format that is being widely used for academic articles on the Web.

Further research

This research focused on only user preferences in Web-based academic articles. Different genres need to be tested. Gathering methods based on real reading tasks instead of examples is required.

References

[1] Dillon, A., "New Technology and the Reading Process," Computers in Libraries, July 1991, 23-26.

[2] Dillon, A., "Readers' models of text structures: The case of academic articles," International Journal of Man-Machine Studies, 35, 1991, 913-925.

[3] Dillon, A., " Expertise and the Perception of Shape in Information," Journal of the American Society for Information Science, 47(10), 1996, 786-788.

[4] Harper, B., Slaughter, L., Norman, K., "Questionnaire administration via the WWW: A validation & reliability study for a user satisfaction questionnaire," Proceedings of Webnet'97, 1997. http://www.lap.umd.edu/QUIS/webnet.pdf

[5] Kirakwsky, J., "Requirement Engineering and Specification in Telematics," Telematics Engineering Project TE 2010, D3.2, 1997. http://www.ucc.ie/hfrg/projects/respect/urmethods/index.html

[6] Kristensen, A., "Formsheets and the XML Forms Language," WWW8 Proceedings, 1999. http://www8.org/w8-papers/1c-xml/formsheets/formsheets.html

[7] Perlman G., "Web-Based User Interface Evaluation with Questionnaires," 1997. http://www.acm.org/~perlman/question.html

[8] Shneiderman, B., "Designing Information-Abundant Websites: Issues and Recommendations," International Journal of Human-Computer Studies, 1997. http://www.cs.umd.edu/projects/hcil/members/bshneiderman/ijhcs/main.html

[9] Slaughter, L., Harper, B., and Norman, K., "Assessing The Equivalence Of The Paper And On-line Formats Of The Quis 5.5". Proceedings of the 2nd Annual Mid-Atlantic Human Factors Conference, Washington, D.C, 1994, 87-91. http://lap.umd.edu/LAPFolder/Papers/SHN.html

[10] Rho, Y., Gedeon, TD, "A Link-Click Lifecycle on the screen," APWeb'98 Proceedings, Beijing, 1998. http://www.cse.unsw.edu.au/~yrho/Publications/apweb98/apweb98_C-tiled_main.htm

[11] Rho, Y., Questionnaire on Reader's preferences in Web-based Academic Articles, 1999. http://www.cse.unsw.edu.au/~yrho/WebPaper/2ndQuestion.htm

[12] Root, R. W., Draper, S., "Questionnaires as a software evaluation tool." CHI'83 Proceedings, 1983, 83-87.

[13] WAMMI Web Usability Questionnaire, 1998. http://www.nomos.se/wammi/hfrg.html