# The Characteristics of Usability and Users' Eye movements in Searching for Information in a Hierarchically Organized Information Structure

**Yoshiko Habuchi (habuchi@tokuyama-u.ac.jp)**
Department of Welfare Information, Tokuyama University,
843-4-2 Gakuendai Shunan Yamaguchi, 745-8566 JAPAN

**Haruhiko Takeuchi (takeuchi.h@aist.go.jp)**
National Institute of Advanced Industrial Science and Technology (AIST),
1-1-1 Higashi Tsukuba Ibaraki, 305-8566 JAPAN

**Abstract**

Finding information by successively following hyperlinks on web pages is a typical task performed on the web. A number of web users search for specific information and several studies have concluded that following the "information scent" is the fundamental process involved in the behavior. The aim of this study was to investigate the relationship between the strength of the "information scent" and user behavior by applying a usability inspection method developed for web usability studies. Three typical usability problems of navigation, (a) a page with a weak scent correct link problem, (b) a page with an unfamiliar correct link problem, and (c) a page with a competing links nested under competing headings problem, were compared with a problem-free page. In this study, we applied the Cognitive Walkthrough for the Web method to simulate a website, and analyze user behavior along with usability problems. Participants were asked to find an article on a simulated encyclopedia website. The success rate, total clicks, total time, fixation count and gaze time were analyzed. The results showed that the critical issues caused by usability problems appear on the target-link page. The results of this study confirm the effect of "information scent" and provide a scientific insight into web navigation.

**Keywords:** information scent; LSA: latent semantic analysis; hierarchical information structure; web usability; eye tracking

## Introduction

Over the last decade, eye-tracking studies have provided detailed insights into the way users interact with websites. In earlier usability studies in web navigation, the level of difficulty of a task for a user was measured by finding the number of clicks required or the time taken to complete the task, where a high difficulty level pointed to usability problems in the web page. Eye-tracking methods have made it possible to analyze user behavior in detail.

Several eye-tracking studies have provided an overall understanding of the behavior of website users based on their eye movements. Nielsen (2006) demonstrated that the eye-movement patterns of website visitors are F-shaped by aggregating a large amount of eye-movement data from web pages. Cutrell & Guan (2007) and Guan & Cutrell (2007) presented a series of detailed studies examining the relationship between the fixation duration, and the ranking and presentation of search results i.e. whether the search results were accompanied by short, medium, or long descriptions. They discovered that providing more information next to the search results significantly improved performance in information tasks, but degraded performance in navigation tasks, where the performance was assessed in terms of selecting the correct search result and the time taken to complete the task. They argued that the difference in performance was because when users were given longer descriptions, they paid more attention to the description and less to the URL of the search result, which would be of help in assessing the usefulness of the search result. These eye-tracking studies have provided a broad understanding of the behavior of web users.

Habuchi, Kitajima, & Takeuchi (2008) focused on users' cognitive activities, especially their goals and knowledge when searching for information on the web. They explored the relationship between usability problems and web users' eye-movements by independently assessing usability problems using Cognitive Walkthrough for the Web (CWW) [1] (Blackmon, Polson, Kitajima, & Lewis, 2002; Blackmon, Kitajima, & Polson, 2003, 2005; Blackmon, Mandalia, Polson, & Kitajima, 2007). They considered three typical usability problems: the weak scent correct links nested under competing headings problem, the unfamiliar correct link problem, and the competing links nested under competing headings problem. They found that if a webpage had any usability problems, user performance deteriorated noticeably in terms of the total number of fixations in the subsequent link selection stage. The performance was worst on pages with a weak scent problem or an unfamiliarity problem, which resulted in longer fixation durations because users examined the link carefully before selecting it. However, no difference in performance was observed in the initial heading selection stage between pages with usability problems and those without. The study showed that it is

[1] Cognitive Walkthrough for the Web (CWW) is a web usability inspection method that can detect several usability problems that a website visitor may encounter while navigating through the website in search of specific information by successively selecting hyperlinks on intermediate navigation pages. It uses a Latent Semantic Analysis (LSA) semantic space (Landauer & Dumais, 1997) to measure the "information scent" i.e. local proximal cues. Web users use the "information scent" to make navigation choices while navigating the web by following hyperlinks (Pirolli & Card, 1999; Blackmon et al., 2002, 2003, 2005; Chi & Suppattanasiri, 2003; Pirolli, 2005).

possible to distinguish between such pages based on the task completion time and total number of fixations. Furthermore, it showed that users' gaze patterns varied depending on the whether the problem was in the link itself, or elsewhere. However, the test material they used consisted of just two samples for each usability problem. To verify their results and draw a general conclusion, results obtained using a larger number of samples must be examined.

The aim of this study is to investigate the relationship between usability problems and information search behavior by applying the CWW method to handle various types of usability problems. In this study, we conducted more experiments and controlled other factors such as familiarity of the desired information. We built a controlled website to study the relationship between usability problems and user behavior.

## Experiment

The task used in this study is an information search task performed within a hierarchically organized information structure. The task difficulty depends primarily on the strength of information scent on each navigation page for a given search target. The three types of usability problems compared are, (1) weak scent correct link problem, (2) unfamiliar correct link problem, and (3) competing links nested under competing headings problem; where problem-free items existed. These usability problems can be detected through CWW.

### Method

**Materials** The Japanese LSA semantic space is necessary for finding usability problems in CWW. It was constructed using corpora from the Japanese language Wikipedia abstract containing 116,038 words and 129,937 contexts. The semantic space of the Japanese LSA consists of 116,038 words and 300 dimensions. Please refer to Takeuchi, & Habuchi (2008) for details.

A simulated encyclopedia website was constructed for this study. It has the same link structure as the Microsoft Encarta website. The menu was two levels deep, with 9 top-level links (called headings) and 93 second level links (called links). Each entry word and its description from the Wikipedia abstract were categorized into one of the 93 links. Usability problems were identified by applying CWW on these articles (Takeuchi, & Habuchi, 2008).

We selected 97 candidate articles, such that each article either had no usability problems, or had one of the following problems: weak scent correct link problem; unfamiliar correct link problem; and competing links nested under competing headings problem (three competing links). The predicted link-click count in this experiment was 2.29 for no usability problems, 3.81 for weak scent, 4.05 for unfamiliar links, and 4.26 for competing links nested under competing headings, according to the CWW model (Blackmon et al. 2007).

The entry words from the selected articles varied in their degree of familiarity for users. As it is desirable for each entry word from the candidate material to have the same level of familiarity, we conducted the word concept familiarity rating task for the 97 entry words with 48 additional subjects (aged between 20 and 33 years with a mean age of 24.3 years). The familiarity rating scale was from 1 (least familiar) to 5 (most familiar). Finally, we selected 20 articles such that each problem type had 5 articles from which unfamiliar entry words would be chosen (see Table 1). These articles were used as the search targets in the experiment.

**Participants** Thirty-five university students participated in this study. Nine of the participants were excluded: six due to not having sufficient eye-tracking data and three others who reported after the experiment that they misunderstood the task. Twenty-six university students (11 males and 15 females; aged between 18 and 24 years with a mean age of 20.6 years) participated in the experiment. All participants were native Japanese speakers and received compensation for their participation. They had normal or corrected to normal vision. They were all regular users of the internet and were used to browsing with Internet Explorer. They were ignorant of the hypothesis being investigated in the study.

**Apparatus and Procedure** The eye-tracking equipment used was a Tobii X60 eye-tracker with Tobii studio 1.24 software. The minimum fixation duration was set to 40 ms and the fixation radius to 20 pixels. Eye movements were recorded at a sampling rate of 60Hz. The experimental website was opened in Internet Explorer 6 under Windows XP. URL visibility events, mouse movement, and click events were recorded.

The task was to find the relevant article within 130 seconds. The search topic, along with a short description was provided to the participants. The expectation was that participants would build a mental representation of the item, which they would use to evaluate the "information scent" of the navigation pages in the website.

Participants were tested individually. They were first seated in front of a screen-mounted eye-tracking system. Next, they were required to search for the desired information on the simulated encyclopedia website. In each test, the participant was given a target word. On clicking, the target word was replaced on the screen by the heading page containing a description of the search target and 9 headings. Participants made their choice by clicking on one of the sections. If a participant reached the target-link, the target page was shown, and then the next trial began when the participant clicked the "next" button (see Figures 1 and 2). To eliminate order effects, the 20 items were allocated among four blocks and the order of the blocks was counterbalanced. Each participant was given four blocks of five trials each. There was a short break of a few minutes between each block. Participants were tested each session lasted approximately 25 minutes.

The dependent variables were the success rate, time taken to complete the task, total number of clicks, total fixation count and gaze time.

Table 1: Properties of experimental materials.

| Property | No. of items | Concept familiarity[1] | Target and Target-link Cosine value | Target-link Vector length |
|---|---|---|---|---|
| Problem-free | 5 | 1.14 (.41) | .52 (.31) | .16 (.30) |
| Weak Scent | 5 | 1.20 (.48) | .09 (.00) | .22 (.22) |
| Unfamiliar | 5 | 1.12 (.42) | .58 (.23) | .08 (.13) |
| Competing links (3) | 5 | 1.18 (.47) | .45 (.27) | .42 (.00) |

[1] Concept familiarity was rated on scale from 1 (least familiar) to 5 (most familiar).

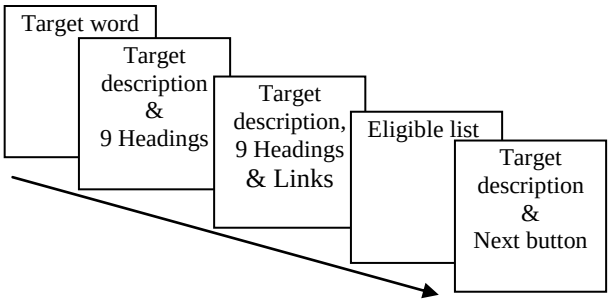


Figure 1: Sequence of events


Figure 2: Menu on the target-link page. The top part of the screen shows the target description and the area under it shows the menu. The links that appear under each heading became visible on clicking on the appropriate heading, and the links were only shown for one heading at a time.

## Results

We analyzed the data from when a heading page was opened to when the target link was clicked. An analysis of variance (ANOVA) and several comparisons were performed separately for the success rate, total click count, and the task completion time. These results are shown in Tables 2 and 3, grouped by usability problem type. An ANOVA and several comparisons were also performed separately for the success rate, total clicks, and total time. The significance level for ANOVA results was $p<.05$. A summary of the success rate, total clicks, and total time results for the experiment is presented in Table 2.

**Success rate** Before an angular transformation was performed, the success rate data were skewed: Problem-free = -2.56, Weak scent problem -1.19, Unfamiliarity problem -.53, Competing links problem 1.36. Post-transformation, the skewness was reduced to -2.56, -1.11, -.03, and 1.36, respectively. An analysis of the variance indicated a significant difference between the angular-transformed success rates [$F(3, 75)=30.20$ $p<.001$]. Post hoc testing indicated that the transformed success rate for the "problem-free" case were significantly higher than for cases with a usability problem.

**Click count** Prior to the log transformation, the total clicks data were skewed: Problem-free 1.03, Weak scent problem 0.97, Unfamiliarity problem 0.24, Competing links problems 0.45. Post-transformation, the skewness was reduced to 0.50, 0.36, -0.31, and -0.38, respectively. An analysis of the variance indicated a significant difference among the log-transformed total clicks for usability problems [$F(3, 75)=51.20$, $p<.001$]. Post hoc testing indicated that the transformed click count for the "problem-free" case were significantly less than for cases with a usability problem.

**Solution time** Prior to the log transformation, the solution time data were skewed: Problem-free 1.18, Weak scent problem 0.25, Unfamiliarity problem 0.12, Competing links problems 0.18. Post-transformation, the skewness was reduced to 0.51, -0.07, -0.64, and -0.38, respectively. An analysis of the variance indicated a significant difference in the log-transformed success time for usability problems [$F(2.58, 64.38)=59.63$, $p<.001$]. Post hoc testing indicated that transformed solution times for the "problem-free" case were significantly less than for cases with a usability problem.

### Eye tracking analysis

We conducted eye-tracking analysis for two regions of the web page, the target description area, and the navigation area during each target selection stage. Tables 4 and 5 show the results of the eye tracking analysis.

**Fixation counts** Before analysis, all data were logarithmically transformed.

Table 2: Success rate, total click count, and solution time.

| | Success rate | Total Clicks | Solution time |
|---|---|---|---|
| Problem-free | 97.7% | 3.92 (1.7) | 17.1 ( 9.7) |
| Weak Scent correct link | 72.3% | 9.45 (2.6) | 53.0 ( 9.7) |
| Unfamiliar correct link | 87.7% | 6.93 (2.2) | 36.6 (15.8) |
| Competing link nested under the competing heading (3) | 84.6% | 7.15 (1.6) | 35.4 ( 7.2) |

Note: Time measurements in seconds. Number shown in parentheses is SD.

**Target-heading selection stage** A two-way analysis of variance was conducted for the fixation counts until the initial heading was clicked. An information area main effect was not obtained [F(1, 25)= 2.23, n.s.]. Also, the usability problem main effect was significant [F(2.77, 69.34)=25.02, p<.001], confirming that the Weak scent problem link pages had significantly more fixations than other pages. Finally, the information area and usability problems interaction was significant [F(2.57, 64.35)= 20.19, p<.001]. A Bonferroni post hoc analysis of this interaction revealed that the Weak scent and Unfamiliar link pages had more fixations than that the other two cases, regardless of page area. In addition, although the Competing links problem pages had fewer fixations in the goal description area, they had more fixations in the navigation area.

**Transit-link stage** Not all participants saw the extra transit-link page; analysis reduced the participant number from 26 to 19. A two-way analysis of variance was conducted for the total fixation count in the transit-link selection page. An information area main effect [F(1, 18)= 111.64, p<.001] was significant, confirming that the navigation area was responsible for far more fixations than the target description area. The usability problem main effect was significant [F(2.48,44.63)=13.02, p<.001], which confirmed that the Weak scent or Competing problem link pages had significantly more fixations than the Problem-free pages. The interaction of information area and usability problems was significant [F(2.65,47.67)=4.81, p<.01]. A Bonferroni post hoc analysis was performed and it showed that in total, the Weak scent, and Competing links problem pages had more fixations than the Problem-free pages.

**Target-link selection stage** A two-way analysis of variance was conducted for the fixation counts in the target-link selection page. An information area main effect [F(1, 25)= 156.28, p<.001] was significant, confirming that the navigation area was subject to significantly more fixations than the goal description area. The usability problem main effect was significant [F(2.98,74.57)=15.87, p<.001], confirming that the Weak scent and Unfamiliarity problem link pages had significantly more fixations than the Problem-free pages. However, Competing links problem pages had fewer fixations than Problem-free pages. The interaction of information area and usability problems was also significant [F(3,75)=5.16, p<.01]. A Bonferroni post hoc analysis confirmed the difference in fixation distribution. Weak scent link problem pages in total had more fixations than Problem-free pages. In the target-description area, Unfamiliar link pages had more fixations than Problem-free pages. However, Weak scent link pages gathered more fixations in the navigation area.

**Gaze time** Before analysis, all data were logarithmically transformed.

**Target heading selection stage** A two-way analysis of variance was conducted for the gaze time until the target-link was clicked. An information area main effect was obtained [F(1, 25)=4.83, p<.05], confirming that the target-description area had significantly longer gaze time than the navigation area. The usability problem main effect was significant [F(2.94, 73.59)=24.18 p<.001], confirming that the Weak scent problem link pages had significantly longer gaze time than the other cases. The information area and usability problems interaction was significant [F(2.55, 63.72)=18.05, p<.001]. A Bonferroni post hoc analysis of this interaction revealed that the gaze time at the target-description area for Weak scent correct link problem items was longer than that in other cases. However, in the navigation area, Weak scent, and Unfamiliar link problem pages had longer gaze time than the Problem-free pages.

**Transit-link stage** We summed up surveyed areas individually and divided total gaze time by total pages opened. As a result, the number of participants was reduced from 26 to 19. A two-way analysis of variance was conducted for the gaze time for transit-link pages. An information area main effect was obtained [F(1, 18)=83.12, p<.001]. The usability problem main effect was significant [F(1.62,29.24)=10.96, p<.001], confirming that Weak scent and Competing links problem pages had significantly longer gaze time than the Problem-free pages. The information area and usability problems interaction was obtained [F(1.44, 25.94)=7.86, p<.01]. A Bonferroni post hoc analysis of this interaction revealed that in the target-description area, all usability problem pages had a longer gaze time than Problem-free pages. However, in the navigation area, Weak scent, and Competing links problem pages had longer gaze times than the Problem-free case.

**Target-link selection stage** A two-way analysis of variance was conducted for the gaze time of the target-link selection page. An information area main effect [F(1, 25)= 7.55, p<.001] was significant, confirming that the navigation area was subject to significantly longer gaze time than the target description area. The usability problem main effect was significant [F(2.1, 52.53)=7.55, p<.001], confirming that Competing links problem pages had a lower gaze time than other cases. Unfamiliar link problem pages had a longer

gaze time than the Problem-free case. The interaction between the information area and usability problems was not significant [F(1.6, 39.88)=2.32, n.s.].

## Discussion

The results indicated clear differences between problem-free and usability problem pages, in the success rate, total clicks and total time. The eye-tracking data showed that users' eye movements for selecting a target link varied depending on the nature of the usability problem. Two main conclusions can be drawn from this study. The first is that usability problems occur from the first heading selection stage in the navigation area. Pages with a weak scent problem had more issues than other pages. Web pages with the latter problems forced users to examine links and the target description more times than for problem-free pages.

The other point was certain characteristics of usability problems were clear at the target-link selection stage. In pages with a weak scent problem or an unfamiliarity problem, it took users longer to select a link. This is presumably because participants had to discern the meaning of the correct link by carefully reading the link label. A possible explanation for this phenomenon is that for the weak scent correct link problem, participants looked at the navigation area several times. In contrast, for the unfamiliar correct link problem, participants remained at one link page for longer. These situations differed from cases of competing links nested under competing headings, where users could immediately select a link because it was a good match for the target and less effort was needed to understand the meaning of the link. We can also say users' eye movement before clicking a link at the target-link selection stage varied according to the usability problem of the targeted link.

Table 3: Mean percentage of shortest path achievement.

| Processing Stage | From the start until Target-heading | From Target-heading until Target-link | From the start until Target-link |
|---|---|---|---|
| Problem-free | 83.8% (13.9) | 72.3% (18.8) | 63.8% (16.0) |
| Weak Scent correct link | 54.6% (18.4) | 34.6% (16.5) | 23.8% (15.0) |
| Unfamiliar correct link | 55.4% (19.8) | 51.5% (24.1) | 23.8% (15.0) |
| Competing links nested under the competing heading (3) | 53.8% (17.7) | 57.7% (16.3) | 36.2% (18.8) |

Note: Number shown in parentheses is SD.

Table 4: Mean Fixation counts.

| Processing Stage | Target-Heading selection stage (n=26) | | Transit Link page (n=19) | | Target-Link selection stage (n=26) | |
|---|---|---|---|---|---|---|
| Area of web page | Target description | Navigation | Target description | Navigation | Target description | Navigation |
| Problem-free | 15.2 (8.1) | 18.9 (12.2) | 7.6 (10.7) | 53.5 (30.0) | 8.1 (6.9) | 31.6 (15.8) |
| Weak Scent correct link | 31.7 (14.3) | 26.8 (14.5) | 23.3 (17.7) | 100.5 (48.7) | 11.1 (7.3) | 55.4 (24. 0) |
| Unfamiliar correct link | 16.0 (8.8) | 21.6 (10.6) | 11.4 (10.7) | 65.5 (37.0) | 15.9 (13.1) | 42.6 (25.2) |
| Competing links nested competing heading (3) | 12.2 (7.6) | 23.2 (12.8) | 16.0 (12.9) | 73.9 (50.2) | 5.4 (5.5) | 25.7 (13.4) |

Note: Number shown in parentheses is SD.

Table 5: Mean gaze time.

| Processing Stage | Target-Heading selection stage (n=26) | | Transit Link page (n=19) | | Target-Link selection stage (n=26) | |
|---|---|---|---|---|---|---|
| Area of web page | Target description | Navigation | Target description | Navigation | Target description | Navigation |
| Problem-free | 2612 (1650) | 3433 (1667) | 1306 (1828) | 10196 (7101) | 1422 (1203) | 6112 (2673) |
| Weak Scent correct link | 5556 (2823) | 4738 (2205) | 4306 (3613) | 16806 (9807) | 1853 (1269) | 9346 (3307) |
| Unfamiliar correct link | 2956 (1606) | 4089 (1798) | 1830 (1882) | 11953 (5767) | 2604 (2153) | 8731 (3754) |
| Competing links nested competing heading (3) | 2173 (1446) | 3960 (1741) | 2668 (2458) | 12057 (6333) | 1017 (1178) | 4521 (1878) |

Note: All measurements are on the millisecond time scale. Number shown in parentheses is SD.

## Conclusion

This study aimed to investigate how website visitors search for desired information in web pages with usability problems. Three types of usability problems, the weak scent correct link problem, the unfamiliar correct link problem, and the competing links nested under competing headings problem, were compared against problem-free pages. The results showed that participants' behavior when faced with usability problems was distinguishable from behavior on problem-free pages in terms of the success rate, total number of clicks and total time taken. In addition, it was possible to distinguish the usability problem facing the user from the fixation counts and gaze times at the target-link selection stage.

The eye-tracking analysis expanded on the information gained in previous studies. This study showed that users' eye movement before clicking a link in the target heading and target-link selection stages varies depending on the usability problem. Pages with competing links nested under competing headings have users spending less time before selecting the correct link, but they involve more clicks due to confusion from incorrect links. Pages with a weak scent problem or an unfamiliarity problem have users taking more time to select the correct link.

This study has attempted to explain the relationship between "information scent" and user behavior. In future work, we aim to extract the discriminative behavior when two or more problems exist simultaneously.

## Acknowledgments

This study was conducted while the first author was in the National Institute of Advanced Industrial Science and Technology (AIST). We are grateful to Muneo Kitajima, Professor, Nagaoka University of Technology, for his support and guidance.

## References

Blackmon, M. H., Polson, P. G., Kitajima, M., & Lewis, C. (2002). Cognitive walkthrough for the web. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 463-470). New York, NY: ACM.

Blackmon, M. H., Kitajima, M., & Polson, P. G. (2003). Repairing usability problems identified by the cognitive walkthrough for the web. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 497-504). New York, NY: ACM.

Blackmon, M. H., Kitajima, M., & Polson, P. G. (2005). Tool for accurately predicting website navigation problems, nonproblems, problem severity, and effectiveness of repairs. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 31-40). New York, NY: ACM.

Blackmon, M. H., Mandalia, D. R., Polson, P. G., & Kitajima, M. (2007). Automating usability evaluation: Cognitive walkthrough for the web puts LSA to work on real-world HCI design problems., In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Chi, E., Rosien, A., & Suppattanasiri, G. (2003). The bloodhound project: Automating the discovery of web usability issues using the infoscent simulator. *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 505-512). Fort Lauderdale, NY: ACM.

Cutrell, E., & Guan, Z. (2007). What are you looking for?: an eye-tracking study of information usage in web search. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 407-416). New York, NY: ACM.

Guan, Z., & Cutrell, E. (2007). An eye tracking study of the effect of target rank on web search. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 417-420). New York, NY: ACM.

Habuchi, Y., Kitajima, M., & Takeuchi, H. (2008). Comparison of eye movements in searching for easy-to-find and hard-to-find information in a hierarchically organized information structure. *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 131-134). New York, NY: ACM.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Nielsen, J. (2006, April 17). F-shaped pattern for reading web content. Jakob Nielsen's Alertbox, Retrieved from http://www.useit.com/alertbox/reading_pattern.html.

Pirolli, P. (2005). Rational analysis of information foraging on the web. *Cognitive Science*, 29, 343-373.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106, 643-675.

Takeuchi, H., & Habuchi, Y. (2008). An automatic evaluation of category names based on Latent Semantic Analysis, *Proceedings of Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on advanced Intelligent Systems* (pp. 423-428). Nagoya, Japan: SCIS & ISIS.

Wikipedia: The Free Encyclopedia. (n.d.). Retrieved from http://ja.wikipedia.org/wiki/