Thinking style impacts on Web search strategies

Gloria Yi-Ming Kao *, Pei-Lan Lei, Chuen-Tsai Sun

Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Rd., Hsinchu, Taiwan 300, Taiwan

Available online 14 September 2007

Abstract

Web searches entail complex cognitive processes influenced by individual differences, and users with similar cognitive or skill factors tend to develop multiple search strategies. The authors analyze such strategies in terms of level of thinking style (global versus local), search targets, and six search behavior indicators and report (a) a significant relationship between different thinking style levels and individual search target types and (b) that different thinking style level conditions can cause significant differences in search behavior performance regarding maximum depth of exploration, revisited pages, and Web pages visited for refining answers. The findings suggest that high global style users tend to disperse their targets to comprehend the search task while high local style users elaborate on a few specific topics. Furthermore, high global style users skim more, require less explicit answers, and are less likely to explore an issue in depth compared to high local style or bi-high style individuals. The results confirm that thinking style level is an important factor affecting search intention. To improve search experiences, search engine designers should incorporate human factors into their products so as to take advantage of personal learning approaches.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Thinking styles; Search strategies; Information seeking; Human factors; Search interface design

1. Introduction

As one of the most prevalent applications in today’s network computing environment, Web search engines are widely used for information seeking and knowledge elaboration.
However, search-related technology has not yet reached a level of maturity, therefore academic and private researchers continue to look for “the perfect search technology” (Bat-telle, 2005). Many researchers are experimenting with ways of predicting user search intentions, with some testing new ideas on presenting information visually so as to help users locate information more efficiently. Our assertion is that the concept of thinking style—a distinguishing human factor—should be incorporated into any search engine interface design for better search intention prediction and to help users comprehend search results.

1.1. Predicting user intention for narrowing search results

Most search engines use keyword-based techniques as part of their primary interface design. This presents a problem: should users search for what they already know or what they do not know? The answer most likely lies somewhere in between—that is, most searches are for what users “partly” know, since they need prior knowledge of precise keywords in order to find the information they desire. According to Bilal (1998), users without this knowledge frequently choose imprecise keywords and therefore must adjust and re-adjust keywords and filter out large numbers of hits in order to locate information of interest. Even individuals with considerable search engine experience and/or good domain knowledge must deal with this issue.

Many search engine users—especially children and people with little Information Technology (IT) experience—have problems selecting precise keywords. Bilal and Kirby (2002) note that children usually fail to find desired information due to an inclination to use complete sentences, misspelled words, or over-generalized terms. They observe that children have problems formulating adequate or alternative keywords for completing search tasks and usually do not evaluate the quality of search results. In an attempt to help inexperienced users by predicting their intentions to create better search experiences, designers of advanced search engines such as Ask.com and A9.com recommend the use of relative search results for locating targeted or more precise information. For instance, users who type in the query “How do elephants sleep?” to Ask.com will be presented with such questions as “Why is an elephant called an elephant?” and “How do elephants eat?” This relieves users of the task of keying in relative keywords to explore core search topics.

1.2. Structured presentation of search results

Regardless of the internal algorithm employed—e.g., Bharat and Mihaila’s (2001) Hill-top, Brin and Page’s (1998) PageRank, Haveliwala’s (2002) topic-sensitive PageRank, or Kleinberg’s (1998) HITS—search results are sorted using relevance-ranking mechanisms that for the most part do not provide significant or structured presentations to help users quickly comprehend the retrieved information. Thus, users are usually required to sift through long lists of excerpts to create an overall picture of the search topic or to glean the best information. Children find it especially difficult to judge and analyze the correctness and value of search results and rarely evaluate or supplement the ones they receive (Hsieh-Yee, 2001).

Categorizing search results is one obvious solution for dealing with information overload. Clustering is one method that allows users to view categorized results without having to deal with the costs and complexities of building taxonomies (see, for example,
the Vivisimo search engine. Zamir and Etzioni (1999) made an empirical comparison of standard ranked-list and clustered presentation systems when designing a search engine interface named Grouper, and reported substantial differences in use patterns between the two. Some researchers who have experimented with highly metaphorical visualizations (e.g., Cugini, Laskowski, & Sebrechts, 2000) present users with structural overviews of result sets and promote visualization as the best approach to dealing with broad search tasks. Visualization structures of this type appear to make it easier for users to locate worthwhile information and to comprehend search results. Based on our hypothesis that thinking style can assist with user interest or intent predictions, our suggestion for search engine designers is to incorporate this human factor into their interfaces to enhance human–computer interaction.

2. Related works

2.1. Individual differences in Web searches

Web searches involve complex cognitive processes that are strongly affected by individual user characteristics. The literature contains many studies focused on differences in cognitive perspective, especially in the area of prior knowledge (see, for example, Last, O’Donnell, & Kelly, 2001; Rouet, 2003; Shapiro, 2000). Kim and Allen (2002) note that cognitive style and task type directly influence search behaviors, and Yuan (1997) adds that search experiences influence search command decisions. Holscher and Strube (2000) and Lazander et al. (2000) are among researchers who have explored differences in information search behaviors associated with different levels of information search expertise, which implies different types or strengths of cognitive factors. According to Bilal and Kirby (2002), a list of such factors should include user comprehension of the search task, individual experience with Web surfing, skill level for manipulating search engines, and the amount of attention an individual gives to a search task. All of the researchers listed in this paragraph have considered how differences in user cognitive or skill perspectives impact search behavior.

Groups of users can still develop search strategies based on shared prior knowledge. Ford, Miller, and Moss (2005) report that attitudes toward the Internet and demographic factors can also affect Web search strategies. In an earlier study, Ford and Miller (1996) observed females who were unable to find their way, frequently became lost or lacked a sense of control, and tended to only look at items suggested to them. Ford and Miller also studied how self-efficacy (in this context, indicating an individual’s judgment of his or her personal ability to find information) impacts perceptions of and approaches to information seeking. Besides human factors, researchers such as Bilal (2000, 2001), Kim and Allen (2002), and Last et al. (2001) state that search task type affects student reactions to hypertext.

The studies cited to this point allow for a summary of human factors that influence search strategies (including cognitive, affective, skill, and demographic) (Fig. 1) and to analyze how thinking style levels (an affective human factor) help determine young students’ search strategies—a topic that has not received proper attention in search behavior studies. This paper also constitutes an attempt to summarize human, search engine, and search task factors that can serve as indicators of how students interact with and respond to search engine interfaces. Combined, all of these indicators influence search strategies.
One current approach to improving the user search experience consists of providing a personalized interface; most search engines use some form of a personal (Google) or social (Yahoo) search history mechanism to achieve this. Data mining-related techniques are used to analyze search histories to recognize search patterns (interests) that reflect human factors. Human factors that can be identified as exerting significant impacts on search behaviors can be used to predict search intentions. As an important human factor that strongly affects daily personal behavior, thinking style has significant potential for impacting information seeking behavior on the Web. Thus, instead of using data mining techniques to explore raw data for recognizing user search patterns, integrating thinking style into search engine interface design may exert a much greater impact on search intention identification.

2.2. Thinking style

Thinking style refers to personal preferences in one’s abilities to deal with problems, not the abilities themselves. Accordingly, people with the same abilities may express different behaviors due to the strengths of their preferences (Sternberg, 1994, 1997). Human mental functions can be discussed in terms of five “mental self-government” dimensions: function, form, level, scope, and leaning. The function dimension involves preferences for formulating ideas, carrying out rules initiated by others, or comparing and evaluating ideas. The form dimension concerns various goal-setting and self-management behavioral styles. The level dimension distinguishes between preferences for dealing with problems at relatively abstract or detailed levels. The scope dimension includes a preference for working alone or with others. The learning dimension addresses a preference for working on tasks that involve novelty and ambiguity or tasks that require adherence to existing rules and procedures (Zhang & Sternberg, 2005).

Sternberg and Grigorenko (1995) suggest that individuals look for learning activities that match their preferred thinking style. With the advent of Internet technology, some researchers are focusing on how thinking styles impact Internet-centered learning contexts. However, to the best of our knowledge the literature does not contain any studies on the impacts of thinking style on Internet-based information seeking behavior (frequently
referred to as “search behavior”). One of our goals is to determine if a specific thinking style emerges over time when conducting Internet searches in the same manner that it emerges as part of other daily life skills and abilities.

Thinking style can affect judgments concerning immediate issues at hand. In the face of different activities that happen concurrently, individuals may initiate different goals or develop different behavioral patterns. Using goal-setting as an example, some people tend toward single-mindedness, others carefully set priorities, and still others are motivated by multiple (often competing) goals perceived as having equal importance. During the search process, some individuals are inclined to grasp the “big picture” of a search task while others focus on a few specific concepts to establish a deeper understanding. The former are satisfied with abstract issues and the latter require detail.

3. Study design

3.1. Participants

Study participants were 355 fifth grade students attending an elementary school in central Taiwan. Each student’s thinking style level was determined using a questionnaire we will describe in a later section. Of the 350 students who completed the questionnaire, 311 were instructed to use Google to search for information on pollution and to fill out a worksheet. All of the participants had two years’ worth of training in computer usage, meaning that they had basic skills with Windows, Microsoft Word, a Web browser, and Web information search techniques.

3.2. Search task

Bilal (2000, 2001) categorizes search tasks as fact-finding or research-based. Fact-finding tasks involve searches for specific answers to simple questions and research-based tasks involve searches for less clear-cut answers to more complex questions. He also notes that different search task types influence children’s cognitive and physical search behaviors. Our aim was not to address the impact of various search task types, but to analyze the impact of various strengths of thinking style level on search target settings and search behaviors. Achieving this required the use of a research-based search task to encourage students to perform more extensive searches for the purpose of attaining comprehensive understandings of their personal preferences.

The topic chosen for the participating students was “pollution”—something that Taiwanese students are well aware of in their daily lives. They had to establish initial search targets in order to attain desired results. After browsing ordered lists of search results, the students made decisions on refining their targets to move closer to their preferred results. They were asked to write down their “search targets” (i.e., Google search keywords) on their worksheets and to regularly revise their sheets according to their current search target interests. Participants were given 80 min to complete the task.

3.3. Procedure

Students were given training on basic search skills using the Google search engine. Specifically, they were asked to type in the keyword “energy resources” as practice to ensure
that they knew how to use a computer mouse and keypad to browse for information. Next, the 355 students in the original sample were asked to complete the “level dimension” of the thinking styles questionnaire described in the following section. Of the 350 students who completed the questionnaire, 311 performed searches on the topic of pollution and completed their worksheets. Searches were recorded using the Camtasia Recorder 3.0 screen capture program for further analysis.

3.4. Data collection instruments and pre-analysis

3.4.1. Investigation of thinking style level

The questionnaire used in this research was adapted from the Sternberg–Wagner Thinking Styles Inventory (Sternberg & Wagner, 1999). A modified version (Huang, 2004) suitable for Taiwanese elementary school students was created to measure the strength of the participants’ style preferences when dealing with relatively large and abstract issues (global) compared to detailed and concrete issues (local). The test consists of 10 items with answers measured along a scale of 1–5. According to the test results ($N = 311$), 72 students constituting the highest 27% of the global group were classified as high global, 66 students constituting the lowest 27% were classified as low global, and the remaining 173 students were classified as medium global. Using the same percentages, the respective numbers of students in the high local, medium local, and low local groups were 65, 184, and 62.

We used representative data due to the complexity of analyzing the search strategies and processes of 311 students. The four conditions that we created were (a) 26 students who were concurrently in the highest 27% of the global group and lowest 27% of the local group, designated as the high global style (HG) group; (b) 32 students who were concurrently in the highest 27% of the local group and lowest 27% of the global group, designated as the high local style (HL) group; (c) six students who were concurrently in the highest 27% of the global and local groups, designated as the bi-high style (Bi-H) group; and (d) six students who were concurrently in the lowest 27% of the global and local groups, designated as the bi-low style (Bi-L) group. The remaining 241 students were excluded from the search behavior analysis.

3.4.2. Investigation of student prior knowledge

To determine if the students’ prior knowledge of natural science affected the search target setting and search behavior variables, their grades for introductory natural and social science courses were collected, averaged, and used to represent their prior knowledge of the pollution topic. The 87 students in the highest 27% grade group were classified as having high prior knowledge, 81 students in the lowest 27% grade group were classified as having low prior knowledge, and the remaining 143 students were classified as having medium prior knowledge.

3.4.3. Investigation of search target settings with worksheets

Students were asked to write down their Google search engine target terms on their personal worksheets and to revise the terms as their search intentions changed. The data were quantified and recorded as number of search targets ($T$), coverage of search targets ($C$), and maximum extension of search targets ($E$). As shown in Fig. 2, the six search targets could be divided into the concept categories of “air pollution” and “noise pollution”, resulting
in a coverage value of 2. Four of the six search targets focused on air pollution and the other two on noise pollution, so the maximum extension value was 4. To apply the search targets to subsequent analyses, we divided them into three types: focused (C ≤ 2 AND E > 2), dispersed (C > 2 AND E ≤ 2), and mixed.

3.4.4. Investigation of search behavior

Files containing data on keyboard and mouse operations were reformatted into navigation flow maps (Lin & Tsai, 2005)—graphic displays of relationships among search keywords, visited Web pages, and task questions. The maps and search target settings recorded on the students’ worksheets were used to analyze their information search behaviors according to six factors adapted from Lin and Tsai: (a) number of keywords (variation in searched information); (b) visited pages (variation in task information sources); (c) maximum depth of exploration; (d) average depth of Web page adoption (average exploration depth for task completion); (e) revisited pages (degree of search navigation recursion); and (f) Web pages for refining answers (frequency of refining or improving answer quality).

4. Analysis and results

4.1. Relationship between search target setting and thinking style level

One of our goals was to determine if the participants’ prior knowledge affected their search target setting patterns (focused, dispersed, or mixed type). Results from a chi-square test indicate no significant relationship between the two variables (χ²(2) = 6.568, p = .161 > .05), therefore prior knowledge was excluded from subsequent analyses. Next, we combined the high, medium, and low global styles into a single independent variable and performed a chi-square test to identify relationships with the search target dependent variable (Table 1). The results indicate a significant relationship (χ²(2) = 25.351, p = .000 < .001). Among the low global style students, only 20.8% dispersed their search targets, 59.7% focused their attention on concept elaboration, and 19.4% showed no preference for either search target setting type. Among the medium global style students, 34.7% dispersed their search targets, 41.6% focused on similar search targets, and
23.7% showed no preference. Among the high global style students, 59.1% dispersed their search targets, 25.8% maintained a steady scope of interest, and 15.2% showed no preference.

Results from a separate chi-square test revealed a significant relationship between local style (all levels) and search target setting ($\chi^2 = 14.174, p = .007 < .01$) (Table 2). Among low local style students, 52.3% dispersed their search targets, 26.2% maintained a steady scope of interest, and 21.5% showed no preference for either search target setting. Among medium local style students, 35.9% dispersed their search targets, 44.6% focused on similar search targets, and 19.6% showed no preference. For high local style, only 22.6% dispersed their search targets, 53.2% focused on search result elaboration, and 24.2% showed no preference.

4.2. Differences among the four conditions

Our small sample size (indicating that nothing was known about the parameters of the variable of interest in the population) required the use of non-parametric methods for the following analyses. Specifically, Spearman’s $r$ was used to express relationships between two variables. Results from a Spearman’s non-parametric test failed to indicate any clear correlations between prior knowledge of the assigned search task and the six indicators listed in Section 3.4.4 (number of keywords: $r = .053$; visited pages: $r = .060$; maximum depth of exploration: $r = .181$; average depth of Web page adoption: $r = -.098$; revisited pages: $r = -.040$; Web pages visited for refining answers: $r = -.053$). Prior knowledge was therefore excluded from subsequent analyses.

Next, the four thinking style level conditions were compared in terms of the mean rank of each search behavior indicator (Table 3). Kruskal–Wallis statistical tests were performed due to the small sample size (HG: $N = 26$, HL: $N = 32$, Bi-H: $N = 6$, Bi-L: $N = 6$). The results indicate no significant differences among the conditions in terms of

<table>
<thead>
<tr>
<th>Pattern type</th>
<th>Style</th>
<th>Low local ($N = 65$)</th>
<th>Medium local ($N = 184$)</th>
<th>High local ($N = 62$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersed</td>
<td></td>
<td>52.3%</td>
<td>35.9%</td>
<td>22.6%</td>
</tr>
<tr>
<td>Focused</td>
<td></td>
<td>26.2%</td>
<td>44.6%</td>
<td>53.2%</td>
</tr>
<tr>
<td>Mixed</td>
<td></td>
<td>21.5%</td>
<td>19.6%</td>
<td>24.2%</td>
</tr>
</tbody>
</table>
the number of keywords ($\chi^2(3) = 2.191$), number of visited pages ($\chi^2(3) = 4.173$), or number of average depth of Web page adoption ($\chi^2(3) = 4.375$), but significant differences for maximum depth of exploration ($\chi^2(3) = 13.378$, $p = .004 < .001$), number of revisited pages ($\chi^2(3) = 8.604$, $p = .035 < .05$), and number of Web pages visited for refining answers ($\chi^2(3) = 9.254$, $p = .026 < .05$). In addition to identifying states of independence among the significant dependent measures, the Spearman test results indicate a correlation between maximum depth of exploration and Web pages visited for refining answers ($r_s = .301$, $p = .011 < .05$); however, no correlation was identified between maximum depth of exploration and revisited pages ($r_s = .226$), or between revisited pages and Web pages visited for refining answers ($r_s = .235$).

When Kruskal–Wallis test results were significant at the 0.05 level, Mann–Whitney $U$ tests were performed to measure contrasts between pairs of conditions. Significant pairs are listed in Table 4. A post hoc contrast of two conditions revealed a significantly higher maximum depth of exploration scores in the HL condition compared to the HG condition ($U = –3.348$, $p < .001$), suggesting that HL students tended to conduct more detailed searches in order to fully understand specific topics. For example, a depth of exploration score of 7 was earned by an HL student who found information on how air pollution was produced and how to prevent it, but an HG student only earned a score of 2 for surveying the broad topic of “water, noise, air, sea, and trash pollution”.

A separate post hoc contrast of two conditions revealed a significantly higher number of visited pages among Bi-H students compared to HG students ($U = –2.611$, $p < .001$), indicating that Bi-H students were more likely to revisit Web pages for purposes of

<table>
<thead>
<tr>
<th>Condition</th>
<th>HG ($N = 26$)</th>
<th>HL ($N = 32$)</th>
<th>Bi-H ($N = 6$)</th>
<th>Bi-L ($N = 6$)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of keywords</td>
<td>34.40</td>
<td>33.77</td>
<td>46.42</td>
<td>38.58</td>
<td>ns</td>
</tr>
<tr>
<td>Visited pages</td>
<td>31.88</td>
<td>38.56</td>
<td>44.67</td>
<td>25.67</td>
<td>ns</td>
</tr>
<tr>
<td>Maximum depth of exploration</td>
<td>27.06</td>
<td>43.70</td>
<td>38.92</td>
<td>24.92</td>
<td>$p = .004$</td>
</tr>
<tr>
<td>Average depth of Web page adoption</td>
<td>30.77</td>
<td>39.91</td>
<td>38.50</td>
<td>29.50</td>
<td>ns</td>
</tr>
<tr>
<td>Revisited Web pages</td>
<td>30.19</td>
<td>38.22</td>
<td>49.50</td>
<td>30.00</td>
<td>$p = .035$</td>
</tr>
<tr>
<td>Web pages visited for refining answers</td>
<td>30.37</td>
<td>39.53</td>
<td>44.25</td>
<td>27.50</td>
<td>$p = .026$</td>
</tr>
</tbody>
</table>

A post hoc contrast of two conditions revealed a significantly higher number of revisited pages among Bi-H students compared to HG students ($U = –2.611$, $p < .001$), indicating that Bi-H students were more likely to revisit Web pages for purposes of
knowledge elaboration than for skimming. One student in the Bi-H group revisited the same page 7 times, but an HG student only revisited the same page once and quickly moved on to other pages. A third post hoc contrast revealed a significantly higher number of HL ($U = -2.324, p < .05$) and Bi-H ($U = -2.412, p < .05$) students who visited a larger number of Web pages to refine their answers compared to HG students. We observed that one HL student made three revisions to an answer, while an HG student made only one.

5. Discussion

The study result confirm that students with different thinking style levels perform variously in terms of three search behavior indicators: maximum depth of exploration, number of revisited pages, and number of Web pages visited for refining answers. Future researchers may be interested in testing other thinking style dimensions to determine their impacts on important search behavior indicators. In order to create better search experiences by predicting user search intention, it is suggested that search engine designers consider incorporating such human factors into preference settings. For instance, after users have chosen their first keywords, instead of forcing them to filter large amounts of search results, search engines can be designed to recommend related information and/or search results that match the users’ personal thinking style levels. For HL or Bi-H users, more focused and detailed search results can be provided to support in-depth understanding or answer refinement. For HG users, related search results in other categories can be provided to satisfy their curiosity for larger or more abstract issues. For Bi-H users who tend to revisit Web pages, recent pages in personal search histories should be made accessible as part of a search result presentation (e.g., a nearby cluster or category), thus eliminating the need to redo searches for useful Web pages.

6. Conclusion

In addition to providing a review of the current literature on how human factors (cognitive, affective, skill, and demographic) influence search strategies, this paper also examined thinking style levels (an affective factor), which in the past has not received proper attention. No attempt was made to analyze how these human factors influence search strategies, but a summary was offered of human factor, search engine, and search task types that can serve as indicators of how students interact with and respond to search engine interfaces.

Our results indicate that thinking style level is indeed reflected in information seeking behavior. HG students are inclined to grasp the overall picture of a search task and HL students tend to investigate and build deeper understandings of specific concepts. Accordingly, HG students are satisfied working on a relatively abstract level and HL students prefer working with details. We therefore suggest that thinking style level influences search target setting and search behavior, and can be used in addition to or apart from data mining techniques to identify user search patterns for predicting search intentions.

The data points to a need for search engine designers to create interfaces that (a) help users narrow their searches to reduce information complexity according to their individual information needs and thinking style differences, and (b) present large bodies of search results in ways that are easier for users to comprehend. Tailoring search engine interfaces to conform to personal information needs will be an important topic for future research.
References


