# Supporting Orientation during Search Result Examination

**Henry Feild**
University of Massachusetts
Amherst, MA 01003
hfeild@cs.umass.edu

**Ryen W. White**
Microsoft Research
Redmond, WA 98052
ryenw@microsoft.com

**Xin Fu**
LinkedIn
Mountain View, CA 94043
xfu@linkedin.com

## ABSTRACT

Search engines help their users decide which results to visit using captions comprising titles, URLs, and snippets containing the query keywords and proximal text from landing pages (the search results linked from the result page). Although caption content can be a key factor in these decisions, snippets provide only basic support for orienting users with landing page content from the search-engine result page (SERP), and no support during the transition to landing pages or once users reach the page following a selection decision. As a result, many searchers must employ inefficient strategies such as skimming and scanning the content of the landing page. In this paper we propose a novel method, called *clickable snippets*, to address this shortcoming. Clickable snippets provide searchers with a direct and actionable link between SERP captions and landing-page content. We describe a user study comparing clickable snippets with extant methods of orientation support such as query-term highlighting on the landing page and thumbnail previews on the SERP. We show that clickable snippets are preferred by participants, and lead to more effective and efficient searching. Our findings have implications for the design of the user experience in search systems.

### Author Keywords

Clickable snippets; Orientation; Search-result examination

### ACM Classification Keywords

H.3.3. [Information Storage and Retrieval]: Information Search and Retrieval–*search process*, *selection process*.

## INTRODUCTION

In response to search queries, Web search systems typically display lists of search result captions comprising surrogate information. The captions may have titles, document snippets containing query terms and their context from the landing page, and URLs. The captions are optimized to facilitate result selection decisions and the development of better snippets has been researched extensively [8,25]. However, if searchers base result selection decisions on snippets, it is likely that they will want to pursue snippet content on the landing pages (we show this is the case). Beyond supporting result selection, search systems must also help searchers situate or *orient* themselves within results they visit.


*CHI 2013,* April 27–May 2, 2013, Paris, France.


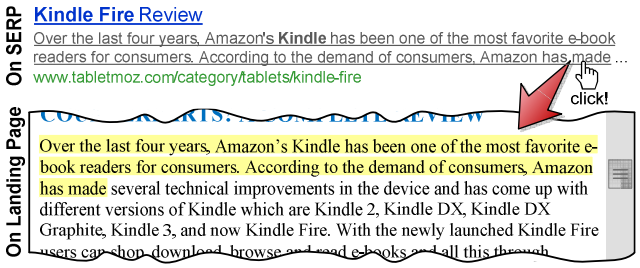


**Figure 1. Clickable snippet on SERP. User clicks on the snippet and is taken direct to that text on landing page.**

There is some limited support for search-result orientation. Browsers offer find-in-page (*Ctrl-f*) functionality, but this is seldom used and training users to adopt this support is difficult [4,17]. Snippets provide some clues about the context of the keywords, but the association between the snippet text and its location in the document is often unclear and does not persist during the transition since searchers are always taken to the top of the landing page. Query-term highlighting is offered on landing pages via search toolbars, or search engine cached pages, but the relationship with the snippet text that motivated the click may be unclear and users may miss it, especially if they spend only a little time reviewing pages [16]. Users may also find such highlighting distracting and obtrusive [7]. Passage highlighting has been explored, but not in terms of SERP-to-landing page transitions [20]. Thumbnail previews [23,26] can offer a condensed snapshot of the landing page, sometimes augmented with the location of the snippet (as currently used by Google, Figure 3). However, the previews may only be used to assess the overall relevance or visual appeal of landing pages [1] and these provide limited support for the *transition* between SERP and landing page, requiring users to remember where on the page the snippet occurred.

To address the shortcomings of existing approaches, we developed a technique called *clickable snippets* that supports orientation at different phases of search result examination. It adds an affordance to SERP snippets giving users the option to click snippet text and transition to it on the landing page (Figure 1). The transition occurs immediately or gradually (over the course of 1-2 seconds), with the goal of the latter reinforcing the relationship between the clicked snippet text and the landing page content. We conjectured that clickable snippets may help people more easily locate information of interest on landing pages. We conducted a 48-participant user study comparing different orientation techniques to assess how their support on the SERP, on the landing page, and

during the transition impacts user preferences, search efficiency and search effectiveness. We show that clickable snippets were preferred by most participants and yielded significant performance gains when employed.

## RELATED WORK

There are four relevant areas of prior research to this study: result summarization, previewing and transitioning to landing pages (results), highlighting terms and snippets on landing pages, and finding relevant landing-page content.

When examining search results, users need some amount of information about each result to decide whether or not they should visit it. Snippets showing sentence fragments matching one or more query terms have become popular and have been found to be useful for search tasks [25] and in information access interfaces in general [15]. Clarke et al. [8] found that results whose snippets were readable and had all query keywords were more likely to be visited by searchers.

Despite their promise, result snippets are typically short, terse, and incomprehensive. Visualizations of result pages have been proposed to aid selection decisions. TileBars [14], for example, simultaneously and compactly indicate relative document length, query term frequency, and query term distribution. The patterns in a TileBars column can be quickly scanned and deciphered, aiding users in making judgments about potential result relevance.

Thumbnails have also been examined to provide additional information about the content of landing pages. Woodruff et al. [26] studied the use of text snippets, unaltered thumbnails, and enhanced thumbnails of landing pages. The enhanced thumbnails, including various treatments such as highlighting and magnifying keywords, led to faster search completion times across several task categories. Teevan et al. [23] examined how different representations of Web pages affected people's ability to recognize new relevant Web content and return to previously-viewed Web pages. Their findings suggested that text snippets help find unseen pages and thumbnails help re-find pages for which users have seen the thumbnail. Aula et al. [1] showed that thumbnails add information about the relevance of Web pages that is not available in their textual summaries. However, showing only thumbnails, with no text information, resulted in poorer performance than showing only textual summaries.

While providing thumbnails can give users a general feel of a page's layout with respect to a set of search terms, it does not provide detailed content. Paek et al. [21] explored a system called WaveLens, which displays dynamic search result snippets that increase in font size and content when hovered over. This allows more results to be shown and richer, more context-filled snippets to be selectively revealed, helping users make better decisions about which results to visit directly from the SERP. The authors found participants preferred the system to static result lists. Beyond search, Zellweger et al. [27] experimented with *fluid documents* which display additional information about the content of embedded hyperlinks on a page when hovered. In a fluid documents system, no explicit support is provided to help users transition to content-of-interest on the landing page once they decide to click.

Once searchers transition to a landing page, they need to locate the particular information of interest. To support this, search engines provide query-term highlighting in browser toolbars and cached pages in result lists. The use of highlights for reviewing and recall of specific information has been suggested in many systems, studies of which have typically found users are faster with highlighting than without [2,12,20,26]. Chi et al. [6] introduced a method that not only highlights keywords, but also highlights sentences if they contain conceptual keywords relevant to query topics. Chi et al. [7] studied the impact of highlighting on gaze. They found that highlighting related passages captured more fixations and resulted in more accurate answers than keyword or no highlighting. Google Quick Scroll [11], an extension for the Chrome browser, disseminates passage highlighting more broadly. On landing pages visited via a Google SERP the system highlights and scrolls to snippets shown on the SERP. However, this does not help orient the user on the SERP nor in the initial landing-page transition.

The most widely available support for orientation on landing pages is Web browsers' find-in-page (*Ctrl-f*) functionality. Although people regard this as an advantage of electronic documents, they rarely use the feature [4,17]. Beyond highlighting, others have studied methods such as the application of relevance profiles or text segmentation to find relevant content in documents. ProfileSkim [11] computes a relevance score for each section and provides overview and within-document navigation options that highlight the most relevant parts of the document. This support is only offered once the user navigates to the document and does not help with the transition between SERP and landing page. Carracciolo and de Rijke [5] developed a "go-read-here" retrieval functionality, which points users to a segment where they can best start reading to learn about the topic of interest. However, their focus was on text segmentation technology rather than the user experience, as we focus on here.

Our research extends previous work in a number of ways. First we focus on orientation during the search process, not just the estimation of relevance from SERP content. Second, we propose clickable snippets, a novel method for orientation during search result examination. Third, we focus on support for orientation at all three phases of result examination: on the SERP, during the transition to the landing page, and on the landing page. Finally, we perform a comparative evaluation of different orientation support methods, including term highlighting and SERP result previews, both of which are offered by current search engines.

## CLICKABLE SNIPPETS

Clickable snippets complement existing orientation methods and direct users from the SERP to the snippet text on the landing page. The support is in three phases.

1. SERP: On the result page, each result is represented by a caption comprising the page title, a query-focused snippet from the landing page, and the URL. The only difference with traditional results is that the snippet text has an additional affordance that allows users to click and be taken directly to that text in the document. For discoverability, we convey this affordance to users by underlining the snippet text as though it were a hyperlink (Figure 1 has an example).

Note that only the snippets with matching content are underlined. If a snippet does not appear in the landing page, it is not underlined on the SERP and is not clickable. There are a number of reasons for snippet mismatches, including if the snippet is based on the HTML META description tag or updates to page content since the engine last crawled it. For the queries used in our study described later, snippets were underlined for 61% of the results on the SERP.

2. Transition: There is either an immediate or a gradual transition from SERP to landing page. In the former case, the landing page is automatically scrolled to the part of the page containing the clicked snippet text and displayed to the user. In the latter case, the transition is performed as a fluid sequence of fades and layer moves. As Figure 2 illustrates, there are four stages to the transition: (i) *isolate* the snippet and fade out other captions; (ii) *scroll* the landing page so that the snippet text is visible in the viewport (although the page is not yet shown); (iii) *move* the snippet so it overlays its source on the landing page; and (iv) *reveal* the landing page by fading it into view while at the same time fading out the snippet. The gradual transition takes under two seconds. It is designed to clarify the relationship between the snippet and the content.

3. Landing page: Once the user transitions to the landing page, the clicked snippet text is highlighted (Figure 5).

## STUDY

We now describe our study to evaluate the performance of different orientation methods, including clickable snippets, at different phases of the SERP-to-landing-page transition.

### Systems

We tested six systems. Each system provided a way of orienting users at each of the three phases outlined in the previous section. Some systems shared the same orientation support at particular phases, but each system had a unique combination of support across all three phases.

#### *Baseline*

Similar to current search engines, this system offers basic orientation support via titles, snippets, and URLs on the SERP with no support on landing pages or in the transition.

#### *ThumbnailPreview*

This system is similar to *Baseline*, but offers a preview of the landing page on the SERP to the right of a result caption when a user hovers on it and clicks a chevron (») shown adjacent to the caption after a short delay. The preview provides an image of the landing page which is similar in size to the preview currently offered on Google SERPs (i.e., 300 × 400 pixels). In a similar way to the Google previews, the snippet text is shown in a callout adjacent to the thumbnail and in a bounding box in the thumbnail itself. Figure 3 has an example of the thumbnail and snippet callout.

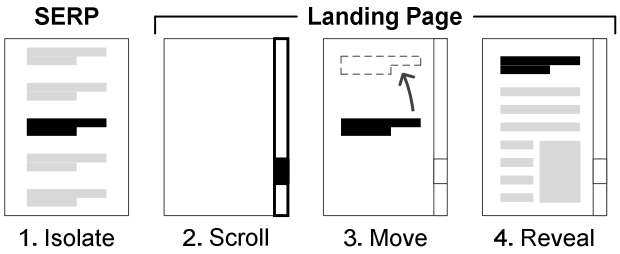


**Figure 2. Schematic illustrating the steps of the SERP to landing page transition in the *ClickableSnippets Gradual* system.**

#### *TermHighlighting*

This system is similar to *Baseline*, but differs in that on the landing page it highlights all instances of each query term in a different color (see Figure 4 for an example).

#### *SnippetHighlighting*

This system is similar to *Baseline*, but differs in that it highlights the snippet text on the landing page (Figure 5).

#### *ClickableSnippets Gradual and Immediate*

Two variants of the clickable snippets system as described in the previous section.

### Research Questions

The goal of this study was to examine how the orientation methods were perceived by users and their impact on search effectiveness. There were three specific research questions:

RQ1 (Preview Strategy): Which type of orientation support from the SERP is most effective and most positively perceived? We study three methods: (i) basic (systems: *Baseline*, *TermHighlighting*, *SnippetHighlighting*); (ii) previews (*ThumbnailPreview*); and (iii) clickable snippets (*ClickableSnippets Gradual* and *Immediate*).

RQ2 (Highlighting Strategy): Which type of orientation support from the landing page is most effective and most positively perceived? We study three methods: (i) none (*Baseline*, *ThumbnailPreview*); (ii) keyword highlighting (*TermHighlighting*); and (iii) snippet highlighting (*SnippetHighlighting*, *ClickableSnippets Gradual* and *Immediate*).

RQ3 (Transition Strategy): Which type of orientation support during the SERP-to-page transition is most effective and most positively perceived? We examine three methods: (i) basic (*Baseline*, *TermHighlighting*, *SnippetHighlighting*, *ThumbnailPreview*); (ii) gradual (*ClickableSnippets Gradual*); and (iii) immediate (*ClickableSnippets Immediate*).

### Study Design

To answer the questions, we performed a within-subjects experiment. Participants completed two distinct tasks on each system, for a total of 12 tasks per participant. Each task consisted of a question, a fixed query, and 10 results. The results originated from a commercial search engine, were scraped before the start of the study, and fixed for all participants. The correct answer was present in exactly one result. No questions were answerable with only the SERP snippet text.

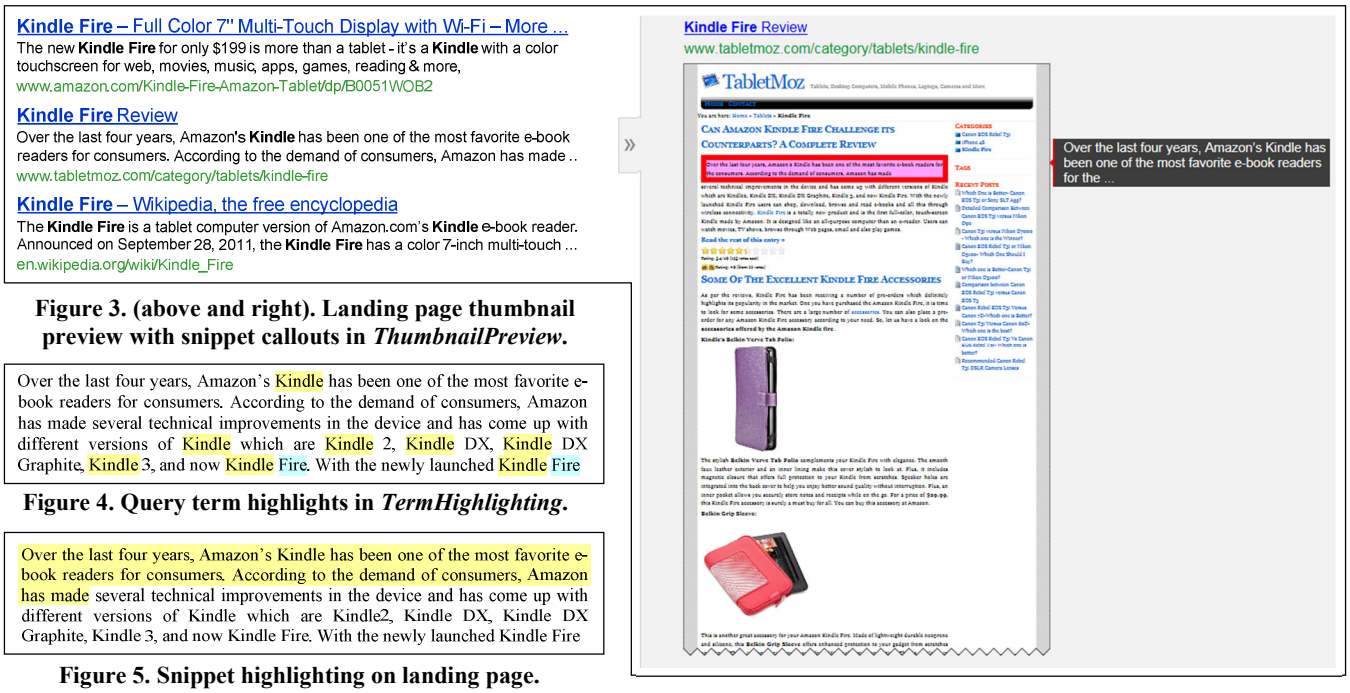


**Figure 3. (above and right). Landing page thumbnail preview with snippet callouts in *ThumbnailPreview*.**

Over the last four years, Amazon's Kindle has been one of the most favorite e-book readers for consumers. According to the demand of consumers, Amazon has made several technical improvements in the device and has come up with different versions of Kindle which are Kindle 2, Kindle DX, Kindle DX Graphite, Kindle 3, and now Kindle Fire. With the newly launched Kindle Fire

**Figure 4. Query term highlights in *TermHighlighting*.**

Over the last four years, Amazon's Kindle has been one of the most favorite e-book readers for consumers. According to the demand of consumers, Amazon has made several technical improvements in the device and has come up with different versions of Kindle which are Kindle2, Kindle DX, Kindle DX Graphite, Kindle 3, and now Kindle Fire. With the newly launched Kindle Fire

**Figure 5. Snippet highlighting on landing page.**

We controlled for learning effects by randomizing the order in which participants used tasks and systems. Eight participants attempted each task-system pairing.

### Procedure

Participants completed the study using a computer outfitted with an eye tracker. Each study lasted one hour. Before each experiment began, the experimenter gave a written overview of the study and then calibrated the eye tracker.

Participants then attempted the tasks using the appropriate systems in a pre-determined order. *For each system*, participants were given a description of its features, an example SERP to see how it worked, the two assigned tasks, followed by a questionnaire qualifying their experience on the system, including eliciting a relative ranking all of the systems they had used up to that point. *For each task* (two per system), participants were instructed to find the answer to a question given the fixed SERP or to abandon if they believed they could not complete, but had spent a reasonable amount of time searching. Participants provided answers through a study dialog. They were then asked to complete a questionnaire about the task. At the end of the study, participants completed a questionnaire providing summary feedback on the systems, including a final system ranking, an explanation for it and suggested improvements. So as to not reveal too much about how they varied, we used pseudonyms comprising a random color and a name (e.g., "blue sky", "black bear") to refer to the systems in experimental materials.

### Participants

There were 48 participants (24 males, 24 females). Participants were randomly selected from an external pool of volunteers managed by the central usability group in our organization. Participants were screened in a phone interview to ensure that they had uncorrected vision (for eye tracking), had used a search engine previously, and had used a search engine at least once in the month before the study. The search experience screening helped ensure that participants were familiar with current search engine technology.

Participants completed an entry questionnaire that elicited limited demographic information and background about their search experience. They were generally in the 22-34 age range (43.8%), most had been searching for 10 years or more (79.2%), and most reported searching 10 or more times per day (60.4%). We also asked about how they oriented themselves on landing pages (we report on that later).

### Tasks

We wanted to assign representative Web search tasks. To do this we started with a week of query-click logs from the Bing search engine (from February 19-25, 2012) and manually selected informational queries, associated with an information-acquisition intent [3], that we believed would make interesting tasks. To identify candidate informational queries, we favored frequent queries with high click entropy (i.e., high variance in the results selected across many users [9]) since they are likely to be informational in nature [24]. We focused on informational queries since they often describe search scenarios where a particular answer is being sought, which is when orientation support may be helpful.

A total of 12 tasks were identified from the set of candidate informational queries. Figure 6 has an example task.

To control experimental variability, we fixed the initial query for each task so that it could not be modified. At most one of the results in the top-10 contained the exact answer to the question. For some queries, we also adjusted the ordering of the result lists so that the page with the answer was distributed over the ranks across search tasks, ameliorating the ef-

***Query (shown to participant on SERP):*** *[kindle fire]*
***Description:*** *Aside from the Kindle Fire, what were four versions of the Kindle that Amazon released to address technical improvements demanded by users?*
***Answer:*** *Kindle 2, Kindle DX, Kindle DX Graphite, Kindle 3*

**Figure 6. Example task used in study, along with the answer and initial query shown to participants, from search log data.**

fects of rank biases. To counteract effects related to the position of the answer on the landing page, we divided the tasks into two categories: for half the tasks, the answer to the question required the landing page with the answer be scrolled (i.e., the answer occurred below the page fold on the browser used for the study); for the other tasks, no scrolling was necessary. Around 7-8 result captions were shown on the SERP before participants needed to scroll.

### Data Capture

We used different methods to capture participant preferences, actions, and task outcomes. We now describe each.

#### *Surveys*

In total, four surveys were created: (i) entry (administered at the start of the experiment), (ii) task (after each task), (iii) system (after each system), and (iv) exit (at the end of the study). The experiment was controlled by a study interface that guided participants through the steps of the experiment.

#### *Web Browser and Answer Provision*

For each task, participants were given the description and a button to open the task interface in a new window. This contained a navigation bar with "Back" and "Forward" buttons, the system description, the task question, an area to place the answer, and a "Done" button. Below the bar was the SERP containing results for the task's fixed query. Participants were allowed to visit the results freely, but links on the landing pages were disabled, preventing participants from diving deeper into the host site. Once they found the answer, they were instructed to fill in the answer box and click the "Done" button, bringing them back to the main study interface. Participants could use the find-in-page functionality using the *Ctrl-f* shortcut or via the browser File menu, which was made visible. To avoid biasing behavior, we did not explicitly remind participants about find-in-page. All pages were prefetched before the study to remove landing-page variability.

#### *System Logging*

We used JavaScript embedded in SERPs and landing pages to log behavior such as clicks and cursor movements. We also manually marked up the landing pages, including adding keyword highlighting, snippet highlighting, and the bounding box of the answer passage (used for later gaze-fixation analysis). We did this manually to ensure correctness and isolate the effect of the orientation methods.

#### *Gaze Tracking*

In addition to tracking the position of the mouse cursor on the SERPs and landing pages, we also employed gaze tracking to monitor eye movements and fixations on the study Web pages. We used a Tobii TX300 eye tracker using 60 Hz tracking frequency and an accuracy of 0.4° visual angle (corresponding to 13 pixels in our setting) on a 1920 × 1080 pixel 23-inch monitor (96 dpi). Gaze position was recorded with respect to the upper left-hand corner of the landing page approximately every 16 milliseconds, facilitating accurate estimates of saccades and fixations on the page.

## FINDINGS

We now present the findings of our study. Parametric and non-parametric statistical testing is performed where appropriate, with $\alpha = 0.05$ unless otherwise stated. Bonferroni corrections adjust $\alpha$ to reduce the likelihood of Type I errors, i.e., incorrectly rejecting null hypotheses, by dividing $\alpha$ by the number of variables. Five-point scales are used to measure preferences, with a higher rating indicating more agreement with attitude statements.

### Current Orientation Strategies

At the outset of our studies we wanted to understand how people oriented themselves during search result examination. As part of the entry questionnaire we described titles, snippets, and URLs and asked participants: *How useful do you generally find each of these in deciding which result to select?* The response options were on a five-point scale ranging from *very useful* (rating=5) to *not useful at all* (rating=1). The responses, summarized in Table 1, reveal that participants considered the titles and the snippets similarly useful, and more useful than URLs. A Friedman test revealed that these differences were significant ($\chi^2(2) = 30.81$, $p < 0.001$; title/snippet vs. URL: $Z \geq 2.20$, $p < 0.01$ with Dunn's post hoc tests; title vs. snippet: $Z = 0.41$, *ns*).

Next we wanted to understand how users transitioned from a SERP to landing pages. Unlike titles (which may be synthesized from other content such as anchor text), snippets are drawn directly from landing pages. We therefore wanted to understand: (i) if searchers sought out snippet content on landing pages (i.e., once they clicked on a result), and (ii) if so, how they located snippet content on landing pages. We asked participants: *How often do you attempt to find text from a result caption on that result's landing page?* Response options: *never*, *for some searches*, *for about half of my searches*, and *for almost all my searches.* Many participants

**Table 1. Participant perceptions of the utility of each of the three caption elements: titles, snippets, and URLs. *N*=48.**

| Caption element | Perceived utility | |
|---|---|---|
| | M | SD |
| Title | 4.46 | 0.88 |
| Snippet | 4.38 | 0.94 |
| URL | 3.44 | 1.03 |

**Table 2. Percentage of participants reporting using each strategy for locating information on landing pages. Parenthesized values show the number of participants. *N*=38.**

| Method | % participants |
|---|---|
| Direct only | 10.53% (4) |
| Indirect only | 68.42% (26) |
| Both | 21.05% (8) |

**Table 3. Participant perceptions of different aspects of the orientation strategies of each system. Ratings are on a five-point scale, with higher ratings being more positive. Bolded values are the highest values in each row. *N*=96 in each of the cells.**

| Orientation location | Measure | Baseline | | Term Highlighting | | Snippet Highlighting | | Thumbnail Preview | | Clickable Snippets | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Gradual | | Immediate | |
| | | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD |
| SERP | Assisted selection | 3.83 | 1.21 | 3.79 | 1.24 | 3.80 | 0.99 | 3.91 | 1.15 | **4.26** | **1.01** | 4.23 | 1.00 |
| | Set expectations | 3.80 | 1.08 | 3.77 | 1.19 | 3.80 | 0.99 | **4.30** | **1.03** | 4.02 | 1.09 | 3.97 | 1.13 |
| Transition | Helped transition | 3.46 | 1.07 | 3.20 | 1.12 | 3.71 | 0.82 | 3.71 | 0.76 | **4.22** | **0.88** | 4.16 | 0.95 |
| | Locate relevant info. | 3.02 | 1.18 | 3.12 | 1.20 | 3.52 | 0.92 | 3.39 | 1.04 | **3.95** | **0.98** | 3.87 | 1.02 |
| | Caption-page assoc. | 3.49 | 1.07 | 3.52 | 1.22 | 3.87 | 0.94 | 3.63 | 1.14 | **4.35** | **0.92** | 4.09 | 0.96 |
| Landing page | Ease of finding info. | 3.45 | 1.04 | 3.44 | 1.27 | 3.81 | 0.94 | 3.59 | 1.17 | **4.11** | **0.85** | 3.88 | 0.96 |
| | Speed of finding info. | 3.35 | 1.04 | 3.42 | 1.29 | 3.76 | 1.00 | 3.47 | 1.21 | **4.09** | **0.95** | 4.04 | 0.90 |
| | Highlight obtrusive | n/a | n/a | **3.04** | **1.43** | 2.13 | 1.13 | n/a | n/a | 2.09 | 1.18 | 2.02 | 1.01 |
| | Highlight useful | n/a | n/a | 2.75 | 1.38 | **3.58** | **1.38** | n/a | n/a | 3.56 | 1.41 | 3.46 | 1.32 |

(56.3%, *N*=27) reported trying to find snippet content on a page for at least half of their searches, and none reported never doing this, suggesting that support for orientation during search is an important area of research focus.

We asked participants to describe in free text how they find snippet content on landing pages. By not listing possible methods, we avoided biasing their opinions. We received a range of responses (e.g., "using the find function, and 'glance-reading' over the page" and "by scrolling and scanning the landing page"). We hand-coded the responses and identified two emergent themes: (i) *direct*, where people use the browser find-in-page functionality, and (ii) *indirect*, where people scan or skim-read the page, leveraging headings and section breaks. The percentage of participants who reported using each strategy is summarized in Table 2. Note that eight participants did not respond to this question and two provided spurious responses, which were dropped.

The findings suggest that direct navigation on the landing page is fairly common, but not as common as indirect navigation. Skimming and scanning is inefficient and since users often make rapid Web page assessments [16,18] indirect navigation may lead to missed answers. Searchers may benefit from rapid direction to relevant landing-page regions.

We now report our findings along different analysis dimensions related to our three research questions. We compare the six different systems using a repeated measures analysis since it is potentially more powerful than an independent measures analysis of the three system groups, and allows for a finer-grained comparison of different systems at each orientation phase. We begin with participant perceptions.

### Participant Perceptions of Orientation Support

We analyzed participant responses to the system questionnaire. In addition to asking people for their responses to attitude statements, we also asked them to describe the features that they liked best and least about each system. We present quotes from participant responses to those questions when we believed that they may help explain the findings.

#### *Effect of Preview Strategy*

To understand the effect of the result preview strategy we asked participants to indicate on Likert scales (5 = strongly agree, 1 = strongly disagree) their agreement with the following: (i) *Result captions helped you decide whether to click on a particular result*; and (ii) *Result captions provided you with insight about what to expect before you visited a result's landing page*. Responses are summarized in the top two rows of Table 3 ($\alpha = 0.025$). They show that the clickable snippets systems helped people decide what results to visit (Friedman: $\chi^2(5) = 15.28$, $p = 0.01$; Dunn's post-hoc tests: all $Z \geq 2.23$, all $p < 0.01$). *Baseline*, *TermHighlighting*, and *SnippetHighlighting* all performed similarly (all $Z \leq 0.20$, *ns*). This was expected since there were no differences in the SERP presentation method for these three systems. The strong performance of the clickable snippets systems may be because they underlined the snippet text present on the page (as a hyperlink) clarifying what they would see before they went to the page (e.g., one participant stated "snippet underlining gave clues about whether snippet was on page"). *ThumbnailPreview* was most effective at setting expectations before visiting the page ($\chi^2(5) = 15.87$, $p < 0.01$; all (*ThumbnailPreview* vs. others): $Z \geq 2.52$, all $p < 0.01$) (e.g., one participant stated "previews were nice because I could see if the page structure was likely to include what I was looking for").

#### *Effect of Transition Strategy*

We also asked participants specifically about the transition between SERP and landing page using Likert scales and these statements: (i) *The system helped you transition between search results and landing page*; (ii) *The transition assisted you in locating relevant information on the landing page*; and (iii) *There was a clear association between the information you saw in a result caption and the result's landing page.* The results are summarized in rows 3-5 of Table 3 ($\alpha = 0.017$). Overall, the results show the clickable snippets systems were preferred (all $\chi^2(5) \geq 16.31$, all $p < 0.01$; Dunn's post-hoc tests: all $Z \geq 2.28$, all $p \leq 0.01$). Participants preferred being taken to the snippet rather than the top of the landing page and the gradual transition was favored over immediate (all $Z \geq 2.26$, all $p \leq 0.01$) even though it took longer. This may be because it more clearly connected snippets with the landing page, as evidenced by the comments (e.g., "highlighted connections between search text and results" and "gives visual context of where to look").

**Table 4. Features of participant engagement with each system. The number of landing page views is averaged at the task level, whereas the others are averaged at the landing-page level. Bolded values are the lowest (most positive) values in each row.**

| Feature | Baseline | | Term Highlighting | | Snippet Highlighting | | Thumbnail Preview | | Clickable Snippets Gradual | | Clickable Snippets Immediate | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD |
| # landing page views | 2.79 | 2.61 | 2.88 | 2.61 | 2.92 | 2.63 | **2.41** | **2.30** | 2.82 | 2.52 | 2.91 | 2.56 |
| Scan path length (px) | 54.5k | 55.6k | 75.0k | 59.1k | 51.7k | 55.5k | 60.8k | 57.2k | 46.5k | 52.8k | **45.7k** | **47.5k** |
| Time until fixation on answer passage (secs) | 1.88 | 1.53 | 1.71 | 1.42 | 1.46 | 1.75 | 1.89 | 1.84 | **1.08** | **1.22** | 1.13 | 1.11 |
| # fixations | 21.12 | 18.86 | 25.36 | 23.32 | 19.65 | 21.47 | 21.61 | 21.03 | **15.67** | **15.45** | 16.45 | 16.90 |
| Scroll distance (px) | 3.8k | 5.3k | 4.6k | 6.9k | 3.2k | 5.1k | 4.6k | 5.9k | **2.6k** | **3.8k** | 2.9k | 3.4k |

The results show preference differences between the four systems with the basic transition strategy (i.e., always taken to top of landing page). Support on the SERP and the landing page may help transitions without supporting them directly. The results show that: (i) *SnippetHighlighting* and *ThumbnailPreview* helped users transition more than other systems (all $Z \geq 2.33$, all $p < 0.01$) whereas the *TermHighlighting* system performed worse than the others (all $Z \geq 2.36$, all $p < 0.01$), and (ii) *SnippetHighlighting* helped users locate relevant information and established clearer associations between captions and page content (all $Z \geq 2.31$, $p < 0.01$).

*Effect of Highlighting Strategy*

Once users clicked on a result caption, they were taken to the landing page. We now focus on their experience once they reached that page. We asked participants: (i) *How easy was it to determine if a landing page was useful for a task*? and (ii) *How quickly were you able to determine if a landing page was useful for a task*? Participant responses are summarized in rows 6-7 of Table 3. The results show that the four systems using snippet highlighting were preferred over *TermHighlighting* and the no-highlighting systems ($\chi^2(5) = 15.51$, $\alpha = 0.025$; $p < 0.01$; all (snippet highlighting systems vs. others): $Z \geq 2.54$, all $p \leq 0.005$). Mirroring previous work [7], we found that highlighting all query terms on the landing page was obtrusive; 60.4% (29 of the 48 participants) found the term highlighting distracting (e.g., "highlighting [was] extremely annoying, distracting and not useful").

In the system questionnaire for the highlighting systems we asked whether participants noticed the highlights. 91.7% of participants reported noticing term highlights, and 75.7% noticed snippet highlights. We also asked participants about the utility and the obtrusiveness of highlighting. As the ratings in Table 3 show ($\alpha = 0.025$), participants found term highlighting more obtrusive than snippet highlighting, perhaps because there were often a large number of highlighted query terms on the landing page (see Figure 3) ($\chi^2(3) = 14.70$, $p < 0.01$; all $Z \geq 2.30$, all $p \leq 0.01$). The results also show that participants found snippet highlighting systems more useful than *TermHighlighting* or no highlighting, perhaps because it drew their attention to the part of the page containing the information they sought when they clicked the search result.

In the next section we turn our attention to our participants' levels of engagement with the SERP and the landing pages.

### Engagement with SERP and Landing Pages

*Effect of Preview Strategy*

On the SERP we focused on clicks on the titles, clicks on the snippets (where available), and the use of the thumbnail previews. The average number of clicks on SERP titles on systems not offering additional assistance on the SERP was 2.86 (SD = 2.65). When clickable snippets were available, the total number of clicks was similar (M = 2.87, SD = 2.58), but clickable snippets cannibalized title clicks: 28.3% of result clicks from SERPs to landing pages come from clickable snippets. There were on average 0.65 (SD = 1.05) snippet clicks per task. 79.2% ($N$=38) of participants used the clickable snippets for 42.2% of the tasks. The average number of result previews viewed per task in *ThumbnailPreview* was 7.61 (SD = 6.27). In total, 58.3% ($N$=28) of participants used the previews for 46.6% of the tasks.

Two statistics are worth computing: (i) the total number of unique page visits on each system and (ii) for clickable snippets, the fraction of successful clicks on the titles versus snippets. The former captures the effect of preview strategy on page selections and the latter helps gauge the benefit to users of engaging with the snippets. The number of page views for each system is shown in the first row of Table 4. The findings show that the total number of landing page views is slightly higher in the systems that do not offer additional orientation support on the SERP, although the differences are not significant with one-way analysis of variance (ANOVA) ($F(5,570) = 0.99$, $p = 0.42$, $\alpha = 0.01$). Comparing clickable-snippet-click versus title-click outcomes, we see that when participants clicked on titles they found the correct answer (and terminated the task) for 69.2% of clicks, whereas clicking on a clickable snippet increased task success to 80.1%. When participants transitioned to the landing page via a snippet, they were more likely to find the answer (McNemar's $\chi^2(1) = 9.44$, $p = 0.003$). Further support was obtained via the phi correlation ($\varphi$) between whether the snippet was clicked (1/0) and whether the participant found the answer on the page (1/0). The value of $\varphi$ was 0.81, signifying a strong relationship between clickable snippet use and task success.

**Table 5. Features of task completion on each system. Values averaged across all 96 tasks on each of the systems. Success is rated via self-report on a five-point scale, with higher ratings being more positive. Bolded values are the highest values in each row.**

| Feature | Baseline | | Term Highlighting | | Snippet Highlighting | | Thumbnail Preview | | Clickable Snippets Gradual | | Clickable Snippets Immediate | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD |
| Success level | 4.27 | 1.31 | 4.49 | 1.10 | 4.46 | 1.19 | 4.47 | 0.93 | **4.84** | **0.82** | 4.81 | 0.95 |
| Time until relevant page visit (min:sec) | 1:36 | 1:23 | 1:35 | 1:24 | 1:37 | 1:23 | **1:55** | **1:25** | 1:13 | 0:57 | 1:16 | 0:55 |
| Duration (min:sec) | 2:44 | 1:55 | **3:23** | **1:47** | 2:44 | 2:01 | 2:43 | 1:53 | 2:22 | 1:53 | 2:27 | 1:58 |
| Answer page visit | 68.75% | | 68.75% | | 68.75% | | 71.87% | | 73.96% | | **73.96%** | |
| Answer correctness | 63.54% | | 64.58% | | 64.58% | | 66.67% | | **72.92%** | | 71.88% | |
| Answer found on page | 92.37% | | 93.93% | | 93.93% | | 92.76% | | **98.59%** | | 97.19% | |

*Effect of Transition and Highlighting Strategies*

We computed a number of features of user interaction with the landing page, averaged per landing page:

- *Scan path length*: Total distance traveled by the eye when examining the landing page (in pixels). Since our eye tracker recorded gaze positions 60 times per second, microsaccades were also included in the scan path length.
- *Total number of fixations*: Total number of fixations on the landing page, where fixations are identified using the velocity threshold identification (I-VT) filter used in Tobii Studio software [10,19].
- *Time until fixation on answer passage*: Time from landing page load until first fixation in answer bounding box.
- *Scroll distance*: Total number of vertical pixels that the user scrolls when examining the landing page; a summation of both upward and downward movement.

The values for these features for each system are shown in the last four rows of Table 4 ($\alpha = 0.01$). In combination, these measures provide an estimate of the amount of effort users employed to find information on landing pages. They show that the clickable snippets systems helped reduce the scan path length, required less scrolling (both $F(5,1600) \geq 4.17$, both $p < 0.001$), and helped users find the answer on pages faster ($F(5,403) = 3.66$, $p = 0.003$), verified with Tukey post-hoc tests. The post-hoc tests show no difference between the two clickable snippets systems for any features (all $p \geq 0.34$). In contrast, the highlighting strategy affected a number of aspects of search interaction. Snippet highlighting led to a reduction in the time until users fixated on the answer passage on the landing page ($F(5,403) = 3.30$, $p = 0.007$) and reduced the scan path length ($F(5,1600) = 3.34$, $p < 0.001$). Other differences between *SnippeHighlighting* and other systems were not significant. In contrast, *TermHighlighting* appeared to have a negative effect on participants' interactions with the landing page: Table 4 shows that scan path length, total number of fixations, and the time to the first fixation on the answer passage were all longer with *TermHighlighting*, signaling that participants may have experienced difficulty in locating information on the landing page (all $F(5,1600) \geq 3.56$, all $p < 0.001$).

*ThumbnailPreview* led to longer gaze trails than all systems other than *TermHighlighting* (Tukey post-hoc tests: all $p \leq 0.01$), a similar time to fixate on the answer passage as the baseline ($p = 0.29$), and a longer time than the other systems (all $p \leq 0.01$), even though it provided an indication of where the relevant content resided on the landing page. Participants were more focused on using the previews to obtain a general sense of landing-page layout (e.g., "can see if the page has paragraphs or images before clicking") and quality (e.g., "you can tell quickly whether a site is a spam site or not") rather than where the snippet appears in the landing page and remembering that location during the transition.

### Task Completion

As well as studying how participants perceived the systems and their engagement, we were also interested in system effects on task completion. We were specifically concerned with the following: (i) whether they reached a landing page that contained the answer; (ii) the time from the start of the task until they visited the relevant landing page; (iii) the duration of the task (until they decided that they were complete); (iv) how successful they believed they were in finding the answer; and (v) the actual correctness of their answer. The correctness of the participants' answers was determined by an experimenter who reviewed the responses provided.

Overall 66.5% of tasks were answered correctly, and participants believed that 72.0% of answers were correct. Table 5 shows the task completion metrics for each system ($\alpha = 0.008$). McNemar's chi-squared and exact tests are used for the binary variables (last three rows in Table 5). Participants were more likely to visit a page with an answer on the clickable snippets systems and *ThumbnailPreview* than with other systems ($\chi^2(5) = 16.23$, $p < 0.001$), and do so in less time with the clickable snippet systems ($F(5,388) = 4.07$, $p = 0.003$; Tukey post-hoc tests: $p < 0.01$). *ThumbnailPreview* led users to a similar percentage of pages with the answer as were visited in the two clickable snippets systems (i.e., 71.9% vs. 74.0%) (all $p \geq 0.34$). However, inspecting the thumbnail previews took longer, so the total time until an answer page was visited was significantly longer ($F(5,570) = 3.85$, $p = 0.002$; Tukey tests: $p < 0.01$).

Overall, participants finished their tasks in less time, reported believing that they were more successful, and were actually

**Table 6. Average relative rankings of the systems and the number of participants ranking the system on top. *N*=48.**

| System | Rel. Rank | # participants who assign rank to system | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| ClickableSnippets Gradual | 2.85 | 16 | 8 | 6 | 7 | 7 | 4 |
| ClickableSnippets Immediate | 3.23 | 9 | 12 | 8 | 4 | 8 | 7 |
| SnippetHighlighting | 3.35 | 6 | 7 | 14 | 11 | 5 | 5 |
| TermHighlighting | 3.69 | 5 | 11 | 5 | 8 | 11 | 8 |
| ThumbnailPreview | 3.75 | 8 | 5 | 6 | 11 | 8 | 10 |
| Baseline | 4.13 | 4 | 5 | 9 | 7 | 9 | 14 |

more successful with clickable snippets systems than the others (all $p \leq 0.006$). There were no differences between the clickable snippets systems (all $p \geq 0.63$) suggesting that the transition had no effect on task outcomes.

Turning attention to the highlighting strategy on the landing page, we see that *TermHighlighting* resulted in longer tasks ($F(5,570) = 3.85$, $p = 0.002$; Tukey tests: $p < 0.01$), perhaps because it was distracting, as was suggested earlier. Tasks with *SnippetHighlighting* took longer than with the clickable snippets systems and participants were less successful. This suggests that the better performance of the clickable snippets was not only due to the highlighting, but also the preview and transition orientation strategies.

Observed differences in whether participants visited answer pages for each of the systems emphasizes the importance of the text on SERPs in directing users toward relevant pages. However, reaching an answer page does not *guarantee* the correct answer will be located on that page. To test this we computed the fraction of occasions where a user found the correct answer after reaching an answer page (see last row of Table 5). The values show that while reaching an answer page helps, there is a significant gain in answer correctness beyond helping people reach answer pages ($\chi^2(5) = 16.82$, $p = 0.005$; all $Z \geq 2.45$, all $p < 0.001$). This demonstrates the practical value of clickable snippets and suggests that the *transition support* may be primarily responsible for the observed gains in task performance from the clickable snippets systems. This also suggests many answers may have been near the snippet text in the landing pages.

### Overall Perceptions

At the end of the study, participants were asked to provide a final ranking of the six systems they had used in terms of their overall preferences. As noted earlier, participants ranked systems progressively, incrementally adding a new system to the overall ranking after they had used it (e.g., after using three systems, they ranked those three). Participants had to rank systems in descending order of preference; no facility was provided for ties. The final relative rankings (1 = best, 6 = worst) are summarized in Table 6.

The findings show that participants preferred the clickable snippet systems to the others ($\chi^2(5) = 13.38$, $p = 0.01$; all $Z \geq 2.10$, $p \leq 0.018$). Over half of the participants in the study (52.0%, *N*=25) preferred one of the clickable snippets and 66.7% (*N*=32) ranked one of the clickable systems in their top two. This shows a clear preference for this type of orientation support. Interestingly, the gradual clickable snippets system is preferred more than the immediate clickable snippets system, despite the transition delay ($Z = 2.23$, $p = 0.013$). Beyond clickable snippets, the findings show *SnippetHighlighting* was preferred over *TermHighlighting* or no highlighting systems (all $Z \geq 2.37$, $p < 0.01$). In addition, the post-hoc testing shows there was no significant difference between *TermHighlighting* and *ThumbnailPreview*, the two commonly used orientation systems, although both ranked above *Baseline* (all $Z \geq 2.45$, all $p < 0.01$), which only four participants favored.

In open feedback about clickable snippets, participants liked the transition (e.g., "the transition really led your eye on the landing page"), including its gradual nature. They also liked how it emphasized the connection between the SERP and the landing page, being shown where on the landing page the snippet appeared, and the guidance on where to look on the landing page. Some participants did not like the delay in the animation and preferred the immediate transition (e.g., "[immediate] is faster and less flashy so it is better"). However, more participants found directness of the transition disorienting (e.g., "it was a little irritating at first to be halfway down the page, because it was hard to determine context").

### Summary

The main findings from this study are that participants:

1. Frequently wanted to locate snippet text in the landing page (56% of participants reported doing this for over half of their searches) (motivation).
2. Reported often using skimming/scanning to locate information on landing pages (motivation).
3. Felt that previews gave them a better idea of what to expect but did not help them navigate to content of interest from the SERP (RQ1).
4. Preferred and were more effective with snippet underlining than other SERP orientation methods (RQ1).
5. Preferred and were more effective with snippet highlighting compared to term highlighting and no highlighting, and found term highlighting to be particularly obtrusive when used on landing pages (RQ2).
6. Preferred gradual transitions from snippet click to page over immediate transitions (with automatic scrolling) and basic transition (no scrolling) (RQ3).

## DISCUSSION AND CONCLUSIONS

Our findings show participants preferred and were more effective with the clickable snippets systems. Of the two clickable-snippet variants, almost twice as many participants preferred the system with the gradual transition, which made the relationship between the SERP and landing page more obvious and was less disorienting than the immediate transition, according to participants. However, the gradual transition requires additional time for the animation and more research is needed to fully understand the effect of transition time on preference and performance.

We manually identified relevant regions in the landing pages to help separate the evaluation of clickable snippets from its implementation. A deployed system, however, requires automatic snippet-content matching that can match snippets spanning multiple HTML elements, different document types, dynamic pages, and ill-formed HTML. Access to search engine crawling and indexing infrastructure may also assist in addressing some of these challenges.

Only 14.6% ($N$=7) of participants used find-in-page functionality on landing pages. While we did permit the use of find-in-page, we did not include it as a baseline because: (i) the functionality could be applied in all systems, and its introduction could affect behavior on systems that followed, and (ii) some of its functionality was already present in *TermHighlighting*. The data we gathered does not allow a direct comparison of clickable snippets and find-in-page functionality; a separate controlled study is needed.

Although participants liked clickable snippets they suggested improvements by combining multiple orientation methods (e.g., show the snippets in context when users hover, similar to [27]). More work is needed to understand the costs and benefits of this technique for landing pages where scrolling is or is not required, as well as navigational and transactional tasks. The methods described could also be expanded to blended and non-text search scenarios [22]. For example, for results represented on the SERP by video thumbnails, the search engine could offer single-click transitions to the exact frame depicted in the thumbnail, not just the start of the video as is current common practice. Methods such as clickable snippets may also work well in mobile settings, where limited screen real estate could make landing-page orientation more challenging than on the desktop.

## REFERENCES


1. Aula, A. et al. (2010). A comparison of visual and textual page previews in judging the helpfulness of web pages. *WWW*, 51–60.
2. Boguraev, B. et al. (1998). Dynamic presentation of document content for rapid on-line skimming. *AAAI Symp. Intel. Text Summarization*, 118–128.
3. Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, 36(2): 3–10.
4. Buchanan, G. and Loizides, F. (2007). Investigating document triage on paper and electronic media. *ECDL*, 416–427.
5. Caracciolo, C. and de Rijke, M. (2006). Generating and retrieving text segments for focused access to scientific documents. *ECIR*, 350–361.
6. Chi, E.H. et al. (2005). ScentHighlights: highlighting conceptually-related sentences during reading. *IUI*, 272–274.
7. Chi, E.H., Gumbrecht, M. and Hong, L. (2007). Visual foraging of highlighted text: an eye-tracking study. *HCI*, 589–598.
8. Clarke, C.L.A. et al. (2007). The influence of caption features on clickthrough patterns in web search. *SIGIR*, 135–142.
9. Dou, Z., Song, R. and Wen, J.R. (2007). A large-scale evaluation and analysis of personalized search strategies. *WWW*, 581–590.
10. Erkelens, C.J. and Vogels, I.M.L.C. (1995). The initial direction and landing position of saccades. In *Eye Movement Research: Mechanisms, Processes, and Applications*, 133–144. New York: Elsevier.
11. Google Quick Scroll. Downloaded from: https://chrome.google.com/webstore/detail/okanipcmceoeemlbjnmnbdibhgpbllgc. (Accessed September 2012).
12. Graham, J. (1999). The reader's helper: a personalized document reading environment. *SIGCHI*, 481–488.
13. Harper, D.J. et al. (2004). Within-document retrieval: a user-centered evaluation of relevance profiling. *Information Retrieval*, 7(3): 265–290.
14. Hearst, M.A. (1995). TileBars: visualization of term distribution information in full text information access. *SIGCHI*, 59–66.
15. Landauer, T.K. et al. (1993). Enhancing the usability of text through computer delivery and formative evaluation. *Hypertext: A Psychological Perspective,* 71–163.
16. Lindgaard, G. et al. (2006) Attention web designers: you have 50 milliseconds to make a good first impression! *Behavior and Information Technology*, 25(2): 115–126.
17. Loizides, F. and Buchanan, G.R. (2008). The myth of find: user behaviour and attitudes towards the basic search feature. *JCDL*, 48–51.
18. Nielsen, J. (1997). *How Users Read on the Web. Useit.com Alertbox*. (Accessed September 2012). http://www.useit.com/alertbox/9710a.html
19. Olsen, A. *The Tobii I-VT Fixation Filter: Algorithm Description*. (Accessed September 2012). http://www.tobii.com/Global/Analysis/Training/WhitePapers/Tobii_WhitePaper_TobiiIVTFixationFilter.pdf
20. Olston, C. and Chi, E.H. (2003). ScentTrails: integrating browsing and searching on the Web. *TOCHI*, 10(3): 177–197.
21. Paek, T., Dumais, S. and Logan, R. (2004). WaveLens: a new view onto internet search results. *SIGCHI*, 727–734.
22. Sushmita, S. et al. (2010). Factors affecting click-through behavior in aggregated search interfaces. *CIKM*, 519–528.
23. Teevan, J. et al. (2009) Visual snippets: summarizing web pages for search and revisitation. *SIGCHI*, 2023–2032.
24. Teevan, J., Dumais, S.T. and Liebling, D. (2008). To personalize or not to personalize: modeling queries with variation in user intent. *SIGIR*, 163–170.
25. Tombros, A. and Sanderson, M. (1998). Advantages of query biased summaries in information retrieval. *SIGIR*, 2–10.
26. Woodruff, A. et al. (2001). Using thumbnails to search the Web. *SIGCHI*, 198–205.
27. Zellweger, P.T. et al. (2000). The impact of fluid documents on reading and browsing: an observational study. *SIGCHI*, 249–256.