A Comparison of Visual and Textual Page Previews in Judging the Helpfulness of Web Pages

Anne Aula, Rehan M. Khan, Zhiwei Guan, Paul Fontes, Peter Hong
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
+1-650-253 0000

{anneaula, rkhan, zguan, pfontes, peterjinhong}@google.com

ABSTRACT

We investigated the efficacy of visual and textual web page previews in predicting the helpfulness of web pages related to a specific topic. We ran two studies in the usability lab and collected data through an online survey. Participants (total of 245) were asked to rate the expected helpfulness of a web page based on a preview (four different thumbnail variations: a textual web page summary, a thumbnail/title/URL combination, a title/URL combination). In the lab studies, the same participants also rated the helpfulness of the actual web pages themselves. In the online study, the web page ratings were collected from a separate group of participants. Our results show that thumbnails add information about the relevance of web pages that is not available in the textual summaries of web pages (title, snippet & URL). However, showing only thumbnails, with no textual information, results in poorer performance than showing only textual summaries. The prediction inaccuracy caused by textual vs. visual previews was different: textual previews tended to make users overestimate the helpfulness of web pages, whereas thumbnails made users underestimate the helpfulness of web pages in most cases. In our study, the best performance was obtained by combining sufficiently large thumbnails (at least 200x200 pixels) with page titles and URLs – and it was better to make users focus primarily on the thumbnail by placing the title and URL below the thumbnail. Our studies highlighted four key aspects that affect the performance of previews: the visual/textual mode of the previews, the zoom level and size of the thumbnail, as well as the positioning of key information elements.

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation**]: User interfaces – *Graphical user interfaces, user-centered design.* H.5.4 [**Hypertext/hypermedia**]: User issues

General Terms

Human Factors.

Keywords

Thumbnails, snippets, web page previews, user study.

1. INTRODUCTION

The usefulness of miniature versions of web pages - web page thumbnails - along with or instead of textual summaries in presenting search results has been a topic of research for a decade. Although studies have suggested that web page or document thumbnails give useful information about the page layout and visual design and that they might help users in relevance evaluation in search tasks (e.g., [4], [5], [7]), the issue is still controversial and thumbnails have not been widely adopted by the major search engines. One specific issue with thumbnails is that they might only be helpful for certain kinds of tasks [21]. In addition to using thumbnails to help users predict the relevance of a new web page, several studies have shown that thumbnails make it easy for users to recognize pages they have seen before (e.g., [3], [9], [6], [20], [22]).

Recently, tools, such as SimilarWeb [15], have been developed that show users related pages when they are browsing the web. In addition to the thumbnail preview, SimilarWeb also shows the page title and URL below the thumbnail. Another tool suggesting web pages to users is StumbleUpon [16]. Like SimilarWeb, StumbleUpon also presents the users with visual aids (either thumbnails or images from the web page) to help users find interesting pages (Figure 1).
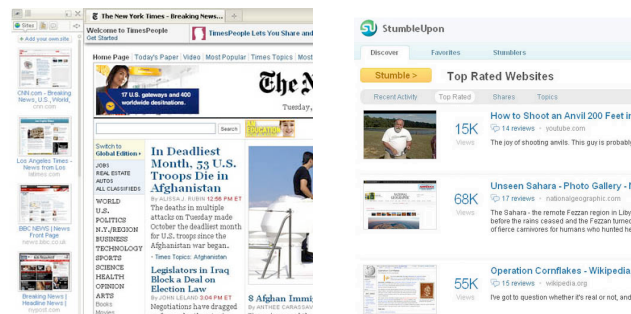

Figure 1: Tools recommending web pages: SimilarWeb (on the left) and StumbleUpon (on the right).

We ran a series of studies to see how thumbnails and text snippets perform in the context of recommending web pages. We report three studies that compared the helpfulness prediction accuracy between different kinds of thumbnails (different sizes and zoom

levels), text snippets, and only titles and URLs. In addition, we were interested in seeing whether the ordering of elements has an effect on the helpfulness evaluations, e.g., does it matter if the textual information is presented above or below the thumbnails? The key design hypothesis was that given a set of pages to browse to, additional information structures and rich visual patterns afforded by the page thumbnail enhance the user's ability to quickly decipher and predict whether a page would be useful to them or not.

Importantly, evaluating the appropriateness of thumbnails in search results was not our goal. Our goal was to understand how thumbnails compare to textual summaries and how these different summaries of the potential helpfulness of pages might interact

2. RELATED WORK

All the major search engines rely on textual snippets (more specifically, query-biased summaries [19]) to represent underlying web pages even though the visual design of the page might markedly influence its overall helpfulness. It is known that when evaluating web pages, users make their evaluations quickly: aesthetic evaluation of the visual appeal of the web page is made within the first 50ms [12]. Furthermore, a pleasing visual design is even known to draw the users' attention away from usability problems and thus, their experience with a visually pleasing page is likely to be better than with a similarly usable page with a less appealing visual design [11].

Thumbnails have often been proposed as an effective memory cue – if the user has seen a web page before, the assumption is that he or she will have an easy time recognizing the page based on it's thumbnail. Kaasten et al. [9] studied how different thumbnail sizes affect the recognition performance. Their study showed that if an 80% recognition rate is needed, the thumbnail size needs to be 160×160 pixels for website recognition and 208×208 pixels in order to recognize the exact page. Later, Won [20] suggested that 235×148-pixel thumbnail was required for accurate identification of previously seen web pages. Compared to text summaries, Teevan [18] et al.'s study suggested that thumbnails might make the identification of previously visited web sites faster, but for the thumbnails to support re-visitation, the participants must have seen the thumbnails, not just the web pages, before. Several researchers, including Jhaveri and Raiha [6], Kaasten and Greenberg [8], and Robertson et al. [13] have used thumbnails successfully in their experimental web page revisitation tools.

Although studies have shown that thumbnails are useful as memory cues when users have visited a web page before, the evidence supporting the usefulness of thumbnails as representations of unfamiliar web pages is less consistent. There are studies that show that even very small (33×34 pixels) thumbnails can provide important cues about the layout of text documents [4]. In the context of web searching, Dziadosz and Chandrasekar [5] studied the accuracy of relevance predictions between three different conditions: text snippets only, thumbnails (160×120 pixels) only, and a combination of text snippets and thumbnails. In their study, participants were asked to make a two-level judgment, relevant or not relevant based on these different preview styles. The authors did not do a statistical analysis of their results, but the results suggested that a combination of thumbnails and text snippets resulted in most correct decisions regarding the relevance of the web page.

Teevan et al. [18] studied the performance of thumbnails, "visual snippets" and text snippets in homepage finding, shopping and medical tasks. Their findings suggested that text snippets are effective for finding new web pages (pages the users have not seen before) and thumbnails work well for re-finding pages seen before, provided that the user has seen the thumbnail version of the page before. Visual snippets, their novel representation combining the essential information of text snippets and the visually distinct features from thumbnails worked well for both finding new information and re-finding previously seen pages. Woodruff et al. [21] showed that the performance of thumbnails was highly variable - in some tasks, they performed extremely well, but there were task categories that were clearly unsuitable for regular thumbnails.

Several researchers have proposed ways to enhance thumbnails in different ways instead of only minimizing the web page. Buscher et al. [2] ran an eye-tracking study which showed that when evaluating web pages, users typically focus on the upper left corner first followed by the center-left region (especially in information-driven tasks). Based on this finding, the authors suggested that by using these most salient regions, it might be possible to create more compact representations of web pages. Woodruff et al. [21] proposed enhancing thumbnails by highlighting keywords on a transparent overlay on the thumbnail; their study showed that the enhanced thumbnails performed more consistently than either regular thumbnails or text summaries (enhanced thumbnails were also used in [17]. When presenting web pages on mobile devices, enhancing thumbnails with readable text fragments was suggested by Lam and Baudisch [10]. Their technique was shown to be preferred by users over a one-column presentation of the web page or a regular thumbnail.

Although enhanced thumbnails have often performed well in laboratory studies and they have typically outperformed simple thumbnails, they can be computationally too expensive to be used in live systems. Thus, we chose to use simple minimized versions of web pages in our studies. In addition, we do believe that the thumbnail alone is not the best possible way to represent web pages and thus, we mostly focused on showing thumbnails with the page titles and URLs (page titles are known to work well in predicting the usefulness of a web page [14]). This presentation style provides a compact presentation that does not take too much space but that provides the users with essential information about the page that is not possible to communicate without combining the visual and textual information.

3. METHODS AND RESULTS
3.1 Study 1
In Study 1, we tested two thumbnail variations. One variation (Zoomout) showed more of the page layout and the other (Zoomin) showed more details by zooming into the top left corner of the web page (this is the area that users typically focus on first as shown by [2]).

3.1.1 Participants
Ten participants (all employees of Google) participated in this study. Participants received gift checks as compensation.

3.1.2 Stimuli
We tested four web page genres: medical, travel, shopping, and images. For each genre, we generated two specific topics, and generated four results per topic. To find web pages, we submitted the topics as queries to Google and selected the "Similar pages"

link for the first result. The four results for each topic were selected from the top similar pages results.

We compared two different thumbnail variations. Both variations are composed of a thumbnail, followed by the page title and URL. One thumbnail shows more of the page layout (20% zoom level, also called 'Zoomout' preview) and another one shows more details of the page (38% zoom level, also called 'Zoomin' preview). With the 38% zoom level, the first level headings of the web page were typically legible. The pixel dimensions used for the two thumbnail versions were 200×250 pixels for the 20% zoom level and 280×250 pixels for the 38% zoom level. Figure 2 shows examples of the two thumbnail variations.
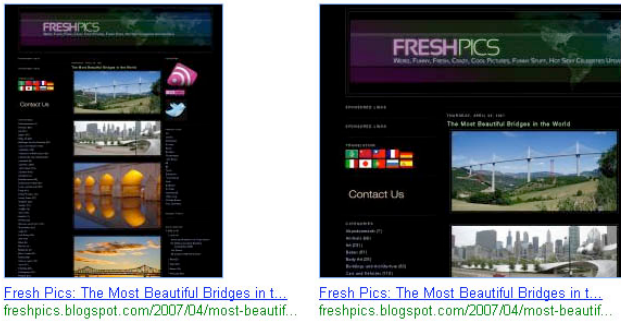

Figure 2: 20% (200x250 pixels - 'zoomout') and 38% (200x250 pixels - 'zoomin') zoom level thumbnails.

3.1.3 Procedure

The participants were given a topic and asked to evaluate the helpfulness of a list of web pages given the goal of learning more about the topic. The four topic categories were medical, travel, shopping, and images. The specific topics were:

- symptoms of high blood pressure
- symptoms of diabetes
- hiking in Virginia
- bike trails in Monterey
- shopping for a custom t-shirt
- shopping for board games
- sunset pictures
- pictures of bridges

Participants were asked to rate the helpfulness of each web page first based on the preview (either visual or text snippet) and then after visiting the actual page. The rating was done on a scale from 1 (not helpful at all) to 5 (extremely helpful) - the scale was presented as radio buttons below each thumbnail image. When rating the previews or web pages, the participants were asked to trust their initial impressions instead of carefully scrutinizing each preview or web page.

The participants first evaluated all 16 previews they saw using one preview variation. After rating all previews, they were presented with a list of URLs that took them to the actual web pages. After that, the procedure was repeated for the other preview variation.

We used a within-subjects design wherein each participant evaluated all four topic categories with both thumbnail variations; the specific topics were different for the different thumbnail variations (the order of seeing the thumbnail × topic combination was counterbalanced). Overall, the participants evaluated 16 thumbnails and web pages with each preview style (4 topic categories, 4 thumbnails each; with two preview styles, the total number of ratings was 32 for each participant).

3.1.4 Results

Both 'Zoomin' and 'Zoomout' previews were effective in predicting the helpfulness of webpages. Figure 3 shows mean actual (y-axis) and predicted (x-axis) helpfulness judgments for each task in both the 'Zoomin' and 'Zoomout' conditions ('Zoomin' $r = 0.85$, $p < 0.02$; 'Zoomout' $r = 0.82$, $p < 0.02$).
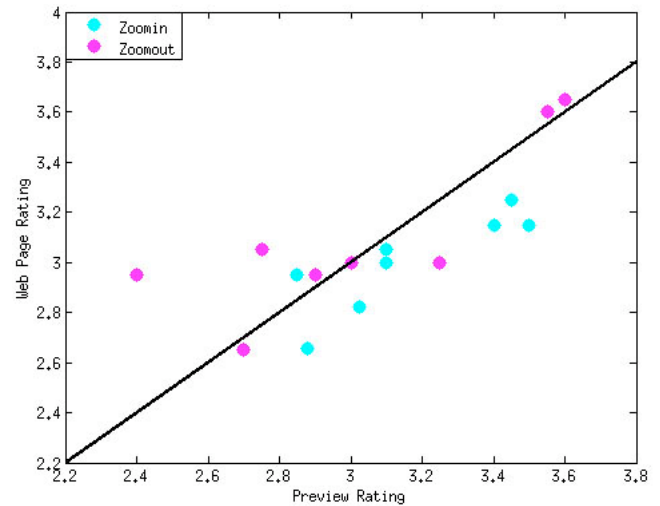

Figure 3: Ratings for the 'Zoomin' (38% zoom level) and 'Zoomout' (20% zoom level) thumbnails. The black line is the unit slope line showing equal ratings for thumbnails and actual web pages (dots are above the line when users underestimated the helpfulness of the page based on the preview and below the line when users overestimated the helpfulness).

Figure 3 shows that seven of the eight tasks for the 'Zoomin' condition fall below the unit slope line, indicating that participants tended to overestimate the helpfulness of web pages based on these thumbnails. In contrast, only three of the eight tasks in the 'Zoomout' condition were below this line suggesting that these thumbnails led to more accurate judgments. To test this observation statistically, we fit a regression model to the actual web page helpfulness judgments based on preview ratings, task type, and user thumbnail condition ('Zoomin' vs 'Zoomout') and controlling for user effects. The difference in thumbnail condition was statistically significant ($F(1,628) = 6.30$ $p < 0.0124$), suggesting that the 'Zoomout' thumbnail condition provided better information for helpfulness judgments.

3.2 Study 2

The results from Study 1 were positive showing that thumbnails can be used successfully to evaluate the helpfulness of web pages related to a certain topic. However, the first study left us wondering how thumbnails perform compared to text snippets. Thumbnails are computationally more expensive to use, so if text snippets resulted in similar prediction performance, they might be a better preview choice for the context of recommending related pages. In addition, we were also interested in seeing if the positioning of the title and URL had an effect on the performance of the thumbnails.

3.2.1 Participants

Twelve participants (all employees of Google) participated in this study. Participants received gift checks as compensation.

3.2.2 Stimuli

We used the same genres and specific topics as we did in Study 1. The procedure of generating the previews was also similar to that used in Study 1. The webpage genres and the actual web pages were reused from Study 1 to maintain consistency and to make it possible to compare the results between the studies.

In Study 2, we compared textual and visual webpage previews. We also wanted to see how the positioning of the title and URL would affect users' prediction accuracy. We used the 20% 'Zoomout' thumbnails from the previous study, but we moved the title and URL above the thumbnail. Textual snippets we used in the study were generic text snippets from Google (type "info:[URL]" on Google to get generic snippets). Figure 4 shows examples of textual and visual previews we used in this study.
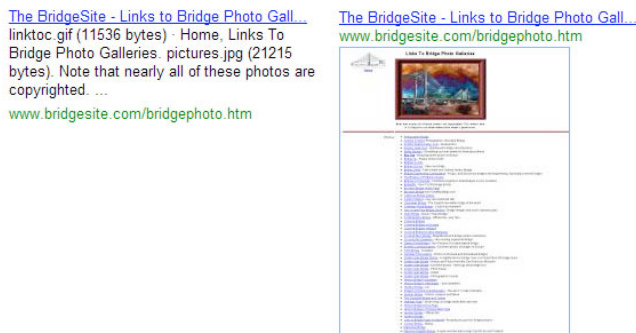

Figure 4: An example of a textual preview and a thumbnail used in Study 2. Note that in this study, the title and URL were above the thumbnail.

3.2.3 Procedure

The same procedure was used as in Study 1.

3.2.4 Results

Thumbnails and snippets were both effective cues for the judging the helpfulness of web pages. Figure 5 shows mean actual (y-axis) and predicted (x-axis) helpfulness judgments for each task in both the thumbnail and textual snippet conditions (thumbnail $r = 0.93$, $p < 0.001$; snippet $r = 0.8$, $p < 0.024$). In this experiment, 7 of the 8 textual snippet tasks fall below the unit slope line, indicating that participants tended to overestimate the helpfulness of web pages based on text-only snippets. In contrast, only 5 of the 8 tasks fell below the unit slope line for the thumbnails, although it appears that with thumbnails, participants tended to overestimate the helpfulness of the most helpful pages and underestimate the helpfulness of the least helpful pages. To test this observation statistically, we fit a regression model to the actual web page helpfulness judgments based on preview ratings, task type, and user preview condition ('thumbnail' vs. 'snippet') and controlling for user effects. The difference in preview condition was not statistically significant ($F(1,380) = 1.12$, $p > 0.1$).
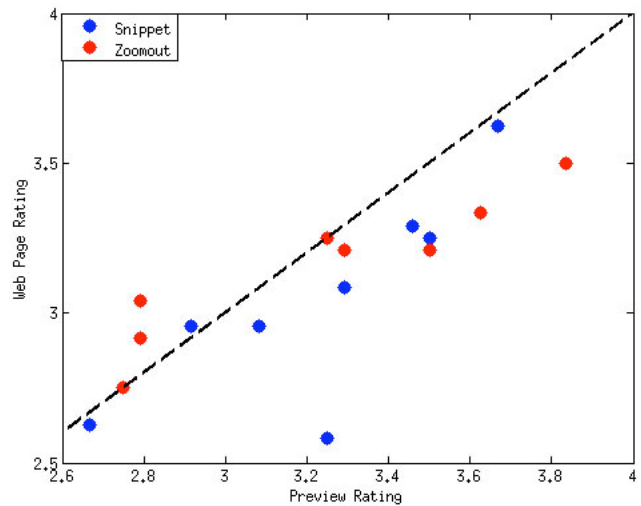

Figure 5: Ratings for the text snippets (blue dots) and thumbnails (red dots).

Somewhat unexpectedly, although the 'Zoomout' thumbnails we used in the first two studies were the same with only the positioning of the titles and URLs being different, we saw some differences in how accurately users could evaluate the helpfulness of the underlying web pages based on these previews. To understand why these thumbnail variations would perform differently, we had 22 users simply select the most helpful-looking preview from a list of four thumbnails for each of the topics we used in the first two studies. While they were doing that, we used an eye-tracker (Tobii 1750) to capture their gaze position. The heatmaps in Figure 6 combine the gaze data from all the tasks the users did with each preview style. The heatmap shows that when titles and URLs were presented below the thumbnail (left column), users spent a lot of time fixating on the thumbnail - but they also check the titles and URLs. When titles and URLs are prioritized by placing them above the thumbnail, users seem to rely mostly on them in their evaluation. As explained above, in the latter case, their prediction performance seems to suffer slightly by making them overestimate the helpfulness of the most helpful pages and underestimate the helpfulness of the least helpful pages. We did not see a similar systematic evaluation error with the other preview variation.
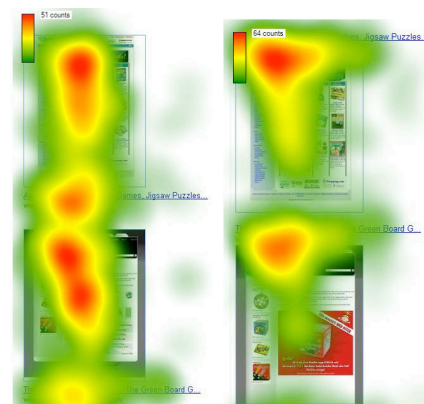

Figure 6: Heatmaps showing how users focused on different parts of the preview when titles and URLs were presented below the thumbnail (on the left) or above the thumbnail (on the right).

3.3 Study 3

The lab studies showed us that the zoomed out (20%) thumbnails resulted in more accurate predictions than the zoomed in (38%) thumbnails. The studies also showed that the positioning of the title and URL has an effect on how users process the visual previews. Thus, we wanted to collect more data with the title and URL below the thumbnail as that seemed to result in more accurate predictions and compare that thumbnail variation to text snippets. We also wanted to address the effects of making the thumbnails smaller, using thumbnails only with no textual information, and also seeing if titles and URLs alone might provide enough information.

In addition, we wanted to address the potential bias in the lab studies that used a within-subjects design: it is possible that the previews performed as well as they did because users wanted to be right – after rating the preview, it is possible they remembered the rating and simply wanted to rate the web page similarly. We do not think this was a serious issue as we saw many cases where participants changed their ratings. However, to make sure that this bias didn't affect the results, we chose to use a between-subjects design in the online study.

3.3.1 Participants

In the online study, we had 223 participants (all participants were employees of Google). As compensation, we raffled 20 gift checks or t-shirts among the participants.

3.3.2 Stimuli

In addition to the four task categories we used in the usability lab studies (medical, travel, shopping, and images), we added two genres: blogs and news for the online study. The specific topics we used for the blogs and news categories were:

- blogs about new cell phones
- funny blogs
- Bay area news
- top news stories

The process of finding the web pages for each topic was similar to what we used in the lab studies.

To collect a larger set of ratings, this study was run as an online experiment. In this study, we used a between-subjects design to alleviate the potential bias of participants being unwilling to change the rating they gave to the preview after seeing the actual web page. The goal of the online study was to focus on the effect of the size of the thumbnail (whether showing more of the page layout helps users in the helpfulness prediction), the titles and URLs, and just textual snippets. In the online study, all the thumbnails were presented at the 20% zoom level and the title and URL was always presented below the thumbnail (or not at all). In this study, we had seven different conditions (the six preview variations are shown in Figure 7):

1. 200×150 pixel thumbnail with the title and URL
2. 200×200 pixel thumbnail with the title and URL
3. 200×250 pixel thumbnail with the title and URL
4. 200×250 pixel thumbnail (without the title and URL)
5. titles and URLs only
6. snippets only
7. the actual web page

3.3.3 Procedure

The participants received an email sent to a company-wide email list inviting them to participate in a quick online study. The participants were explained that they will be given topics (for example, "pictures of puppies") and that their task was to rate web page previews on a 5-point scale in how helpful they think that web page would be for that topic. For those participants who rated web pages, we asked them to rate the web pages in how helpful they were for the topic. We also told participants that they shouldn't spend much time scrutinizing each preview (or web
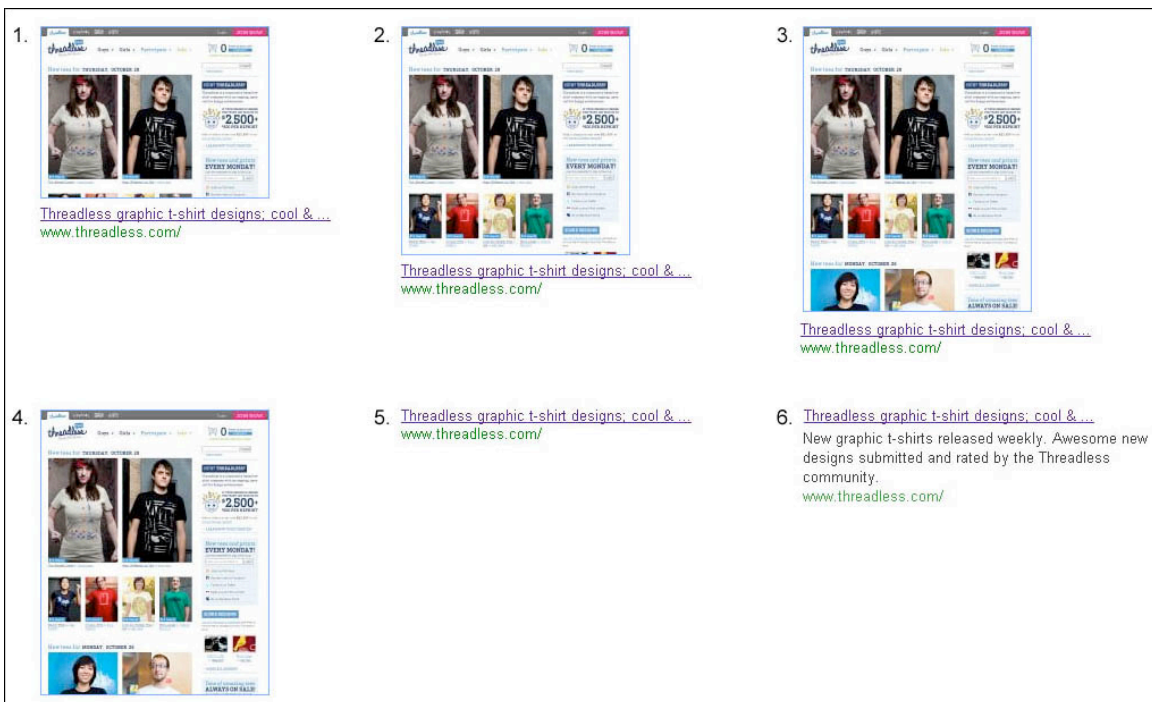

Figure 7: The six preview styles tested in Study 3 (see text for explanations).
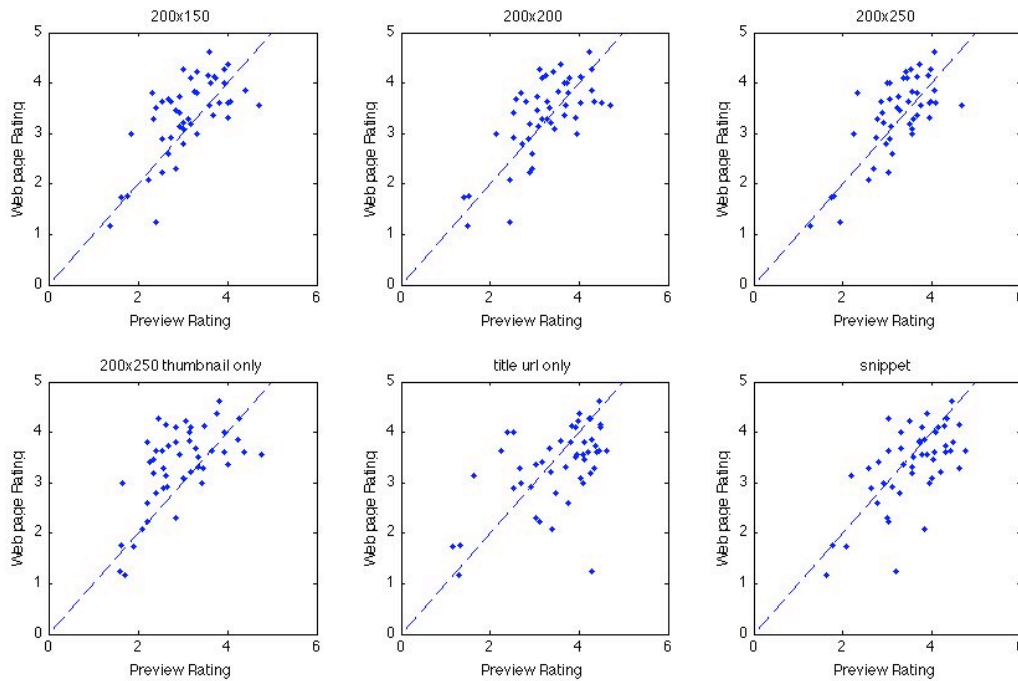
Figure 8: Preview vs. web page ratings in the different conditions.

page) or go to the actual web page (in the preview condition), but instead, they were asked to trust their initial impression based on the preview (or the web page).

Participants were randomly assigned to one of the seven conditions. Each participant rated 48 previews. Participants who rated the web pages (as opposed to the previews) rated only half (24) of the web pages (this was done because the process of rating the actual web pages took more time than rating the previews).

3.3.4 Results

Figure 8 shows the predicted (preview) and actual (web page) judged helpfulness ratings for each task in each of the six preview conditions. The top three panels show the three thumbnail sizes tested, from smallest to largest. Visually, it seems that the smallest thumbnails (200×150) produce the most underestimation of the helpfulness of web pages, and the largest (200×250) thumbnails the least. To test this hypothesis, we conducted paired t-tests between the values of preview judgments for the 200×150 and 200x200 conditions, and between the 200×200 and 200×250 conditions. The difference between the first pair was significant (t(47) = 3.12, p < 0.0031) but the not for the second pair (t(47) = -0.24, p > 0.1). The largest underestimation appears to occur for thumbnails without title and URLs (bottom left panel in the figure); this underestimation was significantly larger than that in the smallest thumbnail condition (t(47) = -2.36, p < 0.023). No effects of task type were observed in regressions of web page helpfulness that included tasks as predictors.

To better see how each of these preview types affects the judged helpfulness we computed a bias score for each preview type by subtracting the preview helpfulness judgments from the actual helpfulness judgments for each task and each preview type.

Average bias scores are shown in Figure 9 (error bars are one standard error of the mean, computed over tasks). A bias score of zero indicates perfect prediction of helpfulness, a positive score

suggests that participants underestimated the helpfulness and a negative score suggests that participants overestimate the helpfulness. As can be seen, thumbnails all produce underestimates of helpfulness, with larger thumbnails showing the least bias. Textual snippets produced overestimates on average. The largest deviation was for the thumbnails only condition, in which participants underestimated the helpfulness of pages the most.

The foregoing results suggest that visual thumbnails and textual snippets produce opposite biases in judging the helpfulness of web pages, and therefore the combination of both types of preview might provide the best user information. In order to understand this, we compared two nested regression models: a full model that predicted web page helpfulness with the preview helpfulness of the snippet and the thumbnail only conditions with a reduced model that included only the snippet preview. The full model was significantly better (F(1,45) = 11.53, p < 0.0014), suggesting that thumbnails contain information not accounted for by snippets alone. To test whether our method was not merely detecting an increase in power with the additional data, we tested a full model which included as predictors snippet, thumbnail 200×250 and thumbnail only and tested it against a reduced model containing only snippet and 200×250 thumbnail. In this case the full model was no better than the reduced model (F(1,45) = 0. 61, p > 0.1).

4. DISCUSSION

Study 1 showed that although both thumbnail versions lead to reasonably accurate helpfulness predictions, the zoomed in thumbnails tended to make users systematically overestimate the helpfulness of web pages. In practice, using this thumbnail style would mean that users would be likely to click on more recommended pages, but they would find the actual web page less helpful than they expected (false alarms). With the zoomed out
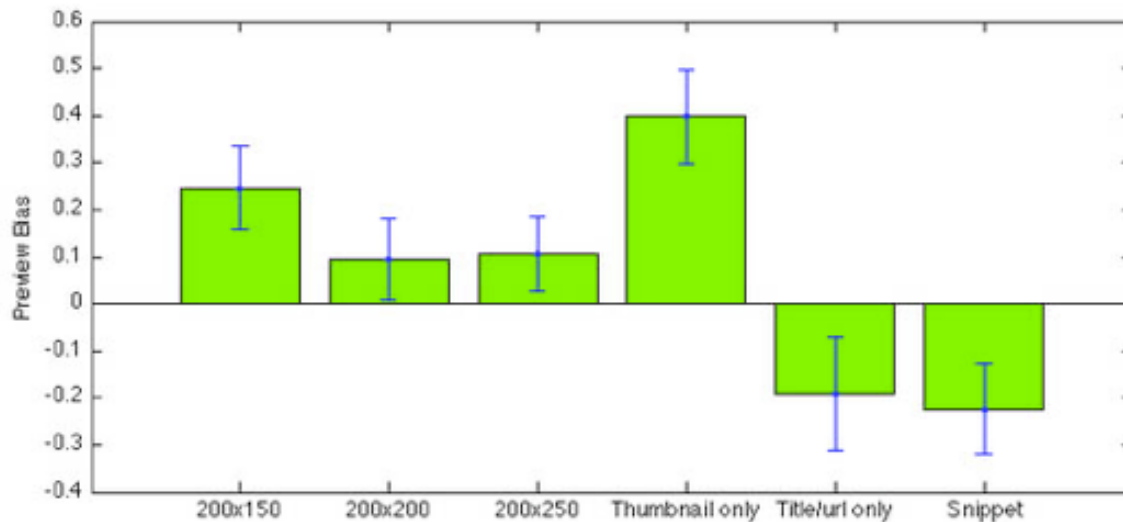
Figure 9: A bias score for each preview type (preview helpfulness rating subtracted from the actual helpfulness rating).

thumbnails, the predictions were more accurate and the thumbnails didn't lead to systematic errors in prediction. Although our data didn't provide definite answers to why we saw differences in the performance of zoomed in and zoomed out thumbnails, one explanation seems plausible in light of all the other data we collected. It might be that when using the zoomed in preview style, users ended up relying on the first level headings (that were legible in most cases) too much. They might have ended up *reading* the thumbnails and expecting the pages to be helpful if the heading seemed promising rather than using the thumbnails for visual information about the page structure and design. As we saw in our other studies, users tend to overestimate the helpfulness of pages when using textual previews. Further research (optimally with eye-tracking) is needed to confirm this hypothesis.

In Study 2, we showed that text only summaries lead participants to overestimate the helpfulness of the page - in the same way than zoomed in thumbnails. The thumbnails we used in this study (zoomed out thumbnails with title and URL above the thumbnail) resulted in more accurate predictions. However, our data showed that with these thumbnails, participants tended to overestimate the helpfulness of the most helpful pages and underestimate the helpfulness of the least helpful pages. In practice, these results mean that users would be more likely to click on recommendations if they were presented as textual snippets – however, they would likely be dissatisfied with the actual pages. With this thumbnail version, users would sometimes click on recommendations that would likely not be quite as helpful as they expected them to be (although they would still be helpful). Eye-tracking data helped us understand the reasons behind the difference in the prediction performance of the two zoomed out thumbnail variations we used in the first two studies. Eye-tracking heat maps showed that users relied heavily on the titles and URLs when they were placed above the thumbnails and spent very little time fixating on the thumbnails. When titles and URLs are placed below the thumbnail, users seem to focus on both pieces of information and that seems to result in more accurate helpfulness predictions.

Study 3 showed that purely visual (thumbnails) or textual previews (title/URL or full snippets) of the web pages result in marked differences in the accuracy of the helpfulness predictions. Thumbnails only made participants largely underestimate the helpfulness of the web pages whereas textual summaries made them overestimate the helpfulness. By adding the titles and URLs below the thumbnails and making sure that the thumbnails are large enough (200×200 or larger), the underestimation bias could be largely alleviated. Our data showed small (non significant) differences in performance between the title/URL condition and the full snippet condition. Although this result is somewhat surprising, Salmoni & Payne's [14] findings were similar.

One preview type we did not study was the combination of thumbnails and full text snippets. We would expect this preview style to result in most accurate predictions, but there is a high risk of adding too much information - if we make the users work too hard and process too much information, they are likely to decrease the workload by simply ignoring some pieces of information. Future studies should address the issue of optimal amount of information: whether previews with large-enough thumbnails, titles and URLs perform best or if providing even more textual information (snippets) would significantly improve the evaluation accuracy.

In our studies, we did not find significant differences in the performance of the preview styles related to the different types of task although we purposefully chose tasks we expected to be more suitable to one or the other (e.g., images and shopping tasks for thumbnails and medical and blog tasks for text summaries). It is possible (and likely) that some web page genres are more suitable for textual vs. visual presentation and the reason we did not see statistically significant differences between the different task types was due to the lack of experimental power. However, we believe that a combination of visual thumbnails and key textual information results in good performance with different kinds of web pages. In addition, as Cockburn et al. [4] showed, seeing the layout of the page is often helpful even for text-only documents.

Earlier research has established that a thumbnail size of approximately 200×200 pixels is needed for reliably recognizing web pages visited before. Surprisingly, the thumbnail size that is
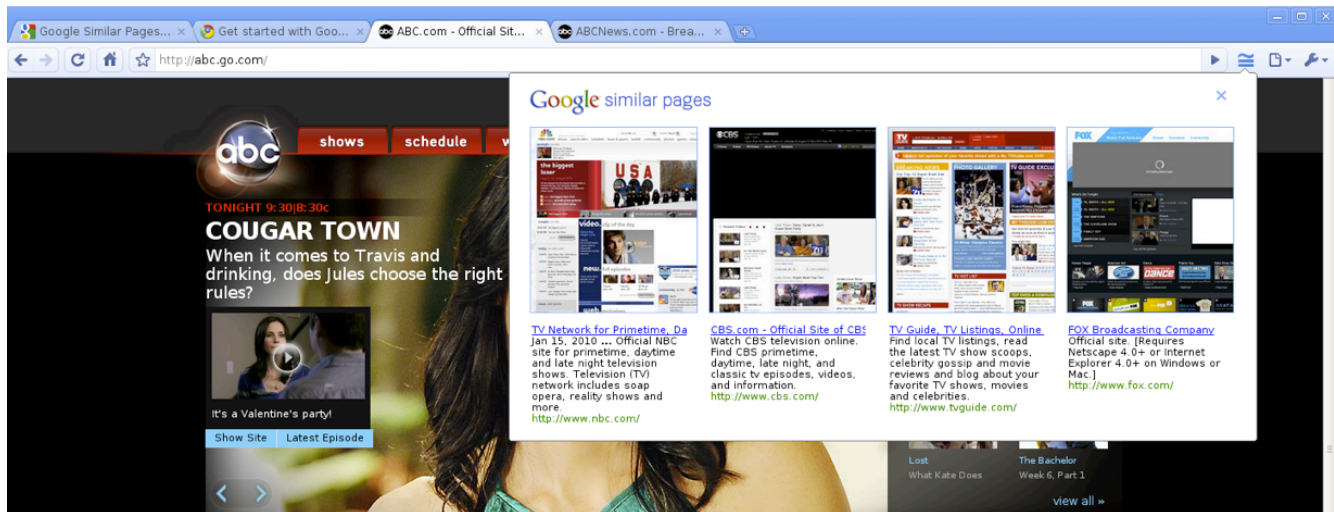
Figure 10: Google similar pages (an extension for the Chrome browser).

needed for effectively utilizing thumbnails when evaluating new web pages has not been systematically addressed by earlier research. Our studies showed that the recommendation by Kaasten et al. [9] seems to hold in this context, as well: thumbnails that were smaller than 200×200 pixels resulted in less accurate predictions, while there was no significant improvement if the thumbnails were made larger (200×250 pixels).

Although we ran these studies in the context of recommending related web pages for a broader topic, it is likely that some of the results would hold for more focused search tasks, as well. In these situations, the text snippets might perform better than they did in this study because it would be possible to use query-biased summaries instead of the generic summaries. Further research is needed to establish guidelines for how to most effectively utilize thumbnails in the context of search engine use.

5. CONCLUSIONS

The three studies we ran focused on the accuracy of users' evaluation of web page helpfulness based on different kinds of page previews – thumbnails of different sizes and zoom levels (showing either details of the page or more of the page layout) and text summaries, thumbnails combined with key textual information, and the positioning of the key information elements. The studies yield us to conclude the following:

- Given the same size thumbnails, the zoomed out (20% zoom level) thumbnails of the entire page, seem to result in more accurate helpfulness predictions than zoomed in (38% zoom level) thumbnails (predictive accuracy improves when showing the entire page over showing details).
- To increase the likelihood that users will pay attention to the thumbnails and enjoy their benefits, it seems better to place the textual information below the thumbnail.
- Combining textual and visual information results in the most accurate relevance predictions: thumbnails alone make users significantly underestimate the helpfulness of the page and textual snippets alone lead to overestimates of their helpfulness.

- A thumbnail size of ~200×200 pixels seems to be sufficient by providing enough information about the page layout – 200×150 pixels results in less accurate predictions.

Based on this and other research, in December 2009, Google released a thumbnail preview extension for the Chrome browser named 'Google Similar Pages' (Figure 10).

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] 50 best websites 2008 – SearchMe. Article in Time; available at http://www.time.com/time/specials/2007/article/0,2880 4,1809858_1809955_1811466,00.html (retrieved Nov 02, 2009).

[2] Buscher, G., Cutrell, E. & Ringel Morris, M. (2009) What do you see when you're surfing? Using eye tracking to predict salient regions of web pages. In Proc. CHI 2009, 21-30.

[3] Czervinski, M.P. can Dantzich, M., Robertson, G., and Hoffman, H. (1999) The contribution of thumbnail image, mouse-over text and spatial location memory to web page retrieval in 3D. Im Proc. Human-Computer Interaction - INTERACT '99, 163 - 170.

[4] Cockburn, A., Gutwin, C. & Alexander, J. (2006) Faster document navigation with space-filling thumbnails. In Proc. CHI 2006.

[5] Dziadosz, S. & Chandrasekar, R. (2002) Do thumbnails help users make better relevance decisions about web search results? In Proc. SIGIR'02, 365-366.

[6] Jhaveri, N. & Räihä, K.-J. (2005) The advantages of cross-session web workspace. In Proc. CHI 2005, 1949-1952.

[7] Joho, H. and Jose, J. (2005) Effectiveness of additional representations for the search results presentation on the web. Information Processing and Management, 44(1), 226-241.

[8] Kaasten, S. & Greenberg, S. (2001) Integrating back, history and bookmarks in web browsers. In Proc. CHI 2001, 379-380.

[9] Kaasten, S., Greenberg, S. & Edwards, C. (2002) How people recognize previously seen web pages from titles, URLs and thumbnails. In X. Faulkner, J. Finlay, F. Detienne (Eds.) People and computers XVI (Proc. Human Computer Interaction 2002), 247-265.

[10] Lam, H. & Baudisch, P. (2005) Summary thumbnails: readable overviews for small screen web browsers. In Proc. CHI 2005, 681-690.

[11] Lindgaard, G. & Dudek, C. (2002) User satisfaction, aesthetics and usability: Beyond reductionism. In: Usability gaining a competitive edge, Proc. IFIP 17th World Computer Congress, Montreal, Canada.

[12] Lindgaard, G., Fernandes, G., Dudek, C. & Brown, J. (2006) Attention web designers: you have 50 milliseconds to make a good first impression! Behavior & Information Technology, 25(2), 115-126.

[13] Robertson, G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D. & van Dantzich, M. (1998) Data mountain: using spatial memory for document management. In Proc. UIST'98, 153-162,

[14] Salmoni, A.J. & Payne, S.J. (2002) Inferences from search engine summaries and headings. In Proc. British HCI 2002, 10-13.

[15] SimilarWeb. http://www.similarweb.com

[16] StumbleUpon. http://www.stumbleupon.com

[17] Suh, B., Woodruff, A., Rosenholtz, R. & Glass, A. (2002) Popout Prism: adding perceptual principles to overview+detail document interfaces. In Proc. CHI 2002, 251-258.

[18] Teevan, J., Cutrell, E., Fisher, D., Drucker, S.M., Ramos, G., Andre (accent!!), P. & Hu, C. (2009) Visual snippets: summarizing web pages for search and revisitation. In Proc. CHI 2009, xx-yy.

[19] Tombros, A. and Sanderson, M. (1998) Advantages of query-biased summaries in information retrieval. In Proc. SIGIR 1998, 2-10.

[20] Won, S.S., Jin, J. & Hong, J.I. (2009) Contextual web history: using visual and contextual cues to improve web browser history. In Proc. CHI'2009, 1457-1466.

[21] Woodruff, A., Faulring, A., Rosenholtz, R., Morrison, J. & Pirolli, P. (2001) Using thumbnails to search the web. In Proc. CHI'2001, 198-205.

[22] Yoong-En, T. & Seong, D.S.K. (2007) Visual history mechanism on mobile Internet browser. In Proc. MC'07 (Mobility'07), 372-379.