

Kickmeier-Rust, M. D., & Albert, D. (2005). Highlighting – Or why Google is that successful. In A. Holzinger & K. H. Weidmann (Eds.), Empowering software quality: How can usability engineering reach these goals? (pp. 45-52). Wien: Österreichische Computer Gesellschaft.

Highlighting – Or Why Google Is That Successful

Michael D. Kickmeier-Rust & Dietrich Albert

Department of Psychology
University of Graz, Austria
michael.kickmeier@uni-graz.at
dietrich.albert@uni-graz.at

Abstract:

In today's information society, searching for information or for functions of a software application is one of the most frequent tasks. Still, visual search tasks are reported to be associated with low efficiency and user satisfaction. A simple method to improve search performance is highlighting. With the current experiment we investigate if search performance and user satisfaction can be significantly improved by presenting an optimal density of highlighted items. As the results of the eye-tracking experiment yielded, presenting 15 percent of highlighted items on a display improved search speed, accuracy, as well as user satisfaction in contrast to presenting no highlighted items or 49 percent. The results are discussed regarding their practical impact, design implications, and regarding emerging questions for future research.

Keywords: Highlighting, Search Performance, Visual Search, Visual Attention, Eye-tracking

1. Introduction

The basics of human visual perception and attention are a major foundation of software usability. For instance in information design, legibility, brightness, contrast, color, saliency, or the structure of information are crucial factors to enable successful and efficient information retrieval.

Searching (e.g., a piece of text on a website or a certain menu entry of a software application) is probably the most common tasks of human-computer interaction. Efficiency, accuracy, and user satisfaction of that interaction are strongly influenced by the visual properties and the amount of visual stimuli of a display. We are facing a vast amount of electronically available information, software applications providing countless functions, and searching is reported to be oftentimes difficult, inefficient, and unsatisfying [2, 7]. The infinite number of stored bits and the variety of software functions, however, are only valuable, if they can be searched, found, and used by humans.

The question arises whether information or functionality can be displayed in a way that performance, accuracy, and user satisfaction of visual search processes are possibly high; even without knowing what someone is searching for.

A well-tryed and successful method to facilitate visual search is highlighting; on paper as well as on electronic displays. There is a variety of research concerning the basic features of highlighting [14] and applied techniques of highlighting [8].

Highlighting is based on early visual processes. In a pre-attentive stage of vision the visual field can be processed as a whole without limitations of capacity. For example, if one searches a red circle among green circles, it can be detected immediately no matter how many green circles surround the red target. In terms of Ulric Neisser [6], the red circle *pops out* of the display. Generally, such pop-out effect occurs if the target is defined by single basic feature such as color, shape, size, or motion [14]. According to the *Guided Search* model [13, 15, 16], these early visual processes create an *activation map*. An activation map is a representation of the visual field determined by continuous levels of activation. The activation for a certain location in the visual field is determined by the visual properties of the item(s) in that location and intentional control. For example, when someone is searching for a target, which is expected to be red, a single red circle among blue ones provokes higher activation for its location than a single green circle. In a later stage of vision, attention is directed towards the peaks of the activation map. This guidance is determined by bottom-up aspects (e.g., saliency) and top-down aspects (e.g., a searchers intentions or expectations). Attention acts like a filter for relevant information to be passed to one or more limited-capacity processes of vision; only locations that are promising for a certain goal are attentively processed.

So far, visual attention refers to focal fixations, e.g. fixating a letter. Information processing, however, does not only occur for focally attended stimuli but also for the periphery of a stimulus. While the fovea is amounting for 2 degree visual angle, which corresponds to about 2 per mille of the visual field, peripheral vision allows the processing of information from the whole visual field. Due to the decrease of neurons related to retinal areas with increase eccentricity, however, visual performance generally decreases with eccentricity [13]. Still, there is evidence for the importance of peripheral vision in search tasks. Besides pre-attentive processing of the whole visual field, peripheral vision contributes to feature detection [1] or letter recognition [12] during search episodes.

Past research on highlighting, primarily focused on valid highlighting. Fisher and colleagues [3, 4] argued that a key variable of facilitating visual search through highlighting is the frequency a target, as opposed to a distractor, is highlighted. The aim of the current experiment is to investigate the effects of the density of highlighted items in a display.

On the basis of mechanisms of visual search, as introduced very briefly, we assume that presenting an *optimal* density of highlighted items increases visual search performance. The distribution of highlighted items facilitates the creation of an activation map in a pre-attentive stage of vision and, subsequently, guides the gaze through the display. Due to mechanisms like the *inhibition of return* [10] we can assume that these are attended only once along a serial path. A target can be detected either if itself is highlighted or, by means of peripheral information processing, if it surrounds a fixated (highlighted) item in a certain distance. This guidance is important because the effort a searcher spends on surveying a display is limited by means of a cost-benefit analysis, as outlined in the *Information foraging Theory* [9]. Not all objects within an area to search are attended during a search episode. When people are searching information they roughly scan through the visual field. However, if pre-attentive guidance corresponds to these limitations, the full search capacity can be utilized for target detection.

With the current eye-tracking experiment we investigate the impact of the total density of highlighted items in a display on search performance, accuracy, and user satisfaction.

2. An Eye-Tracking Experiment

2.1. Experimental Design

In an eye-tracking experiment participants were presented 48 panels consisting of 140 simple nouns each. The task was to determine as fast and as accurate as possible, whether a given search word was present among the nouns of the panel or not. In 26 tasks the target was present and in 22 tasks the target was absent. On the basis of a previous study on density effects of highlighting [5] which reported highest search performance between 15 and 30 percent highlighting, three different densities were realized: no highlighted items (0 percent) as a too small portion of highlighting, 15 percent, which was supposed to be an optimal density, and 49 percent as too high portion of highlighting. We realized 49 percent in order to avoid an inflection point at about 50 percent when the number of highlighted items starts to exceed the number of non-highlighted items and non-highlighted items start becoming more salient than highlighted ones. The sequence of tasks and the distributions of words and highlighting were created randomly for each participant. As dependent variables response times, answer accuracy (correct and incorrect responses), and eye movements were recorded.

2.2. Participants

The data of 36 participants were analyzed, thereof 27 females (75 percent) and 9 males (25 percent). The average age was 22.53 years ($SD = 3.14$), the youngest participant was 19 years of age, the oldest 31. In total, 21 participants reported normal vision, 15 reported vision corrected to normal. One participant of the “corrected to normal” group additionally reported a minor color vision defect.

2.3. Material and Apparatus

We used the SMI Eyelink I eye-tracker (SensoMotoric Instruments GmbH, Germany; www.smi.de) to record binocular eye-movements. The eye-tracking experiment was developed using Microsoft Visual C++ 6.0 and the SMI Eyelink API templates. The test software presented introduction and briefing screens and, subsequently, generated automatically random search panels from a pool of simple nouns. To counter-balance position effects, the different search tasks were automatically presented in random order for each participant. The experimental sequence was presented on a 21” computer monitor with a screen resolution of 1280x1024. At a distance of 58 cm from the screen a chin rest was mounted to stabilize head movements and the distance between eyes and screen. Responses were given using the Y key (yes, the target is present) and the N key (no, the target is absent) on a regular computer keyboard with German layout. The keys were prepared with stickers that allowed haptic detection in order to prevent participants from looking down to the keyboard during the experimental tasks.

2.4. Procedure

Participants were first asked to conveniently take place in front of the eye-tracker’s subject monitor. Then the participants were briefed regarding the aims of the experiment and the eye-tracking procedure. In a next step a paper-pencil data form was filled out, asking for age, sex, education, vision (normal, corrected to normal, vision below 100 percent, and vision defects), and Internet and search engine usage. Then participants were instructed to determine as fast and as

correct as possible whether a given word was present in the display or not. Additionally, the participants were briefed to inform the experimenter about any events during search tasks, e.g. forgetting the target word or usage of the wrong response key. After initial briefings the eye-tracker's head-mounted camera system was applied carefully and eye cameras were adjusted. After dousing lights, eye cameras were focused on the pupils and cameras were calibrated and validated. The camera adjustment, calibrations, and validations required about 10 to 15 minutes.

The experiment started with three training tasks. After successful completion of the training, participants performed the first block of 24 search tasks. Each task was initiated with the fixation of a center point on the screen. During this fixation the participants were told the target word. The participants started the search task by pressing the space key. To terminate a search task participants pressed either the Y key (yes, the target is present) or the N key (no, the target is absent).

After the first block of 24 search tasks a ten minutes break followed. During this break the uncomfortable head-mounted camera system was removed. The second block of search tasks was performed identical to the first one. Finally, the camera system was removed and participants rated three questions regarding their satisfaction with the different densities of highlighting. Arising comments of participants were note by the experimenter.

3. Results

3.1. Response Times

As hypothesized, we found significantly shorter response times when presenting 15 percent highlighted items in the search panel, in contrast to no highlighted items or 49 percent. In target-present tasks mean response time for the 15-percent group was 12.62 seconds ($SD = 7.25$), in the 0-percent group 14.85 seconds ($SD = 5.75$), and in the 49-percent group 15.69 seconds ($SD = 7.35$). In target-absent tasks mean response time for the 15-percent group was 23.6 seconds ($SD = 6.33$), in the 0-percent group 26.01 seconds ($SD = 7.84$), and in the 49-percent group 23.23 seconds ($SD = 6.70$). As shown in Figure 1a, in total the 15-percent group yielded a mean response time of 18.11 seconds ($SD = 6.79$), the 49-percent group 19.46 seconds ($SD = 7.03$), and the 0-percent group 20.43 seconds ($SD = 6.80$). An analysis of variance (ANOVA) reported a significant effect of the density of highlighted items ($F(2, 1132) = 2.28, p = .013$).

3.2. Accuracy

Focusing on target-present tasks only, also for answer accuracy (relative frequencies of correct responses) we found expected results. Answer accuracy was significantly higher when presenting 15 percent highlighted items in the search panel, in contrast to no highlighted items or 49 percent. As shown in Figure 1b, presenting 15 percent of highlighted items resulted in a relative frequency of correct responses of .79; for the 0-percent group we found a relative frequency of .67 and in the in the 49-percent group a relative frequency of .65. An ANOVA reported a significant difference between the three highlighting densities ($F(2, 1132) = 43.71, p < .001$).

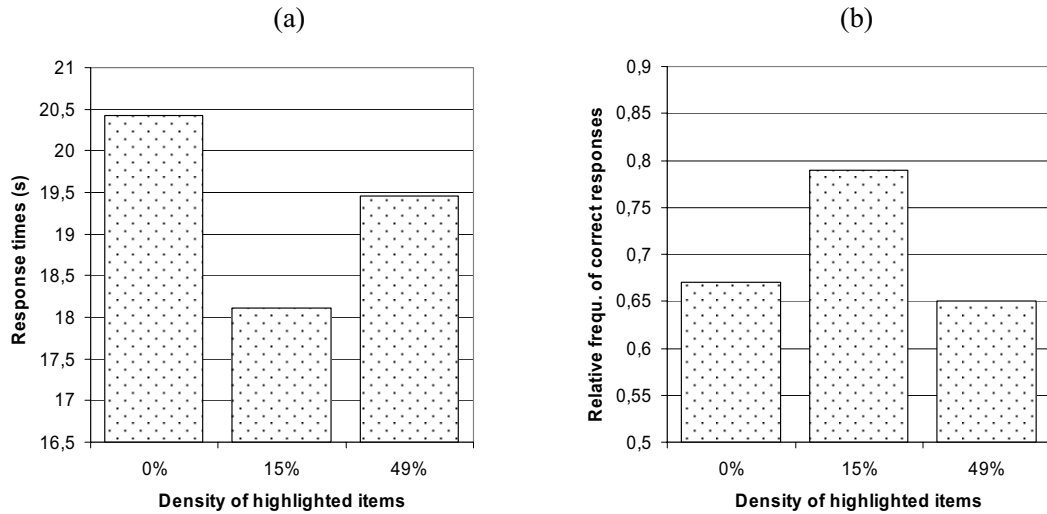


Figure 1. Panel (a) shows the response times for different densities of highlighted items in the search panel; panel (b) the relative frequencies of correct responses.

3.3. Scan Paths

Analysis of eye movements yielded that the gaze was immediately shifted from the initial center fixation to the item in top left corner of the search panel. This distinct saccade appeared in about 90 percent of search trials. Subsequently, scan paths were performed in reading-like direction from the top left corner row-wise to the bottom right corner. As shown in Figure 2a the relative frequency of saccades in reading-like direction (.62) was higher than the relative frequency of a reverse directed saccade (.38). An ANOVA reported a significant difference between these relative frequencies ($F(1,2310) = 4702.65, p < .001$). Eye-tracking data, moreover, reported that fixations on highlighted items were significantly more frequent than fixations on non-highlighted items. To assess the fixation ratio, we applied a coefficient Φ to obtain a meaningful value of the portion of fixations on an item type in relation to the total number of fixations and the portion of this type in the search panel. Φ combines the values of the number of total fixations ϕ_T , the number of fixations on one item type ϕ_I , the total number of items in the display I_T , and the number of items of a type I_I as

$$\Phi = \frac{\phi_I}{\phi_T} \cdot \left(1 - \frac{I_I}{I_T} \right).$$

Φ results in values between 0 and 1, whereat 0 means a low Φ -frequency of fixations and 1 a high frequency. The coefficient is 0 in two conditions: (a) If no fixations occur on an item type or and/or (b) if the number of items of a type equals the number of total items.

We applied this coefficient on highlighted words (Φ_H) and on standard, non-highlighted words (Φ_S). As shown in Figure 2b, the Φ -frequencies were higher at highlighted words ($\Phi = .19$) than at non-highlighted words ($\Phi = .14$). An ANOVA reported a significant difference between the Φ -frequencies of highlighted objects and standard, non-highlighted words ($F(1,2310) = 144.42, p < .001$).

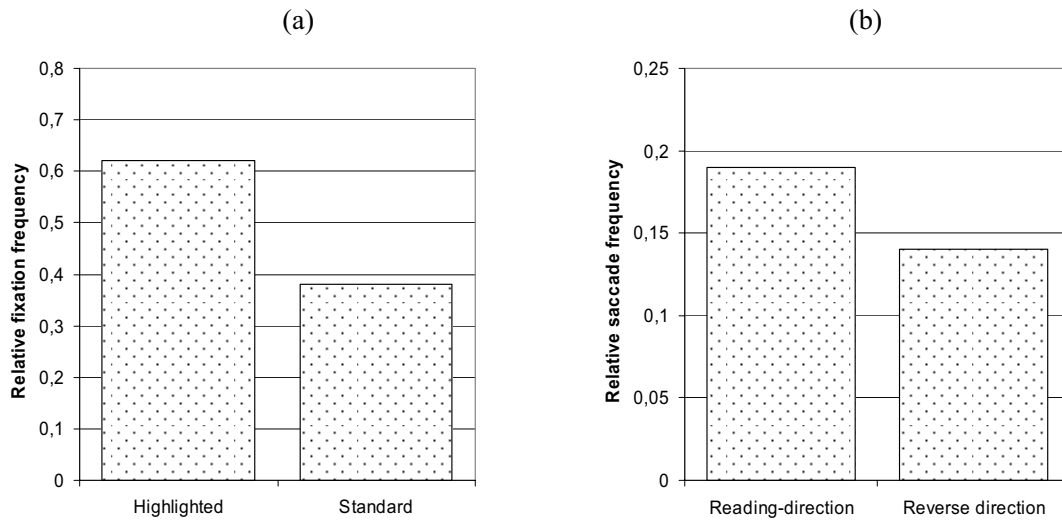


Figure 2. Panel (a) shows the relative frequencies of fixations highlighted items in contrast to non-highlighted items. Panel (b) shows the relative Φ -frequencies of saccades in reading-like and reverse directions.

3.4. User Satisfaction

In terms of software usability, user satisfaction is a crucial factor. To investigate not only performance but also user satisfaction, we analyzed the participant's rating of three questions at the end of the test session. As summarized in Table 1, presenting 15 percent of highlighted items in the search panel resulted in higher user satisfaction throughout. Summarized, the 15-percent group yielded a rating of 4.00 (SD = 1.35), the 0-percent group a rating of 3.62 (SD = 1.10), and the 49-percent group a rating of 2.75 (SD = 1.13). The differences between the groups was significant ($X^2(10, N = 315) = 70.08, p < .01$).

	Highlighting Density		
	0%	15%	49%
In your opinion, how good could you determine whether the target was present or not?	4.03	4.48	2.63
How convenient was the presentation of the search panels?	3.74	4.00	2.60
How easy was it to determine whether the target was present or not?	3.09	3.54	3.03

Table 1. This table shows the average rating on a six-step rating scale; 1 corresponds to a negative rating, 6 is the best rating.

4. Discussion

It's an important aim to increase search speed, accuracy, and user satisfaction of search tasks. With the current experiment we could demonstrate that this can be achieved by presenting an optimal density of highlighting. It's a remarkable finding, since when measuring speed and

accuracy one would expect a speed-accuracy trade-off. As the results show, on average no such trade-off was found. Longer response times did not result in higher accuracy. The reason for this effect might be the limited effort a searcher spends on surveying a display. Such limitations are described in the Information Foraging Theory [9] which refers to cost-benefit analyses of information's value and the expected efforts to reach this information. Consequently, searching is mostly a hasty and rough scanning of contents accompanied by fast decisions on the relevance of contents. For the current experiment this means that participants clearly limited their efforts; although participants were asked to search fast and accurate, none fixated all or at least almost all word in the search panel. On average 65 fixations were applied to scan through the 140 words within about 19 seconds. Within these limitations, a speed-accuracy trade-off does not significantly impact on search performance.

In summary we found clear indications for the hypothesized mechanisms of visual search. Highlighting facilitates - independent from its validity - the pre-attentive creation of activation maps. The peaks of activation are fixated serially in an attentive stage of vision. As the results show, fixations on highlighted items were significantly more frequent than fixations on non-highlighted items. The deployment of fixations generally occurred in reading-like direction from the top left row-wise to the bottom right. These results might be valid only for western culture. It would be an interesting question for future research to compare search paths of different cultures.

Eye movements along the described search paths enable target detection either if the target itself is highlighted or, by means of peripheral vision, if the target surrounds a fixated item. Across target-present and -absent tasks in the 15-percent group response times were 1.35 seconds faster than in the 0-percent group and 2.32 seconds faster than in the 49-percent group. For target-present tasks, the 15-percent group yielded 2.23 faster response times than the 0-percent group and 3.07 seconds faster response times than the 49-percent group. Moreover, the 15-percent group reported 12 percent higher accuracy than the 0-percent group and 14 percent higher accuracy than the 49-percent group.

The question remains, whether 15 percent highlighting is actually the optimal density for the current search tasks. This value was arbitrary selected on the basis of a previous study, which reported a distinct plateau of search performance between 15 and 30 percent highlighting. Future research should address the question, which density is optimal for different kinds and different amounts of material.

These results, furthermore, have significant practical impact. Referring to basic mechanisms of human visual perception and attention, visual search speed as well as search accuracy can be significantly improved with very simple methods. A crucial factor is that not only performance can be increased but also a searchers satisfaction with a search episode and the convenience with the presentation of contents.

This leads back to the sensationalism of this article's title. It's an interesting fact, that *Google*, likely the most successful and popular Internet search engine, presents the search results in a way that corresponds quite well to an optimal density of highlighting as reported in the current and in previous [5] experiments. For example, the percentage of highlighting (bolding, blue and green colors) of the first 10 results for when searching for "usability" is about 23%. Of course, it's far over-the-top to assume that Google's design of search results, which by now is a very popular design, is the only reason for the success of the company. But the results of the current experiment suggest that highlighting contributed to that success to a certain extent.

5. References

1. Ben-Av, M. B., Sagi, D., & Braun, J., Visual attention and perceptual grouping. *Perception and Psychophysics*. 52 (1992).
2. Bullinger, H.-J., Ziegler, J., & Bauer, W., Intuitive Human-Computer Interaction - Toward a User-Friendly Information Society. *International Journal of Human-Computer Interaction*. 14 (2002).
3. Fisher, D. L., & Tan, K. C., Visual displays: The highlighting paradox. *Human Factors*. 31 (1989).
4. Fisher, D. L., Courey, B. G., Tengs, T. O., & Duffy, S. A., Minimizing the time to search visual displays: The role of highlighting. *Human Factors*. 31 (1989).
5. Kickmeier, M. D., & Albert, D., The effects of scanability on information search: An online experiment, in: Ph. Gray, H. Johnson, & E. O'Neill (Eds.), *Proceedings of HCI 2003: Designing for Society*, Volume 2, 2003.
6. Neisser, U., *Cognitive Psychology*, New York: Appleton, Century, Crofts 1967.
7. Padovani, S., & Lansdale, M., Balancing search and retrieval in hypertext: context-specific trade-offs in navigational tool use. *International Journal of Human-Computer Studies*. 58 (2003).
8. Philipsen, G., Effects of six different highlighting modes on visual search performance in menu options. *International Journal of Human-Computer Interaction*. 6 (1994).
9. Pirolli, P., & Card, S., Information foraging. *Psychological Review*. 106 (1999).
10. Posner, M. I., & Cohen, Y., Components of attention, in: H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X*, Hillsdale: Erlbaum 1984.
11. Strasburger, H., Indirektes Sehen. Formerkennung im zentralen und peripheren Gesichtsfeld, in: A. Albert, M. Oswald, K. Pawlik, & Stapf, K.-H. (Eds.), *Lehr- und Forschungstexte Psychologie*, Göttingen: Hogrefe 2003.
12. Wang, Q., Cavanagh, P., & Green M., Familiarity and pop-out in visual search. *Perception and Psychophysics*. 56 (1994).
13. Wolfe, J. M., Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*. 1 (1994).
14. Wolfe, J. M., Visual Attention, in: K. De Valois (Ed.), *Seeing*, 2nd ed., San Diego: Academic Press 2000.
15. Wolfe, J. M., Guided Search 4.0: A guided search model that does not require memory for rejected distracters [Abstract]. *Journal of Vision*. 1 (2001).
16. Wolfe, J. M., & Gancarz, G., Guided Search 3.0, in: V. Lakshminarayanan (Ed.), *Basic and Clinical Applications of Vision Science*, Dordrecht: Kluwer Academic 1996.