Image Search Versus Text Search Revisited: A Simple Experiment using a Kahoot Quiz

Signe Aanderaa Eide[1], Ana-Maria Poljac[1] and Frode Eika Sandnes[1,2][0000-0001-7781-748X]

[1] Oslo Metropolitan University, N-0130 Oslo, Norway
[2] Kristiania University College, N-0153 Oslo, Norway
s333781@oslomet.no, s333745@oslomet.no,frodes@oslomet.no

Abstract. Information search is a common task when interacting with computers. Many studies have investigated the characteristics that facilitate effective search, specifically the use of icons and/or text. This study reports a simple experiment involving Kahoot where the goal was to observe searching performance with images and text. The results show that image search was faster, less error prone and preferred by participants compared to text search.

Keywords: image search, text search, visual scanning

1 Introduction

Visual search is needed when operating graphical computer interfaces, for example to select an option from a menu. Much have been written about visual search both in terms of text and graphical representations, especially icons.

Photographs and photorealistic images are quite different from icons in that they often are rendered as simplified canonical representations of objects while photograph are real representations. The human visual system is trained to perceive and interpret the real world, while on the other end of the scale technical drawings require time to decipher. Icons are often viewed as "user friendly", but as simplified renderings of real objects they need to be learned, especially if they represent abstract concepts such as traffic signs.

On the other hand, text represents a visual coding in which most people have many years of training and experience. Readers tend to recognize entire words instead of individual letters when these are presented using lowercase characters. Trained readers perceive and interpret text at very high rates even though text are advanced sequences of abstract symbols. The question thus arises whether to use text, icons, images, or a combination to facilitate the easy identification of various elements in user interfaces.

This study had two objectives: First, we wanted to collect empiric evidence on the search performance of users when dealing with photorealistic images versus text. Second, we wanted to explore the suitability of using an off-the shelf web-based application Kahoot for setting up and conducting simple experiments.

The rest of this paper is organized as follows. The following section reviews related studies into visual search using text and images. Section 3 presents the methodology. This is followed by results and discussion. The paper closes with concluding remarks in Section 6.

2 Related work

Information search is a huge research field and visual search is one part of this. Moreover, semiotics, or the study of symbols, is another established field. More recently, symbols have also been studied in more applied domains such as product design and information technology [1, 2] although Fleetwood [3] in 2002 claimed that there were few studies on icons despite the longevity of graphical user interfaces. Icons have also been studied outside the computing domain, for example the visual search for warning labels on products [4], where it was found that the participants preferred icons and signal words in red.

In one of the early writings on computer icons Byrne [5] claimed that simple icons with clear and few characteristics is the most beneficial, and that complex icons are no better than simple rectangles. Fleetwood and Bryne studied the relationship between search performance and icon quality [3, 6]. Klöckner, Wirschum, and Jameson [7] investigated results lists from search engines and discussed two search list exploration strategies taken by users, namely breadth first and depth first search. Everett and Byrne [8] studied how spacing between icons in a grid affects visual search performance.

Focusing on the visual appearance of icons Arledge [9] compared the effects of icons represented by outlines versus those that are filled. Garcia, Badre and Stasko [10] investigated the level of abstraction of icons and found that context is affecting the understanding of icons. Cho et al. [11] probed senior users' preference for level of realism and level of abstraction. Icon location has also been found to be an important factor in addition to the icon shape [12]. Blankenberger and Hahn [13] claimed that trained users remember where icons are located, that icon design has little effect, and that icon design is more influential for novice users.

Holloway and Bailey [14] compared students and developers' perceptions of icons and found that developers had a different understanding of the icons compared to the students. They concluded that the selection of icons should involve the target audience and not only developers. The involvement of the target audience in icon selection is also echoed by Berget and Sandnes [15] who tested how well students recognized common icons in the public domain. They found that many of the participants did not successfully recognize seemingly common objects such as wrenches and parachutes.

Studies have also addressed the effect of contrast and sharpness for icon search. A study using an eye-tracking methodology [16] found that contrast was more important than sharpness for icons arranged in a grid. Luminance contrast has also been studied extensively for the readability of text in relation to the background [17, 18, 19]. Clearly legible text is a prerequisite for visual search [20]. The choice of colors is affected by the color pickers used [21] and different approaches to color pickers have been proposed to facilitate color combinations that ensure enough luminance contrast [22, 23,

24]. Huang [25] studied specific figure-ground color combinations for icons as well as outline width and icon shape. It was wound that colors influenced search time. Outline width, border and icon shape was found to have no effect on error rates. The combination of text in visual search has also been explored. In an experiment of visual search with structured layouts Hornof [26] found that the search was more efficient when structured groups were labelled compared to unlabeled. The benefit of redundant coding with icons and text was also observed in menu selection tasks [27]. Although no difference in response time was observed, they did find fewer errors when icons and text were used together.

Icons, images, and text has also been studied in context of disability. For instance, researchers have explored whether icons are more effective than text for users with dyslexia [28] and for users with a diagnosis on the autism spectrum [29].

Juola, Ward and McNamara [30] studied participants' ability to detect absence or presence of letters in strings. Implications of knowledge about search behavior include the manner and order in which search engine results are presented [31].

More recently there has been an interest in icons in connection with the emergence of smartphones and other handheld devices that have limited display real estate. Issues studied includes how users organize their icons [32] and change blindness [33], more specifically how the inability of noticing icon changes communicating status changes is related to the number of icons.

3 Method

3.1 Experimental design

A controlled experiment with one within-group independent variable and three dependent variables was designed. The within-group variable had two levels, namely search in images and search in text. The three dependent variables included mean search completion time, error rate and preference. The preference variable was dichotomous.

3.2 Participants

A total of 15 participants was recruited randomly from the campus of the authors' university. All the participants were students in the age range of 19 to 25 years with an approximate gender balance. None of the respondents reported any issues related to reduced vision such as color blindness or cognition such as dyslexia [34, 35].

Fig. 1. Example of image search task (identify the image without a mobile phone).

Fig. 2. Example of textual search task (identify the text with the letter “d”).

3.3 Materials

A total of 16 tasks were designed of which eight involved image search and eight textual searches. Each task entailed presenting four alternatives of which one was the correct answer. The image search tasks typically comprised identifying the picture with or without a given object, or quantity of objects (see Fig. 1). The images were of varying levels of difficulty such as simple images, stylized cartoons, artistic drawings and for instance an owl camouflaged in its natural habitat. The text search task involved identifying texts with or without a given letter (see Fig. 2) similar to the procedure reported in [30], identifying a specific word or a word denoting a certain concept such as objects

that are blue. The text strings were of similar lengths. Instructions were given in Norwegian. Steps were taken to ensure the objects were universally familiar. The tasks were designed to have varying levels of difficulty. The tasks were presented in increasing level of difficulty alternating between image tasks and text tasks.

The experiment was implemented using Kahoot which is intended as an educational and entertaining quiz tool [36, 37, 38]. Kahoot was generally well-known among the cohort of students recruited. Kahoot facilitates four options as images or text and provides time-taking functionality and bookkeeping of correct and incorrect responses.

3.4 Procedure

The participants were first briefed about the experiment before the participants gave their consent to participate. Participation was voluntary and anonymous. No personal information was collected. Participants had the freedom to withdraw from the experiment at any time without having to provide a reason.

The participants were tested individually in an undisturbed meeting room with the two first authors present as facilitators. Each session lasted approximately 15 minutes. A maximum time limit of 30 seconds was set for each task and a countdown-timer displayed the remaining time. The participants responded to the quiz using a smartphone but a laptop computer with the same quiz mirrored was also placed in front of the participant. After completing the 16 tasks the participants were asked to indicate their preferred mode of searching.

3.5 Analysis

The median time for each task type was computed for each participant. The errors were analyzed in terms of incorrect responses and number of missing responses within the 30 second window. The observations were checked for adherence to normal distribution using Shapiro-Wilk tests. The results were analyzed using JASP 0.12.0.0 [39].

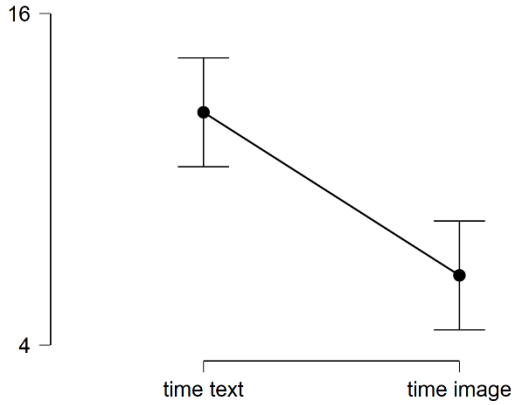

Fig. 3. Mean task complete time in seconds. Error bars show 95% confidence intervals.

4 Results

Fig. 3 shows the mean task completion times in seconds for the textual search and image search. Clearly, task completion times are much larger with textual search (M = 12.4, SD = 5.6) compared to image search (M = 6.5, SD = 1.6). The two confidence intervals are not overlapping. A paired t-test confirms that the search times are significantly different ($t(14) = 4.54$, $p < .001$, Cohen's d = 1.173, 95% CI [0.497, 1.826]).

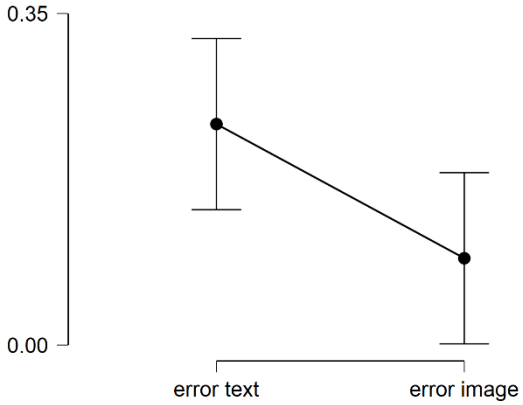

Fig. 4. Error rate. Error bars show 95% confidence intervals.

Fig. 4 shows the error rates for the two modes of searching. The error rate is much higher with textual search (M = 23.3%, SD = 18.2%) compared to image-based search (M = 9.2%, SD = 7.4%). A paired t-test shows that the two error rates are significantly different ($t(14) = 3.3$, $p = .032$, Cohen's d = 0.614, 95% CI [0.051, 1.159]).

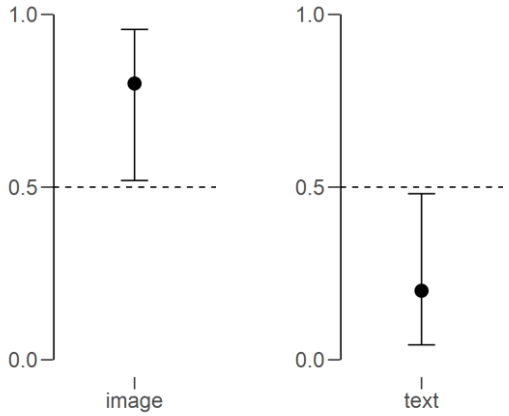

Fig. 5. Preferences for image and text search. Error bars show 95% confidence intervals.

Of the 15 participants a majority of 12 participant preferred image-based search while 3 participants indicated a preference for the text-based search. This 80% preference for image search had a 95% CI of 51.9% to 95.7%, $p = .035$, while the 20% preference for textual search had a 95% CI of 4.3% to 48.1%, $p = .035$ (see Fig. 5). A majority thus prefer image search.

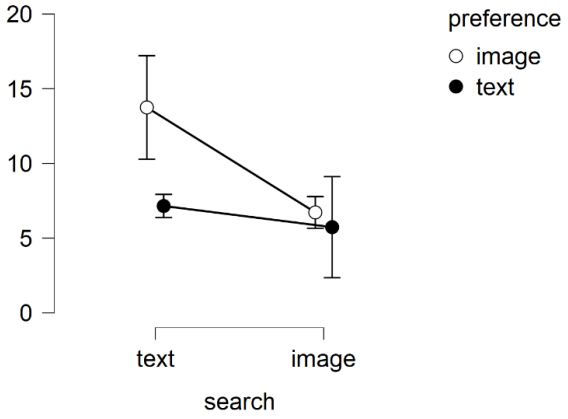

Fig. 6. Task completion time plotted according to preference. Error bars show 95% CI.

As only three participants preferred text search it was not feasible to conduct further statistical analyses. However, the task completion times for the two search modes were plotted according to preference to get an indication the potential trend (see Fig. 6). The results show that the participants who prefer textual search perform very similar under both conditions, while participants who prefer image search appears to take much shorter time with textual search. Moreover, the 95% confidence intervals reveal that the spread is larger in the least preferred mode compared to the preferred mode for both preferences.

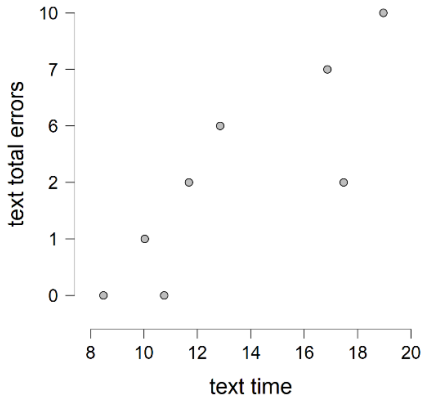

Fig. 7. Scatter plot showing the relationship between the total number of errors and the task completion time.

The responses were also analyzed according to question. The results confirm that the questions had different levels of difficulty as intended. The total number of errors, that is, the sum of actual errors and number of missing responses correlated positively and significantly with the response time for both textual search ($r(16) = 0.778$, $p = .023$, 95% CI [0.163, 0.958]) and visual search ($r(16) = 0.965$, $p < .001$, 95% CI [0.811, 0.994]). Moreover, the number of errors also correlated positively and significantly for both image search ($r_s(16) = 0.756$, $p = .030$, 95% CI [0.110, 0.953]) and text search ($r_s(16) = 0.793$, $p = .019$, 95% CI[0.200, 0.961]). The connection between search time and the number of errors in the textual questions is illustrated in Fig. 7. Only observations for the text tasks are shown since text search were associated with most errors.

5 Discussions

The results show that image search is much faster than text search. In time-critical tasks it is important to facilitate search tasks such that the delay associated with the visual search is reduced. Not only is the search speed critical, but also the error rate. It is not enough for a search to be quick if the result is incorrect. The search needs to be quick and preferably error free. However, the results show that the error rates follow the same trend as the search time. Image-based search was associated with a much lower error rate compared to textual search. These results partially agree with the results obtained when contrasting icons and text where lower error rates were achieved with icons although not difference in search time [27]. This may be because icons are simple not photorealistic representations.

The results give a strong indication that users in general prefer image-oriented search. It is useful from a practical perspective that users' preferences are consistent with performance characteristics, that is, for the types of search studied users prefer the search modes with the shortest task completion times and lowest error rate. Although generalizable, the benefit of image search is not suitable for all users [17, 28, 29]. The convention of redundant coding where image is used together with text and other characteristics seems sensible [26, 27].

The battery of questions was intentionally varied in level of difficulty, and these variations were observed in the results. Difficult search tasks took longer than easier search tasks. The difficult search tasks also associated with more errors and missing responses than easier search tasks.

Kahoot proved itself as a convenient platform for conducting controlled experiments as it was easy to set up experiments. Kahoot managed the stopwatch tasks and tallied correct answers which allowed the answers to be conveniently downloaded into an Excel sheet for subsequent statistical analysis. Kahoot also supports randomization of questions, although we did not exploit this function. The participants provided positive feedback on participating in the Kahoot experiment. It is thus possible that Kahoot may be more engaging than traditional electronic questionnaires. However, it is also possible that the timed tasks were perceived as somewhat stressful. Perhaps the element of competition could have been reduced by not displaying the count down clock during the tasks?

5.1 Limitations

One limitation of this study is the small population size. Hence, it was not possible to conduct detailed analysis of performance according to preference. To perform such analyses a larger sample of participants with a preference for textual search is needed which require pre-screening of the participations. The cohort was also limited. Although the results clearly show distinct trends for image and text searches it is possible that a different cohort would yield different results.

Another limitation of this study is that the presentation order was fixed. Hence, it is possible that there may be some confounding effects. However, attempts were taken to limit this by intertwining the presentation of image-based and textual search tasks.

There was also an issue using the mobile handset with the image search tasks as the display made it hard to see details although a laptop computer also displayed the tasks. For Detail-rich images the experiment should perhaps be employed on experimental setups with a larger display only.

6 Conclusions

A simple experiment was conducted were Kahoot was used to measure users' search performance in images and texts, respectively. The results show that users search time and error rates are smaller with image search compared to textual search. Most participants were in favor of image search. Our experiences revealed that Kahoot was straightforward to configure and is therefore a convenient platform for conducting certain types of controlled experiments.

References

1. Gittins, D.: Icon-based human-computer interaction. International Journal of Man-Machine Studies **24**(6), 519-543 (1986).
2. Lin, R., Kreifeldt, J. G.: Understanding the image functions for icon design. In Proceedings of the Human Factors Society Annual Meeting **36**(4), pp. 341-345. Sage CA: Los Angeles, CA: SAGE Publications (1992).
3. Fleetwood, M. D., Byrne, M. D.: Modeling icon search in ACT-R/PM. Cognitive Systems Research **3**(1), 25-33 (2002).
4. Bzostek, J. A., Wogalter, M. S.: Measuring visual search time for a product warning label as a function of icon, color, column and vertical placement. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting **43**(16), pp. 888-892. Sage CA: Los Angeles, CA: Sage Publications (1999).
5. Byrne, M. D.: Using icons to find documents: simplicity is critical. In: Proc. ACM INTERCHI'93 Conference on Human Factors in Computing Systems, pp. 446-453 (1993).
6. Fleetwood, M. D., Byrne, M. D.: Modeling the visual search of displays: A revised ACT-R/PM model of icon search based on eye-tracking and experimental data. In: Proceedings of the Fifth International Conference on Cognitive Modeling, pp. 87-92. Bamberg,, Germany: Universitas-Verlag Bamberg (2003).

7. Klöckner, K., Wirschum, N., Jameson, A.: Depth-and breadth-first processing of search result lists. In: CHI'04 extended abstracts on Human factors in computing systems, pp. 1539-1539. ACM (2004).
8. Everett, S. P., Byrne, M. D.: Unintended effects: Varying icon spacing changes users' visual search strategy. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 695-702. ACM (2004).
9. Arledge, C.: Filled-in vs. Outline Icons: The Impact of Icon Style on Usability. Masters thesis (2014).
10. García, M., Badre, A. N., Stasko, J. T.: Development and validation of icons varying in their abstractness. Interacting with computers **6**(2), 191-211 (1994).
11. Cho, M., Kwon, S., Na, N., Suk, H. J., Lee, K.: The elders preference for skeuomorphism as app icon style. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 899-904. ACM (2015).
12. Moyes, J.: When users do and don't rely on icon shape. In: Conference companion on Human factors in computing systems, pp. 283-284. (1994).
13. Blankenberger, S., Hahn, K.: Effects of icon design on human-computer interaction. International Journal of Man-Machine Studies **35**(3), 363-377 (1991).
14. Holloway, J. B., Bailey, J. H.: Don't use a product's developers for icon testing. In: Conference Companion on Human Factors in Computing Systems, pp. 309-310. ACM (1996).
15. Berget, G., Sandnes, F. E.: On the Understandability of Public Domain Icons: Effects of Gender and Age. In: International Conference on Universal Access in Human-Computer Interaction, pp. 387-396. Springer, Cham. (2015).
16. Näsänen, R., Ojanpää, H.: Effect of image contrast and sharpness on visual search for computer icons. Displays **24**(3), 137-144 (2003).
17. Sandnes, F. E.: Understanding WCAG2. 0 color contrast requirements through 3D color space visualization. Stud. Health Technol. Inform **229**, 366-375 (2016).
18. Sandnes, F. E.: On-screen colour contrast for visually impaired readers (pp. 405-416). Routledge (2017).
19. Sandnes, F. E.: An image-based visual strategy for working with color contrasts during design. In: International Conference on Computers Helping People with Special Needs, pp. 35-42. Springer, Cham (2018).
20. Ling, J., Van Schaik, P.: The effect of text and background colour on visual search of Web pages. Displays **23**(5), 223-230 (2002).
21. Brathovde, K., Farner, M. B., Brun, F. K., Sandnes, F. E.: Effectiveness of Color-Picking Interfaces among Non-Designers. In: International Conference on Cooperative Design, Visualization and Engineering, pp. 181-189. Springer, Cham. (2019).
22. Sandnes, F. E., Zhao, A.: A contrast colour selection scheme for WCAG2. 0-compliant web designs based on HSV-half-planes. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 1233-1237. IEEE (2015).
23. Sandnes, F. E., Zhao, A.: An interactive color picker that ensures WCAG2. 0 compliant color contrast levels. Procedia Computer Science **67**, 87-94 (2015).
24. Hansen, F., Krivan, J. J., Sandnes, F. E.: Still Not Readable? An Interactive Tool for Recommending Color Pairs with Sufficient Contrast based on Existing Visual Designs. In: The 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 636-638. ACM (2019).
25. Huang, K. C., Chiu, T. L.: Visual search performance on an LCD monitor: effects of color combination of figure and icon background, shape of icon, and line width of icon border. Perceptual and motor skills **104**(2), 562-574 (2007).

26. Hornof, A. J.: Visual search and mouse pointing in labeled versus unlabeled two-dimensional visual hierarchies. ACM Transactions on Computer-Human Interaction **8**, 171-197 (2001).
27. Kacmar, C. J., Carey, J. M.: Assessing the usability of icons in user interfaces. Behaviour & Information Technology **10**(6), 443-457 (1991).
28. Berget, G., Sandnes, F. E.: The effect of dyslexia on searching visual and textual content: are icons really useful?. In: International Conference on Universal Access in Human-Computer Interaction, pp. 616-625. Springer, Cham. (2015).
29. Joseph, R. M., Keehn, B., Connolly, C., Wolfe, J. M., Horowitz, T. S.: Why is visual search superior in autism spectrum disorder?. Developmental science **12**(6), 1083-1096 (2009).
30. Juola, J. F., Ward, N. J., McNamara, T.: Visual search and reading of rapid serial presentations of letter strings, words, and text. Journal of Experimental Psychology: General **111**(2), 208 (1982).
31. Hsu, W. H., Kennedy, L. S., Chang, S. F.: Reranking methods for visual search. IEEE MultiMedia **14**(3), 14-22 (2007).
32. Böhmer, M., Krüger, A.: A study on icon arrangement by smartphone users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2137-2146. ACM (2013).
33. Davies, T., Beeharee, A.: The case of the missed icon: change blindness on mobile devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1451-1460. ACM (2012).
34. Berget, G., Mulvey, F., Sandnes, F. E.: Is visual content in textual search interfaces beneficial to dyslexic users?. International Journal of Human-Computer Studies **92**, 17-29 (2016).
35. Berget, G., Sandnes, F. E.: Searching Databases without Query-Building Aids: Implications for Dyslexic Users. Information Research: An International Electronic Journal **20**(4), (2015).
36. Dellos, R.: Kahoot! A digital game resource for learning. International Journal of Instructional Technology and Distance Learning **12**(4), 49-52 (2015).
37. Zarzycka-Piskorz, E.: Kahoot it or not? Can games be motivating in learning grammar?. Teaching English with Technology **16**(3), 17-36 (2016).
38. Licorish, S. A., Owen, H. E., Daniel, B., George, J. L.: Students' perception of Kahoot!'s influence on teaching and learning. Research and Practice in Technology Enhanced Learning **13**(1), (2018).
39. JASP Team (2020). JASP (Version 0.12.0)[Computer software].