

How Does Knowing What You Are Looking For Change Visual Search Behavior?

Duncan P. Brumby, Anna L. Cox, Jacqueline Chung, Byron Fernandes

UCL Interaction Centre
University College London
London WC1E 6BT UK
Brumby@cs.ucl.ac.uk

ABSTRACT

When searching a display, users sometimes know what the target is but sometimes do not. It has generally been assumed that for this latter case people must engage in a deeper semantic evaluation of items during the search process. This idea is central to Information Foraging theory. But do people actually spend longer assessing items when engaged in a semantically demanding search task? We investigate this by having participants locate target items in 16-item menus. Participants were either told exactly what to look for (known-item search) or they were told the category that the target belonged to (semantic search). Participants were faster and more accurate at known-item searches. Eye-movement data show that this was because participants were more likely to skip over items when performing known-item searches. Contrary to expectation, we found limited empirical evidence to support the idea that deeper semantic evaluations of items lead to longer gaze durations (this occurred only when items were arranged very close together). This finding is important because it reveals how people adopt different eye gaze strategies depending on the kind of search activity they are engaged in.

Author Keywords

Visual search; menus; information foraging; eye-tracking.

ACM Classification Keywords

H.5.2 User Interfaces: evaluation/methodology, screen design, theory and methods; H.1.2 User/Machine Systems: human factors, human information processing.

INTRODUCTION

Computer users spend a great deal of time looking at and searching displays, whether that be launching an application from a menu on a smartphone or finding the cheapest flight to Toronto on a webpage. It is perhaps not surprising then

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.

Copyright ©ACM 978-1-4503-2473-1/14/04...\$15.00.

<http://dx.doi.org/10.1145/2556288.2557064>

that since the inception of HCI there has been a sustained interest in studying visual search [3,4,6,7,9-11,13,15].

Across the literature, a basic distinction can be made in the kind of search tasks that users have been asked to perform. In one tradition, the user is given a precise description of the target item that must be found. We refer to this as *known-item* search, and a number of studies have investigated this in detail [3,6,7,15]. An everyday example of this kind of search might be launching the Angry Birds application on a smartphone. This is a relatively simple search task because the user knows precisely what it is they are looking for. It has therefore been assumed that items need only be assessed in terms of their surface features [5,10,17,19]

In contrast to known-item search, there are many instances in which the user has a clear goal in mind but only a vague idea of how the target will be specified. We refer to this as *semantic* search. Studies have investigated this kind of search, with a particular focus given to how people look for information on the Internet [2,4,8]. Information foraging theory [9,16] has been highly influential for understanding this activity because it describes how people go about assessing the semantic relatedness of items to their search goal – that is, the *information scent*. Given that it takes time to retrieve information from memory, it has long been speculated that assessments of items should be longer during semantic search than during known-item search [17,19].

In this paper, we investigate differences in how people perform known-item and semantic searches. These tasks have been extensively studied in the literature, but separately – there have been surprisingly few attempts to compare how people perform these different kinds of search tasks. What few studies there have been have found that known-item searches are faster and more accurate than semantic searches [12,13]. However, these early studies did not make use of eye-tracking techniques to investigate underlying changes in search strategy. If semantically assessing an item incurs an added time cost we might expect this to be reflected in eye movement data: Fixations to items should be longer during semantic searches than during known-item searches. We provide the first direct test of this hypothesis.

A further aspect of search behavior that concerns us here is the observation that people often do not sequentially look at every item in turn when scanning through a menu – occasionally items are skipped. This item skipping behavior has been documented in known-item searches [5,10] as well as semantic searches [4]. It is generally assumed that the reason people skip over items while searching is that they accumulate enough visual information to rule-out an item while it is in peripheral vision. We test a novel prediction of this account: People should be more likely to skip items when performing known-item searches than semantic searches. This is because items can be assessed on the basis of visual features during known-item searches. Hence, items can be assessed quickly and on the basis of partial information available from peripheral vision. In contrast, assessing the semantic meaning of an item requires it to be read carefully. Hence, items must be directly fixated. We investigate this skipping hypothesis by varying the physical distance between menu items, as this will determine how much visual information is available from each eye movement: more items can be assessed within a single fixation when items are placed closer together [10].

METHOD

Participants

Thirty-three participants (six male) were recruited through the UCL Psychology Subject Pool to take part in the study. Participants were between 19 and 21 years of age. All were native English speakers, with normal or corrected-to-normal vision, and experienced computer users.

Materials

A target item had to be found in a menu containing 16 items. Each item in the menu was a single word, which was sampled from a dataset (adapted from [18]) of 560 unique words (with character length $M = 6.2$, $SD = 2.0$, range: 2 – 12). The words in this dataset were organized into natural categories. There were 14 top-level categories (e.g., animal, building, entertainment), with each of these separated into four mid-level categories (e.g., bird, farm animal, tropical fish, and wild animal are all members of the animal category). Finally, for each mid-level category there were 10 instances (e.g., bluebird, canary, starling, eagle, hawk are all members of the bird sub-category). Using this dataset it was possible to generate menus in which every item in the menu came from a different mid-level category.

Menu items were arranged in a single vertical list, one item per line. The vertical separation between items was varied between conditions (see design). The first item in the menu was positioned 80 pixels from the top of the screen. All text was presented in an Arial font, size 10. Stimuli were presented on a 17-inch TFT monitor set at a resolution of 1280 by 1024 pixels. Eye movements were recorded using a Tobii 1750 eye tracker. The sampling rate of the eye tracker was 50 Hz, with gaze point accuracy of less than 0.5 degrees of visual angle. The experimental software ran on a

Dell Optiplex machine with 1 GB of RAM running Microsoft Windows XP. An optical mouse was used, set at the ‘medium’ speed via the system control panel.

Design

A 2 x 3 (search type x layout) within-subjects design was used. We varied the target description that was given to participants between different search type conditions. For known-item searches, participants were given the target word (e.g., find the word canary). For semantic searches, participants were given a mid-level category description of the target (e.g., find a type of bird). For the manipulation of menu layout the vertical separation between items was systematically varied from a small gap (17 pixels, 0.40° visual angle), to a medium gap (35.5 pixels, 0.85° visual angle), to a large gap (55 pixels, 1.3° visual angle).

Procedure

Participants were informed that they would be required to search menus to locate a target item and that they would be given either the precise label for the item that had to be found in the menu or they would be given a category label for the item. This target description was given at the start of each search trial. After reading it, participants started the trial by clicking on a button on the screen. This made the target description disappear, and the menu was displayed beneath it. Participants were told that there was only a single target item in each menu and that they had to locate it as quickly as possible. Participants selected an item by clicking on it with the mouse. If an incorrect selection was made, participants were instructed to make another selection from the same menu after being shown the target information again. The trial ended when the participant had successfully selected the target. Before proceeding to the next trial, participants were given feedback on the time taken to locate the target. The purpose of this feedback was to reinforce the instruction to locate the target quickly.

Each participant completed 84 menu search trials, divided into six blocks of 14 trials each (grouped by each of the search type x gap size experimental conditions). Each block had an initial set of four practice trials, followed by 10 critical trials. Each trial consisted of a unique (and automatically generated) menu. The order in which each condition was experienced was partially counter-balanced between participants: Half of the participants did all of the known-item searches before the semantic searches, or vice versa. Within each search type, the order in which each display layout was experienced was randomized across participants (but held consistent across conditions). Participants were given the opportunity for a short break after completing each block of 14 trials. The eye tracker was calibrated before the start of each block to ensure accurate gaze tracking. Participants were seated approximately 60-70 cm from the monitor throughout the experiment. The entire experiment took approximately 30 minutes to complete.

Figure 1. Search time as function of target position.

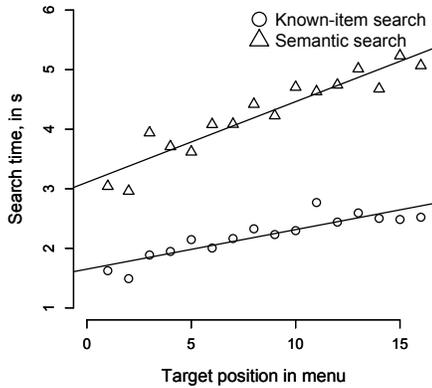


Figure 2. Duration of item visits by condition.

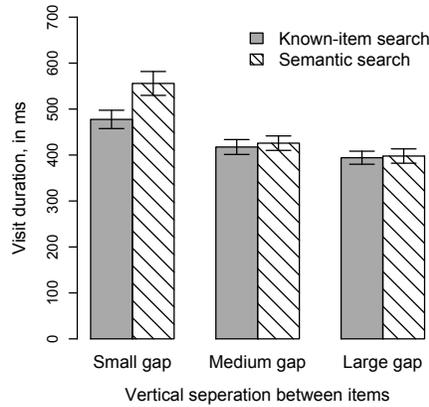
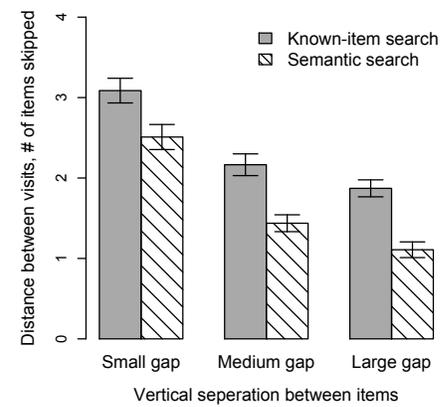


Figure 3. Distance between item visits by condition.



RESULTS

A 2 x 3 (search type x layout) repeated measures ANOVA was used, with a .05 significance level for judging the significance of effects. We excluded trials from the analysis in which the participant did not correctly select the target item on the first attempt. Accuracy was generally very high, but participants did make more errors when performing semantic searches ($M = 18\%$, $SD = 12\%$) than known-item searches ($M = 1\%$, $SD = 4\%$), $F(1, 32) = 139.20$, $p < .001$. There was no significant main effect of layout nor was there a significant interaction.

In terms of search time, participants were faster at locating targets when performing known-item searches ($M = 2.21$, $SD = 0.45$) than semantic searches ($M = 3.68$, $SD = 0.77$), $F(1, 32) = 269.80$, $p < .001$. Search times also became slower as the size of the vertical gap between items increased from a small gap ($M = 2.76$, $SD = 0.93$), to a medium gap ($M = 2.88$, $SD = 0.87$), to a large gap ($M = 3.20$, $SD = 1.05$), $F(2, 64) = 20.18$, $p < .001$. There was no significant interaction. In addition to this basic analysis, we also consider how search time varies as a function of the position of the target within the menu. Figure 1 shows that there is a simple linear relationship between search time and target position for both known-item searches, $y = .14x + 3.11$, $R^2 = 0.80$, and for semantic searches, $y = .07x + 1.65$, $R^2 = 0.89$. The larger slope and intercept for the semantic search model suggest a slower and more deliberate assessment strategy during this kind of search.

We next consider eye movement data to understand how participants were assessing items during search. For each search trial, we define a *visit sequence*. To do this, we map each fixation to the nearest item in the menu (up to a maximum distance threshold), before collapsing multiple contiguous fixations to the same item together. The resulting visit sequence describes the order in which items in the menu were assessed as well as the duration of each of these assessments. From this we can calculate for each trial the mean duration of visits as well as the mean distance between consecutive visits.

Figure 2 shows the mean duration of item visits across conditions. It can be seen that there is a significant search-

type by layout interaction, $F(2, 64) = 6.14$, $p < .01$. Follow-up tests of this interaction show that item visits were longer for semantic searches than known-item searches, but only when there was a small gap between each item, $F(1, 32) = 11.91$, $p < .005$. When there was a larger gap between items (medium or large), there was no effect of search type on the duration of item visits, F 's < 1 . It can also be seen that the duration of item visits increased when items were closer together; that is, there was a main effect of gap size on visit durations, $F(2, 64) = 49.51$, $p < .001$.

Figure 3 shows the distance between consecutive item visits. This provides an indication of the mean number of items that were skipped over from one visit to the next during a search trial. For example, imagine a sequence in which items 1, 3, 6, 7 are visited. The number of items skipped between each visit is then 1, 2, 0, giving a mean of 1 item being skipped between visits for this sequence. Figure 3 shows that participants tended to skip over more items between visits when performing known-item searches than when performing semantic searches, $F(1, 32) = 23.52$, $p < .001$. Participants also skipped over more items when they were closer together, $F(2, 64) = 30.83$, $p < .001$. The interaction was not significant.

Finally, we consider how often participants selected the target item straight after visiting it for the first time – we refer to this as self-terminating search [11]. This measure gives an index of how often the target was missed on the first pass. There were more self-terminating search trials for known-item searches ($M = 86\%$, $SD = 20\%$) than for semantic searches ($M = 61\%$, $SD = 21\%$), $F(1, 32) = 89.94$, $p < .001$. This suggests that when performing known-item searches participants rarely missed the target when it is within their gaze. There was no effect of gap size nor was there a significant interaction.

DISCUSSION

The results of this study show how eye gaze strategy is influenced by the kind of search activity a user is performing. We found that during known-item searches participants were more likely to skip over items when searching a menu and were more likely to select the target immediately after locating it. In contrast, during semantic

searches participants were more likely to assess items in turn. This finding could potentially inform the development of real-time gaze tracking systems that try to infer a user's search intentions based on their eye movements [20].

We found limited empirical evidence to support the intuitive idea that assessing the semantics of a label during search – inferring an item's *information scent* – is reflected in longer item gazes. Our results show that item gazes were longer for semantic searches but only when items were arranged close together. These longer gazes probably occurred because the close proximity of items allowed participants to assess multiple items with a single fixation. Consistent with this we also found that participants were more likely to skip over items when searching tightly packed menus. In contrast, when items were placed farther apart, assessments had to occur individually, and we found no evidence to suggest that semantic assessments incur an added time cost per assessment in this situation. We should however remain mindful that it is difficult to rule out the possibility that participants might have been accessing the semantic meaning of an item from memory *after* the focus of gaze had moved on to the next item in the menu [1].

One way to explore these explanations in more detail is to model these data. There is a long tradition of modeling menu selection tasks and various accounts of the underlying cognitive and perpetual-motor processes have been put forward [3,6,7,9-11]. We might speculate that Hornof's model [10] should be able to account for the item skipping behavior observed here. However, this model currently lacks a mechanistic account of the memory retrieval process required for semantic search. This could be reconciled by incorporating elements from Fu and Pirolli's model [9], which provides a detailed account of this process. Of particular interest to us though is recent work by Myers et al. [14] who use Reinforcement Learning techniques to explain the emergence of various eye movement strategies in a visual search task. Future work is necessary to ascertain whether these models can capture the data reported here.

ACKNOWLEDGEMENTS

We thank Gilles Bailly, Antti Oulasvirta, and Andrew Howes for encouraging us to publish this work, as well as our reviewers for their constructive feedback.

REFERENCES

1. Anderson, J.R., Bothell, D., & Douglass, S. Eye movements do not reflect retrieval: Limits of the eye-mind hypothesis. *Psychological Science*, 15, (2004), 225-231.
2. Aula, A., Khan, R.M., & Guan, Z. How does search behavior change as search becomes more difficult? *Proc. CHI '10*, ACM (2010), 35-44.
3. Bailly, G., Oulasvirta, A., Brumby, D.P., & Howes, A. Model of visual search and selection time in linear menus. *Proc. CHI '14*, ACM (2014).
4. Brumby, D.P. & Howes, A. Strategies for guiding interactive search: An empirical investigation into the consequences of label relevance for assessment and selection. *Human-Computer Interaction*, 23, (2008), 1-46.
5. Byrne, M.D. ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies*, 55, (2001), 41-84.
6. Card, S.K., Visual search of computer command menus. In H. Bouma and D. Bouwhuis (Eds.) *Attention and Performance X*, LEA, 1984, 97-108.
7. Cockburn, A., Gutwin, C., & Greenberg, S. A predictive model of menu performance. *Proc. CHI '07*, ACM (2007), 627-636.
8. Cutrell, E. & Guan, Z. What are you looking for? An eye-tracking study of information usage in web search. *Proc. CHI '07*, ACM (2007), 407-416.
9. Fu, W.-T. & Pirolli, P. SNIF-ACT: A cognitive model of user navigation on the world wide web. *Human-Computer Interaction*, 22, (2007), 355-412.
10. Hornof, A.J. Cognitive strategies for the visual search of hierarchical computer displays. *Human-Computer Interaction*, 19, (2004), 183-223.
11. Lee, E. & MacGregor, J. Minimizing user search time in menu retrieval systems. *Human Factors*, 27, (1985), 157-162.
12. McDonald, J.E., Stone, J.D., & Liebelt, L.S. Searching for items in menus: The effects of organization and type of target. *Proc. Human Factors and Ergonomics Society Annual Meeting*, HFES (1983), 834-837.
13. Mehlenbacher, B., Duffy, T.M., & Palmer, J. Finding information on a menu: Linking menu organization to the user's goals. *Human-Computer Interaction*, 4, (1989), 231-251.
14. Myers, C.W., Lewis, R.L., & Howes, A. Bounded optimal state estimation and control in visual search: Explaining distractor ratio effects. *Proc. CogSci 2013*, Cognitive Science Society (2013), 1049-1054.
15. Nilsen, E.L. *Perceptual-motor control in human-computer interaction*. Report 37, The Cognitive Sci. & Machine Intelligence Lab. Univ. of Michigan (1991).
16. Pirolli, P., *Information Foraging Theory: Adaptive Interaction with Information*. Oxford University Press, 2009.
17. Rieman, J., Young, R.M., & Howes, A. A dual-space model of iteratively deepening exploratory learning. *International Journal of Human-Computer Studies*, 44, (1996), 743-775.
18. Yoon, C., Feinberg, F., Hu, P., et al. Category norms as a function of culture and age: Comparisons of item responses to 105 categories by American and Chinese adults. *Psychology and Aging*, 19, (2004), 379-393.
19. Young, R.M., Rational analysis of exploratory choice. In M. Oaksford and N. Chater (Eds.), *Rational Models of Cognition*. Oxford University Press, 1998, 469-500.
20. The all-telling eye, *The Economist*, 22 Oct 2011. <http://www.economist.com/node/21533362>.