To appear as Work-in-Progress at CHI 2008,
April 5-10, 2008, Florence, Italy

# The Effects of Semantic Grouping on Visual Search

**Tim Halverson and Anthony J. Hornof**
Department of Computer and Information Science
University of Oregon
Eugene, OR 97405 USA
{thalvers,hornof}@cs.uoregon.edu

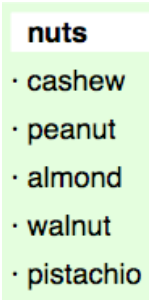


**Figure 1.** A semantically cohesive group from an experimental layout. "Nuts" is the group label. The remaining words are the menu items.

## Abstract
This paper reports on work-in-progress to better understand how users visually interact with hierarchically organized semantic information. Experimental reaction time and eye movement data are reported that give insight into strategies people employ while searching visual layouts containing words that are either grouped by category (i.e. semantically cohesive) or randomly grouped. Additionally, sometimes the category labels of the cohesive groups are displayed as part of the group. Preliminary results suggest that: (a) When groups are cohesive, people tend to search labeled and unlabeled layouts similarly. (b) People seem to trust the categorical information of labels more than non-labels. This work will be used to extend current computational models of visual search to better predict users visual interaction with interfaces.


CHI 2008, April 5 – April 10, 2008, Florence, Italy
ACM 1-xxxxxxxxxxxxxxxxx.

## Keywords
Visual search, semantics, visual grouping, cognitive strategies, eye movements

## ACM Classification Keywords
H5.2. [Information Interfaces and Presentation]: User Interfaces – screen design, theory and methods

## Introduction
For a great deal of human-computer interaction (HCI), visual search is an essential process. However, our theoretical understanding of how users will interact with everyday devices is lacking. In this paper, we present preliminary results from research that (a) deepens our understanding of how people visually interact with interfaces as a function of the semantic content of the interface and (b) extends current computational models that predict users' interaction with visual interfaces.

Previous research has shown how group labels, like the word "nuts" in Figure 1, affect visual search strategies and aid the user in finding a target when the target label is known [5]. Other research has shown how the semantic information in menu items affects visual

search [1]. But what happens when the exact group label is not known? How does the semantic content of group labels, or the grouping of menu items, guide visual search? This research investigates how semantics may affect users' visual search strategies in visually and semantically grouped layouts of menu items.

Brumby and Howes [1] found that when searching a menu, people tend to search fewer items when distractor menu items are less similar to the goal and when the target is more similar to the goal. Further, people tend to revisit smaller and smaller groups of menu items as visual search progresses. Both Brumby and Howes [1] and Fu and Pirolli [3] propose models that account for the effects of semantic content on visual search. In both cases, the models assume no visual hierarchy and that menu items are sequentially inspected. However, many user interfaces contain some form of visual hierarchy, semantic hierarchy, or both. In fact, earlier research found menu search was initially faster when items were functionally grouped [7].

The remainder of this paper discusses current and future work that investigates the effects of visual and semantic grouping on users' visual interaction. The task is described, the results of an experiment are discussed, and conclusions and future work – including computation modeling – are discussed.

## Method

Eighteen native English speaking adults, nine male and nine female, ages 20 to 62 years of age (mean = 29.1) were paid to participate in the study. Visual stimuli were presented on an LCD display with Dual 2GHz PowerMac G5 running OS X 10.4.7. Eye movements were recorded using an LC Technologies Eyegaze System dual-camera configuration, a 120 Hz pupil-center / corneal-reflection eye tracker. A chinrest was used to maintain a constant eye-to-screen distance.

Figure 2 shows a layout from one trial. A total of six structured layouts were used. Three variables were manipulated in the layouts: the semantic cohesion of groups of words, the presence of group labels, and the use of background color. Groups of words were semantically related (e.g. cashew, peanut, almond, …) or randomly grouped (e.g. attic, dodo, polyester, ….) Groups were labeled (e.g. nuts) or not. Background color divided the groups into four common regions or not. When the common regions were used in a semantically-grouped layout, groups in the same region were further semantically related (e.g. clothing and cosmetics). The words in all layouts were selected from a hierarchical list of words based on categories used in a study of word category norms [9].

Each trial of the experiment proceeded as follows: The participant studied the precue (i.e. the target word); clicked on the precue to make the precue disappear; found the target word; moved the cursor to the target word; and clicked on it. The trials were blocked by experimental condition. Each block contained 40 trials, preceded by 5 practice trials. The blocks were countered balanced with a Latin Square. Only correct trials are analyzed.

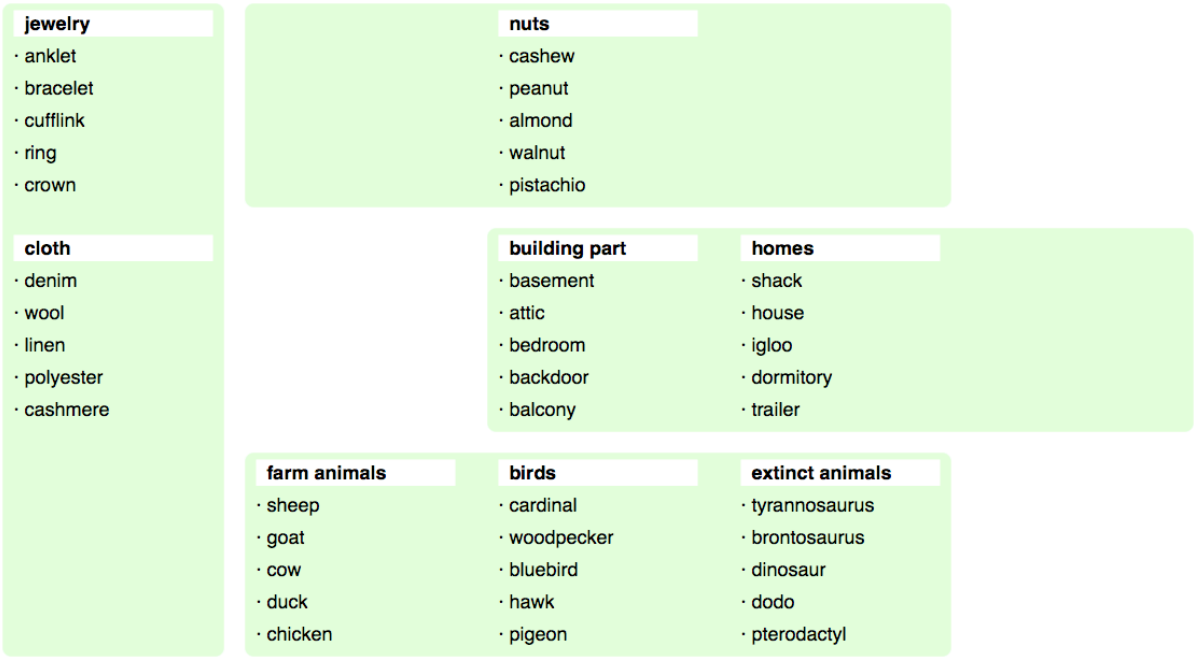


**Figure 2.** An example layout with semantically cohesive groups, group labels and background color.

## Results

Search time began when the participant clicked on the precue and ended when the participant started moving the mouse [4]. Eye movement data were included starting from the first fixation that started after the precue was clicked and first fixation that stopped before the mouse started moving.

Error rates were fairly low, ranging from 10-13% across all conditions. Error rates for the trials in which labels were absent varied by less than 0.4%. Error rates for the trial in which labels appeared were 2% higher. However, this increase in error appears to be due at least in part to participants incorrectly clicking on the group labels, which were never valid targets. We do not believe that observed differences in reaction time and eye movements are a result of a speed-accuracy tradeoff.

Search time and eye movement data were analyzed using mixed-model ANOVAs and contrast analysis. The type of layout affected all dependent variables:

search time, $F(3,1497) = 46.39$, $p < .0001$
fixations per trial, $F(3,1496) = 32.77$, $p < .0001$
fixation duration, $F(3,621) = 5.76$, $p = .0007$
saccade distance, $F(3,1477) = 4.44$, $p = .0041$

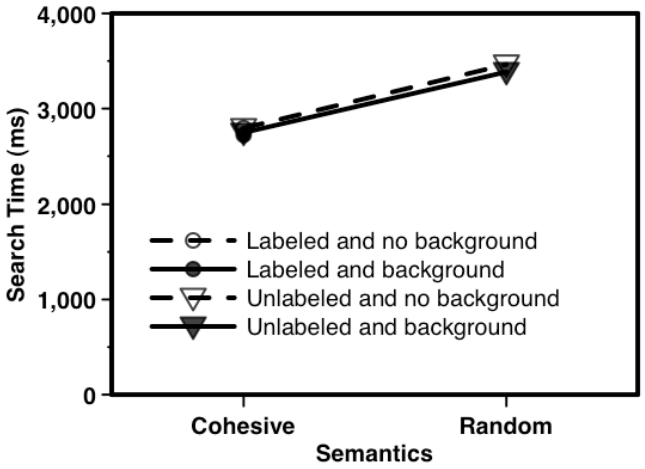


**Figure 3.** Mean search times. The means for the labeled layout are nearly identical to those of the unlabeled. The standard error is too small for the error bars to be seen.

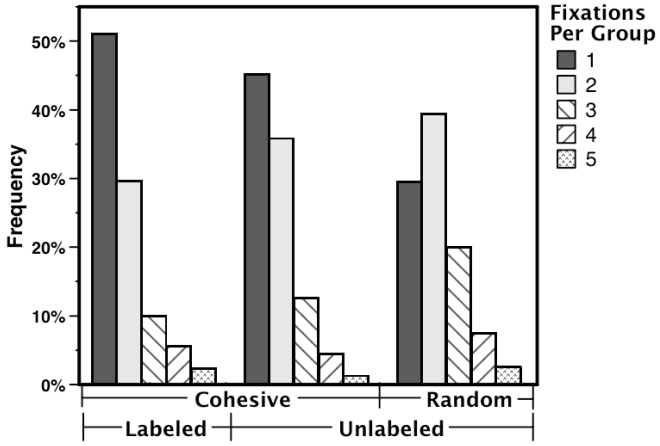


**Figure 4**. Frequency of the number of fixations per group visit. Notice how one fixation is often enough for a cohesive group, especially in the labeled layouts. Whereas, two fixations are typically needed for a random group. Data from layouts without background colors are shown. The

However, for all measures and across all layouts, only the effect of semantic grouping was significant. Figure 3 shows the search time results. Participants tended to take less time to find the targets when the layouts were semantically organized, $t(1474) = 10.06$, $p < .0001$. This was largely because the semantically organized layouts required fewer fixations to find the target, $t(1474) = 9.06$, $p < .0001$. Participants also tended to make longer saccades, $t(1413) = 2.88$, $p = .004$, and use shorter fixations, $t(207) = -2.25$, $p = .0253$, when the layouts were semantically organized. The use of background colors and group labels had no significant effect on any of the measures, all $p > .0988$.

A qualitative analysis is given based on the fixations-per-group results shown in Figure 4. Fixations per group measures the number of contiguous fixations on a group. As is shown in Figure 4, the participants tended to use just one fixation per group when the groups were semantically cohesive and two fixations when randomly organized. Within the cohesive groups, participants were more likely to use one fixation when labels were present. Participants were also more likely to make four or more fixations per group when the groups were labeled.

## Discussion

This study investigates the effects of semantic content and visual indicators of semantic relations on visual search. The data strongly support the expected result that people capitalize on the meaningful structure provided by the semantic content of the words in the layout to guide their visual search. The novel results

are what the data suggest about how visual indicators of semantic relations are used in peoples visual search strategies.

People search layouts faster when the groups are semantically cohesive. This is not surprising considering that in the semantically cohesive layouts, the meaning of non-targets provide strong cues about the target location, and no similar information is provided in the random layouts. As seen in Figure 4, people are more likely to make just a single fixation to a group in the cohesive layouts. This suggests that people tend to judge the semantic relevance of all objects in a group with that one fixation. This allows the participants to "explore" more of the layout per fixation and thus reduces the number of fixations required to find the target. Conversely, without the semantic content, it is more difficult or impossible to discount an entire group of objects with just one fixation. Both the fixations per trial and saccade distances also support this conclusion.

While the semantic content seems to provide useful information, a first pass of the data suggests that the group labels provided no additional useful information. The data show no significant difference in any measure as a function of label presence. This null result would seem to contradict previous research that showed the importance of group labels in users' visual search strategies [5]. This previous finding was supported by results from a task in which no useful semantic information was involved in the search. So, can we conclude from this study that labels are not useful when layouts are semantically cohesive? While the labels did not affect the search time or the total number of fixations needed to find a target, the real story is in the detail of the eye movements.

Further analysis of the participants' eye movements supports previous claims of the importance of group labels in visual search strategies and extends this role to non-label words in semantically organized layouts. As shown in Figure 4, when the groups were unlabeled *and* cohesive, people behave much more like they do when searching labeled groups. The participants tended to make just one fixation, presumably evaluating all words in the group based on the words processed in that one fixation. One way to interpret these results are that people were using any word in unlabeled and cohesive groups as the label for that group.

The eye movement data also differentiate the use of labels and non-labels as semantic indicators. While the semantic grouping had more of an effect than the labels, if we look at the distributions in Figure 4, we can see that people were more likely to use one fixation per group in the semantically organized layouts when the groups were labeled. People were also more likely to make four or more fixations per group when the groups were labeled. It appears as if people had more "trust" in the group labels. That is, people were more likely to discount the contents of groups based on the group label, thus more one-fixation group visits, and more committed to searching a group when they believed the target to be in a group based on the label, thus more four-or-greater fixation group visits.

## Future Work

This research is in progress. Additional work is needed in at least two areas. First, analogous to varying the distractor-goal similarity in Brumby and Howes [1], an experiment in which the cohesiveness of the words grouped together is systematically varied should be conducted. This will allow a clearer understanding of

when people use group labels or, in the absence of labels, when people discount an entire group based on the information from one fixation.

Second, computational cognitive models are being built, using the EPIC cognitive architecture [6], to better understand the strategies people used in this task. Additionally, EPIC has been extended to include retrieval of semantic similarity estimates using the Measures of Semantic Relatedness (MSR) server [8].

The modeling will help to answer remaining questions. Previous modeling of a different task [2] suggests that people continue to scan a group with a probability inversely proportional to the number of items already scanned in the group. Such a strategy may not account for the current results in which people were most likely to abandon searching a group after only one fixation.

## Conclusion

This research investigates the effects of semantically grouping visual layouts. This work in progress has already found useful results that give us some insight into the strategies users employ when visually searching user interfaces, such as hierarchically organized web pages and application interfaces. These insights into users' use of hierarchically-organized semantic information in visual search will be useful for expanding current computational models of user behavior and informing theory for HCI.

## Acknowledgements

This research was supported by the Office of Naval Research and the National Science Foundation.

## References

[1] Brumby, D. P., & Howes, A. (2004). Good Enough but I'll Just Check: Web-page Search as Attentional Refocusing. In *Proc. ICCM 2004*, Lawrence Erlbaum Associates.

[2] Budiu, R., Pirolli, P., & Fleetwood, M. (2006). Navigation in Degree of Interest Trees. In *Proc. AVI*.

[3] Fu, W.-T., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the world wide web. *Human-Computer Interaction*, 22(4), 355 - 412.

[4] Hornof, A. J. (2001). Visual search and mouse pointing in labeled versus unlabeled two-dimensional visual hierarchies. *ACM Transactions on Computer-Human Interaction*, 8(3), 171-197.

[5] Hornof, A. J. (2004). Cognitive Strategies for the Visual Search of Hierarchical Computer Displays. *Human-Computer Interaction*, 19(3), 183-223.

[6] Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4), 391-438.

[7] McDonald, J. E., Stone, J. D., & Liebelt, L. S. (1983). Searching for items in menus: The effects of organization and type of target. In *Proc. HFES 1983*.

[8] Veksler, V. D., Grintsvayg, A., Lindsey, R., & Gray, W. D. (2007). A proxy for all your semantic needs. In *Proc. CogSci 2007*.

[9] Yoon, C., et al. (2004). 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(3), 379-393.