Cognitive Strategies and Eye Movements for Searching Hierarchical Computer Displays

Anthony J. Hornof and Tim Halverson
Department of Computer and Information Science
University of Oregon
Eugene, OR 97403 USA
+1 541 346 1372
{hornof, thalvers}@cs.uoregon.edu

ABSTRACT

This research investigates the cognitive strategies and eye movements that people use to search for a known item in a hierarchical computer display. Computational cognitive models were built to simulate the visual-perceptual and oculomotor processing required to search hierarchical and nonhierarchical displays. Eye movement data were collected and compared on over a dozen measures with the *a priori* predictions of the models. Though it is well accepted that hierarchical layouts are easier to search than nonhierarchical layouts, the underlying cognitive basis for this design heuristic has not yet been established. This work combines cognitive modeling and eye tracking to explain this and numerous other visual design guidelines. This research also demonstrates the power of cognitive modeling for predicting, explaining, and interpreting eye movement data, and how to use eye tracking data to confirm and disconfirm modeling details.

Categories and subject descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces -- Evaluation/methodology, eye tracking, graphical user interfaces (GUI), screen design, theory and methods; H.1.2 [Models and Principles]: User/Machine Systems; I.2.0 [Artificial Intelligence]: General – Cognitive simulation; I.6.4 [Simulation and Modeling]: Validation and Analysis; General Terms: Design, Human Factors, Measurement, Verification; Keywords: Cognitive modeling, cognitive strategies, EPIC, eye movements, eye tracking, hierarchical menus, screen design, visual search.

INTRODUCTION

Many computer tasks consist of visually navigating through screens, searching displays of data, and finding things. A Google.com web search, for example, is typically followed by a human visual search of the resulting hits. Due to the high bandwidth of human visual perception and the ever-increasing bandwidth of visual display devices, human-computer *visual* interaction—in which people execute computer tasks by simply moving

their eyes—will likely increase in importance and relevance in human-computer interaction (HCI).

To facilitate a more efficient visual navigation and search, some web pages and screen layouts are organized into clearly labeled visual hierarchies, analogous to the headings and subheadings in this article. Designers recommend using a visual hierarchy to "establish a plan for moving the attention to some interesting portion of the display for a more detailed reading" [10, p.94]. Figure 1 shows a portion of the ACM SIGCHI home page that incorporates a useful visual hierarchy.

More About SIGCHI	**General Information**
Membership	More About SIGCHI
Involvement	Membership
Documents, Policies	Involvement
CHI Awards	Documents, Policies
	CHI Awards
Officers & Committees	**SIGCHI People**
Mailing Lists	Officers & Committees
Local SIGs	Mailing Lists
Photo History of SIGCHI	Local SIGs
	Photo History of SIGCHI
Conferences	**HCI Information**
Publications	Conferences
HCI-Sites	Publications
HCI Bibliography	HCI-Sites
	HCI Bibliography
Accessibility	**Special Interest Areas**
Education	Accessibility
Intercultural Issues	Education
Kids and Computers	Intercultural Issues
World Wide Web	Kids and Computers
	World Wide Web

Figure 1. It seems to be harder to find the link to "HCI Bibliography" in the layout on the left. The visual hierarchy on the right seems to help. (Adapted from www.acm.org/sigchi/.)

There are a number of reasons that a well-organized layout makes it easier to find things. This paper specifically examines the performance benefits at the lowest levels of processing, in the cognitive strategies that plan and coordinate the very eye movements that occur during the search. Though visual search has been studied at great length [see 1, for example], and the benefits of organizing with a visual hierarchy have been known for decades, the

effect of adding a useful visual hierarchy has not yet been scientifically examined or explained.

Two subfields of HCI are combined in this investigation: Eye tracking and cognitive modeling. Eye tracking is used to identify the patterns of visual activity that people exhibit when searching layouts with and without a useful hierarchy. Cognitive modeling is used to predict and explain the visual activity.

Eye tracking holds great promise as an observational technique for understanding how people interact with computers, but this potential is not yet fully realized for a number of reasons. Aside from the labor-intensive challenges of collecting eye movement data and preparing them for analysis, there is the challenge of interpreting the data in a meaningful way [8].

Cognitive modeling, the practice of building computer programs that behave in some way like humans, is useful to HCI in part because it can reveal patterns of human performance at a level of detail not otherwise available (as in [3]). This paper demonstrates that cognitive modeling can also be very useful to HCI because it provides a theoretical framework for explaining and predicting eye movement data. Eye tracking and cognitive modeling have much to offer each other. Cognitive models can be used to help interpret eye movement data, such as to identify the problem-solving strategy a student uses to solve math problems [11]. Eye movement data can guide the development of cognitive models, such as models of menu search [2].

This paper utilizes eye tracking and cognitive modeling to examine the core perceptual, cognitive, and motor processing involved in searching a visual layout, and how this processing changes when a useful hierarchy is introduced. The task specifically avoids reading and semantic processing. The task is analogous to a real-world task in which a user knows the exact content of what they are looking for, and the precise heading under which it will fall, but has not memorized the locations of items in the layout. Clearly, this is sometimes but not always the case.

The visual search task discussed here was first presented to sixteen participants and eye tracking data was *not* collected. The results, discussed in Hornof [5], demonstrate that people can search a visual layout more quickly when it is organized with a clearly labeled visual hierarchy. Computational cognitive models, discussed in this paper and in more detail in Hornof [4], demonstrate that people use a fundamentally different strategy when a useful hierarchy is present.

The study lends itself particularly well to eye tracking for two reasons. First, because significant and interesting phenomena have been identified using easier-to-use dependent measure of search time. Second, because the eye tracking data can be used to evaluate established theory, namely the models. The question is not simply an open-ended "How do people search a visual hierarchy?" but rather "Do people search hierarchical and nonhierarchical layouts as predicted by the cognitive models?"

THE VISUAL SEARCH EXPERIMENT

The visual search task studied here is to find a known target in a hierarchically organized visual layout. Layout items are grouped, and sometimes the groups have useful headings.

Experimental Procedure

Figure 2 shows a sample layout from the experiment. This layout has six groups of items, and each group is *labeled* with a heading of X*n*X, where *n* is a single numerical digit. In the figure, the groups are annotated with the letters A through F, though these letters did not appear in the experiment.
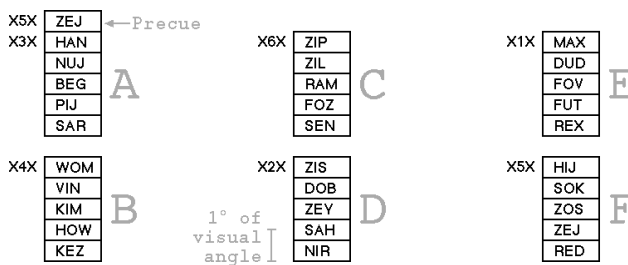

Figure 2. A *6-group labeled* layout. The precue, in the top left, would have disappeared when the layout appeared. The target is in group F. The gray text did not appear during the experiment.

Participants searched eight different screen layouts for a precued target object. Each layout contained one, two, four, or six groups. Each group contained five objects. The groups always appeared at the same physical locations on the screen. One-group layouts used group A. Two-group layouts used groups A and B. Four-group layouts used groups A through D. In each trial, the entire layout was displayed at the same moment, permitting any search order.

Layouts were either labeled or unlabeled. In unlabeled layouts, the X*n*X group labels did not appear. Each unique layout (such as *6-group labeled*) was presented in a separate block of trials.

Target and distractor items were three-letter words or pseudo-words, randomly selected for each trial. Group labels were randomly reordered for each trial. The target position was randomly selected for each trial. Participants were precued with the target object and, for labeled layouts, the label of the group that would contain the target.

Each trial proceeded as follows: The participant studied and clicked on the precue; the precue disappeared and the layout appeared; the participant found the target, moved the mouse to the target, and clicked on the target; the layout disappeared and the next precue appeared.

Search time was separated from mouse movement time by using a *point-completion deadline* [5]. The experiment is an exact reproduction of the experiment described in [5], except that eye tracking data were collected.

Sixteen participants completed the experiment. All were experienced computer users. Five were female. The mean age was 23. Eight participants were screened out for one of

the following reasons: calibration difficulties (*n*=4); failed stereopsis exam (*n*=2); too many trials interrupted by the point-completion deadline (*n*=2). Participants were financially motivated to perform each trial as quickly as possible while maintaining a very low error rate.

Eye Tracking Procedure

Eye movements were recorded using the LC Technologies Eyegaze System, a 60 Hz eye tracker that tracks eye movements using the pupil-center and corneal-reflection. The experimental stimuli were displayed using an Apple Power Mac G4 computer running Mac OS 9.2 at 733 Mhz and a ViewSonic VE170 LCD display with a 0.264 mm pixel pitch. A chinrest maintained an eye-to-screen distance of 56 cm, such that 1° of visual angle subtended 38.4 pixels. Screen objects were 25 pixels high. Columns were 300 pixels center-to-center. The precue always appeared at the same location, level with the participant's eyes.

Fixations were identified using a dispersion-based algorithm, with a minimum fixation duration of 100 ms and a deviation threshold of 0.5° of visual angle in radius. Systematic error in the eye tracking data was reduced using the *required fixation location* technique introduced in Hornof and Halverson [6], which also provides additional details of the experiment.[1]

OBSERVED SEARCH TIMES

Figure 3 shows the search times observed when the experiment was run with eye tracking. There were no meaningful differences in the search time without or with eye tracking. It appears as if the presence of the eye tracker did not substantially affect search performance.
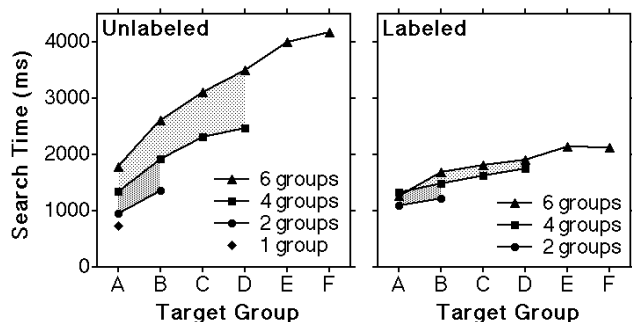

Figure 3. Mean observed search time as a function of the group containing the target, for unlabeled layouts (left) and labeled layouts (right). The shaded area shows the number-of-groups effect. Groups A-F are defined in Figure 2.

The three most salient trends in the search time data are: (a) Smaller layouts were faster. (b) Labeled layouts were faster. (c) Unlabeled layouts had a larger number-of-groups effect. The number-of-groups effect measures how much longer it takes to find an item in the same group as the

---

[1] The eye tracking data and a tool for viewing the data are downloadable at http://www.cs.uoregon.edu/~hornof.

layout gets bigger, and suggests an element of noise or randomness in the search process [5].

DESCRIPTION OF THE COGNITIVE MODELS

A number of computational cognitive models were built, using the EPIC cognitive architecture (Executive Process-Interactive Control, [9]). EPIC captures human perceptual, cognitive, and motor processing constraints in a computational framework that is used to build simulations of human information processing and task execution. EPIC constrains the models that can be built, and the predictions that can be made, based on fundamental human processing capabilities and limitations.

As is required to use the architecture, we encoded into EPIC (a) a reproduction of the task environment, (b) the visual-perceptual features associated with each of the screen objects and (c) the cognitive strategies that guide the visual search. These components were specified based on task analysis, human performance capabilities, previous visual search models, and parsimony.

After these components were encoded into the architecture, EPIC executed the task, simulated the perceptual-motor processing and interactions, and generated search time predictions for every target position in every screen layout. The models are described in more detail in Hornof [4].

EPIC simulates oculomotor processing, including the fast ballistic eye movements known as *saccades*, as well as the fixations during which the eyes are stationary and information is perceived. The models assume a foveal region 2° of visual angle in diameter which allows the model to perceive, with a single fixation, the fixated screen object as well as one or two adjacent objects. All foveated objects move to visual working memory in parallel.

The cognitive strategy, encoded as production rules, is a core component of each model. The strategy represents a plausible explanation of how people recruit and coordinate their perceptual-motor processes to accomplish the task. A modeling project typically involves searching through a space of plausible cognitive strategies, eventually finding the strategy that best explains the data.

Eight different strategies were written to examine how people searched unlabeled and labeled layouts. Each strategy was encoded into EPIC, which executed the strategy and generated predictions that were compared to the observed data. Two strategies that provide a good fit with the search time data are described here.

Noisy-Systematic Search Strategy

The noisy-systematic search strategy for unlabeled layouts assumes that people attempt to make a *maximally-efficient foveal sweep* [7], in which the eyes move to capture everything in the high resolution foveal vision with as few fixations as possible.

Noise is introduced into the strategy by having it sometimes overestimate how far the eyes can move and still foveate everything with successive fixations. If the target is missed, another sweep is required, substantially increasing the search time for that trial.

To vary the noise in the strategy, it was run with eighty-four different fixation distributions. In the model evaluated below, the first fixation is on the first or second item in group A. Subsequent fixations are made to a randomly chosen item 3 to 7 items down. The "down" direction assumes people searched down the first column, down the second, down the third, and back to the first. This order is suggested by the slope in the search time data.

Mostly-Systematic Two-Tiered Search Strategy

The mostly-systematic two-tiered search strategy for labeled layouts assumes that people search the group labels until they find the target group, and then search within that group. The strategy was based on task analysis and the significantly faster search times for labeled layouts. It is "mostly" systematic because it searches the labels in the A-B-C-D-E-F order 75% of the time, and in random order 25% of the time.

PREDICTED SEARCH TIMES

Figure 4 shows the search time predictions. The models explain the search time data rather well, predicting unlabeled layout search time with an average absolute error (AAE) of 8%, and labeled layout search time with an AAE of 6%. The only discrepancy that can be seen is a divergence between the predicted and observed data when the target is in groups C through F in a *6-group unlabeled* layout.

OBSERVED AND PREDICTED EYE MOVEMENTS

This section will present a detailed analysis of how the participants and the models searched the layouts. It will compare unlabeled and labeled search. It will evaluate how well the models predict the observed data.
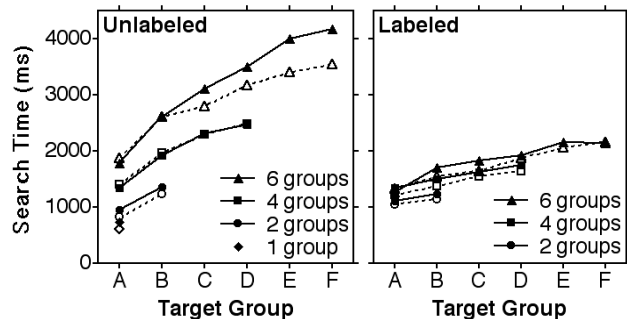

Figure 4. Search times observed (solid lines) and predicted (dashed lines) by the noisy-systematic model for unlabeled layouts (left) and the mostly-systematic two-tiered search model for labeled layouts (right).

Figure 5 shows the predicted and observed eye movements from one trial with an unlabeled layout, and one trial with a labeled layout. There were many variations in the observed data, but these are somewhat typical trials. The figure gives an idea of the similarities and differences between (a) the observed and the predicted and (b) unlabeled search and labeled search.

Table 1 summarizes comparisons between the observed and predicted eye movements. The comparisons will be elaborated in this section. It should be noted that the models were built and the predictions were generated before eye movement data were collected. The models were in no way modified to fit the eye movement data.
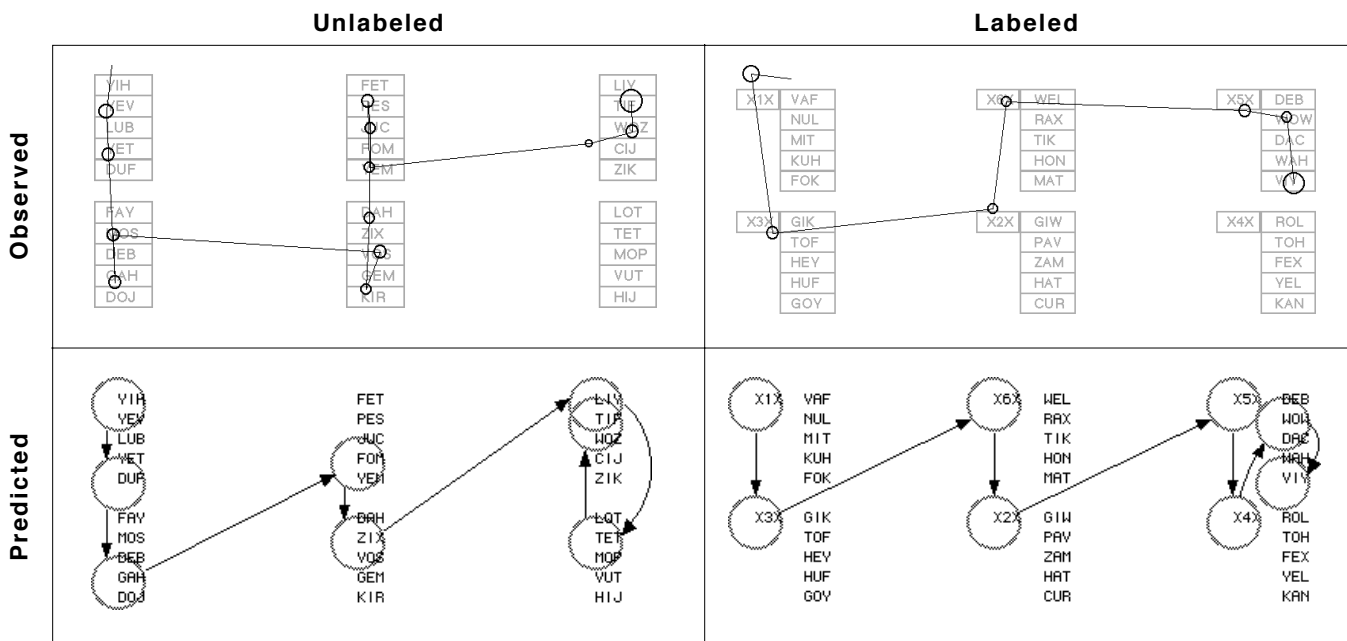

Figure 5. Fixations observed (top) and predicted (bottom) from one trial with an unlabeled layout (left) and one trial with a labeled layout (right). In the observed, the diameters of the circles represent the fixation duration. In the predicted, the circles represents the foveal region. The unlabeled layout fixations are predicted by the noisy-systematic search strategy. The labeled layout fixations are predicted by the two-tiered systematic search strategy.

Table 1. A summary of the observed and predicted eye movements. Plus signs indicate correct predictions.

Eye Movements	Observed	Predicted
Across All Layouts		
Fixations per trial (+)	7.4	7.9
Fixation duration (+)	264 ms	228 ms
Number of scan paths	Many	One
Anticipatory fixations (+)	Yes	Yes
Respond to layout onset (+)	Yes	Yes
Ignore white space (+)	Yes	Yes
Ignore text shape (+)	Yes	Yes
Overshoot the target	Rarely	Yes
Unlabeled Layouts		
Fixations per group	2.1	1.1
Groups revisited per trial	0.69	4.4
Items examined per fixation (+)	2.4	2.6
Labeled Layouts		
Use group labels (+)	Yes	Yes
Groups revisited per trial	0.29	1.2

All Layouts

Fixations Per Trial

Figure 6 shows the number of fixations per trial that were observed and predicted. There is a high correlation between fixations per trial and search time, both in the observed data ($r^2 = 0.89$) and in the models ($r^2 = 0.99$).

As can be seen in Figure 6, the models and the participants made a similar number of fixations per trial. The model overestimates the number of fixations per trial for unlabeled layouts, with an AAE of 18.0%. The model predicts an additional 1.1 fixations per trial, perhaps due to overshooting the target, discussed later. If 1.1 fixations are removed from each trial, the AAE drops to 5.4%. The model accurately predicts the number of fixations per trial for labeled layouts, with an AAE of 5.1%.
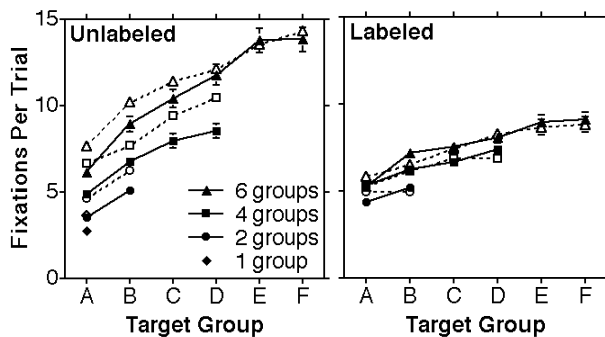
Figure 6. The average number of fixations observed (solid lines) and predicted (dashed lines) for each trial, as a function of the target group.
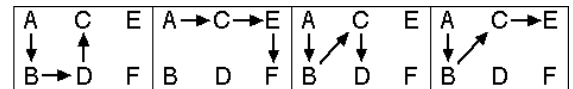
Fixation Duration

The average fixation duration exhibited by participants while searching was 264 ms ($SD = 146$). This excludes first and last fixations, which tended to be longer because the participant was either studying and clicking on the precue, or pointing and clicking on the target. The average fixation duration was a little longer for unlabeled layouts (283 ms, $SD = 152$) than for labeled layouts (238 ms, $SD = 134$).

While searching, the models average one fixation every 228 ms ($SD = 64$). There is a small discrepancy of 36 ms between the observed and predicted fixation times, but the two are very close.

Scan Paths

Figure 7 shows the most common orders in which the participants and the models searched the layouts. Both tended to start in the top left (group A), near the precue. People searched the groups in many different orders, but the most common search order, for both unlabeled and labeled, was to scan from group A down to B, over to D, and up to C. Scanning across the top, from A to C to E, was also rather common. This is not terribly surprising for the labeled layouts in which people jumped from group label to group label, but it is mildly surprising for the unlabeled layouts, in which the single-item gap in each column appears to have created two perceptual groups that people sometimes searched separately.

Observed

Unlabeled: 30%	12%	17%	5%
Labeled: 19%	12%	6%	7%

Predicted

Unlabeled: 0%	1%	70%	12%
Labeled: 0%	0%	75%	0%

Figure 7. The most common orders in which participants (observed) and the models (predicted) started searching six-group layouts, and how often each path was taken. Paths over 10% are in bold.

The models search the groups in the same order for most trials, moving from group A to B to C to D. This behavior was based on the assumption that participants would search down one column before proceeding to the next, an assumption that is clearly incorrect.

The models correctly predict that the search will start in group A. In 91% of all observed trials, the first fixation fell in or within 1° of visual angle of Group A. In the models, 91% of the initial fixations fell in the first group.

Anticipatory Fixations

On about half of all trials, participants made *anticipatory fixations*, which are eye movements from the precue to the layout before the layout actually appears. These eye movements occurred at roughly the same time as the

mouseclick that made the layout appear. In 48% of the trials, a fixation started ±100 ms of the mouseclick, which is too early for the eyes to have been responding to the layout onset. Fixations occurred during this interval slightly more often in labeled than in unlabeled trials (52% versus 44%).

Figure 8 shows how the destinations of anticipatory fixations were distributed across the first nine layout positions in unlabeled layouts. The distribution is almost identical to that reported by Byrne [2] for visual search of randomly-ordered menus of letters and numerical digits. The similarity suggests that some oculomotor strategic decisions will persist across a variety of search tasks.
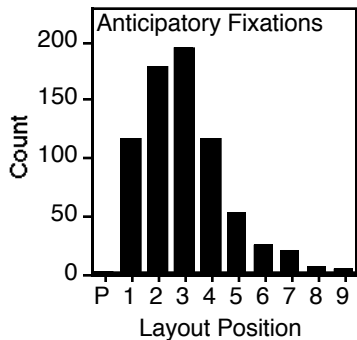

Figure 8. A histogram showing the distribution of anticipatory fixations for unlabeled layouts. "P" is the precue, and "1" through "9" are the first nine items in the layout. Fixations that did not fall directly on a screen object are excluded.

Anticipatory fixations appear to be less useful for labeled layouts. Only 67% fell within 1° of visual angle of a group label. Many fell on the precue, somewhere between the first and second group label, or in the first group, rather than on a group label. Once the layout appeared, though, participants more reliably fixated the group labels.

The EPIC models incorporate anticipatory fixations, building on previous cognitive modeling of menu search. Eye tracking data now validates this assumption, and the parameters of the fixations can be fine-tuned. Currently the noisy-systematic model for unlabeled layouts makes anticipatory fixations to positions 1 or 2 an equal number of times. The two-tiered model, for labeled layouts, makes an initial fixation to the first group label 79% of the time, and to the other group labels an equal number of times.

Response to Layout Onset

There is a strikingly high frequency of eye movements roughly 200 ms after the onset of the layout. Figure 9 shows the start time of the second fixation after the click on the precue. There are very few eye movements around 140 ms, and then a sudden burst of eye movement activity starting around 160 ms, peeking at 200 ms. It appears as if this spike of activity is in response to the appearance of the layout onset. The models similarly predict that an eye movement will occur in response to the layout onset. In the models, the movement typically occurs around 280 ms.
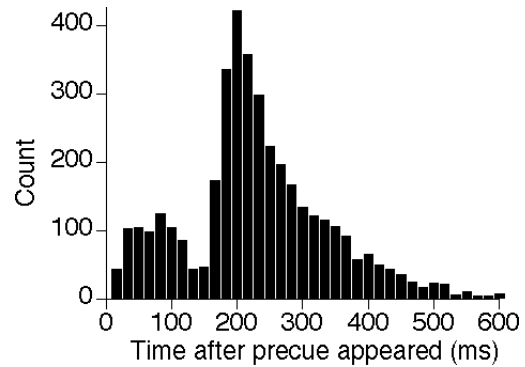

Figure 9. A histogram showing the distribution of start times for the second fixation of every trial.

Effect of White Space

Participants tended to fixate objects on the screen and to effectively ignore the white space. Across all participants and all trials, 99% of all fixations were either directly on screen objects or within 1° of visual angle of a screen object. The blank space between the columns and groups tended to be traversed with a single jump, with a long saccade that usually landed directly on the next column, but occasionally fell slightly short and was followed by another saccade along the same trajectory. Similarly, all fixations in the models are to screen objects. No fixations land on the white space between the objects.

Effect of Text Shape

Based on the eye movement data, it appears as if participants ignored the shape of text. On a trial by trial basis, we counted the average number of fixations on all distractors, and the average number on all distractors with one or two letters-in-position in common with the target (*JAX* and *SOX* have one, for example).

Across all trials, any randomly chosen distractor received an average of 0.228 fixations. Any randomly chosen distractor with one or two letters-in-common with the target received an average of 0.241 fixations. Having a letter-in-common with the target increased the likelihood of receiving a fixation by only 5.7%. Similarly, the models ignore text shape.

Target Overshoot

Participants tended to stop moving their eyes as soon as they found the target. They rarely overshot the target. We counted the number of trials in which participants fixated within 1° of visual angle from the center of the target, moved their eyes more than 1° away from the center of the target, moved their eyes back within 1° of the target, and ended the trial. People overshot the target on only 6.9% of all successful trials.

The models, on the other hand, overshoot the target on every trial. This behavior is an interesting artifact of the EPIC cognitive architecture used to build the models. The time required by the architecture to process a visual stimulus and to make its features available in visual working memory exceeds the time required to load and execute the next eye movement.

Unlabeled Layouts

Fixations Per Group
People searched unlabeled layouts going from group to group, making fixations within each group. For unlabeled layouts, there were an average of 2.1 fixations per group ($SD = 0.9$) up until the target group was reached, and then an average of 2.4 fixations within the target group ($SD = 1.0$).

The noisy-systematic model for unlabeled layouts behaves quite differently, exhibiting an average of 1.1 fixations per group ($SD = 0.4$) up until the final group is fixated, and then 1.3 fixations within the target group ($SD = 0.7$). This is one of the fundamental discrepancies between the model and the data, and is discussed later.

Group Revisits
When searching unlabeled layouts, participants usually found the target the first time they visited the target group. Participants required relatively few revisits, averaging 0.69 revisits per trial ($SD = 1.9$). Sixteen percent of the time that a participant moved to a group, it was a revisit.

The models, however, usually do not find the target on the first visit. When searching an unlabeled layout, the model revisits an already-examined group an average of 4.4 times per trial ($SD = 6.5$). Thirty-nine percent of the time the model moves to a group, it is a revisit.

Items Examined Per Fixation
Searching the unlabeled layouts, it appears as if participants examined two or three items with each fixation. This is derived from the fact that participants averaged 2.1 fixations per group, there are five items per group ($5 \div 2.1 = 2.4$), and participants typically found the target on the first visit to a group.

When searching the unlabeled layouts, the noisy-systematic model examines two or three items with each fixation because two or three items fit in EPIC's fovea. This is perhaps one of the most interesting confirmations of the model provided by the eye tracking data.

Labeled Layouts

Effect of Group Labels
For labeled layouts, participants tended to search the group labels until they found the target group, and then confine their search within the target group. This can be seen by examining the eye movements, as in Figure 5, and also in the following measures. Up until the target group was reached, 64% of all groups were visited with a single fixation, and 80% of all fixations were recorded within 1° of visual angle of a group label.[2] The models, similarly, made one fixation per group, always on the group label, up until the target group was found.

Group Revisits
When searching labeled layouts, participants usually found the target group with a single pass of the group labels. Participants averaged only 0.29 revisits per trial ($SD = 0.7$).

---

[2] One participant (P24) did not follow group labels and is excluded from the analysis.

The models, however, averaged 1.2 revisits per trial ($SD = 1.5$), roughly one extra revisit per trial because the *target overshoot* also occurred while searching group labels; after the target group label was found, the eyes typically continued to the next group, and then returned directly to search in the target group.

DISCUSSION

Searching Hierarchical Computer Displays
The eye tracking data reveal a number interesting patterns of visual search behavior which lend themselves to a number of specific design recommendations.

People use a hierarchy. When a useful hierarchy is present, people can focus their attention on one level of the hierarchy at a time. They adopt a more systematic search strategy which requires fewer revisits and fewer fixations overall. *Design recommendation:* Support a multitiered search. Do not make users infer the label by studying a few items in the group. Make labels salient. Make it easy to move from label to label with a single eye movement.

People can examine multiple items with a single fixation. *Design recommendation:* Facilitate foveal coverage of more than one item at a time. Position relevant options next to each other, in vertical lists instead of horizontal lists. Left-justify keywords.

People can navigate white space with single eye movements. Users keep their eyes on relevant screen items, and jump over stretches of white space—in this case, up to 8° of visual angle between columns—with single saccades. *Design recommendation:* Follow the advice of graphic designers and organize visual layouts with white space. This study shows how the human visual system interacts with this design element.

People anticipate visual locations. People can move their eyes well into an anticipated layout and start "searching" before the layout had even appeared. *Design recommendation:* Provide consistent screen layouts that accommodate anticipatory eye movements to known locations. Consistency is critical even at the level of visual-perceptual and oculomotor processing.

People prime for visual onset. When anticipating the appearance of a layout, it appears as if the perceptual-oculomotor system is primed and ready, waiting for the onset. *Design recommendation:* Design guidelines already recommend providing response times under 100 ms. Perhaps computer responses over 100 ms are annoying not only because they delay task completion and break the illusion of control, but because they effectively hold the perceptual-oculomotor system hostage.

Cognitive Strategies and Eye Movements
The eye tracking data confirm that many aspects of the models are correct, such as anticipatory fixations, considering more than one item with each fixation, and ignoring text shape. The data also identify a number of improvements that can be made to these and other cognitive models of visual search.

ACT-R/PM models of menu search [2] assume that 50% or 74% (depending on the model) of all fixations will be made

to text shaped like the target, far more than the 5.7% observed here.

Models need to accommodate *global* search strategies, as is done here in the strategic decision to search group labels first. Models also need to incorporate somewhat flexible and opportunistic *local* search strategies that make other decisions during the search, such as which group to visit next, and how many fixations to make in a group. The rigid global search order assumed by these models is incorrect.

The interplay between the data and the models demonstrates that noise enters the process at many different levels. Previous modeling demonstrated that the number-of-groups effect can be explained by incorporating some randomness in the model [2, 7]. The noisy-systematic model discussed here introduces one major element of noise—randomly skipped over and missing items while searching. Occasionally, participants thrashed all over the layout revisiting group after group as did the noisy-systematic model, but there were also many other sources of randomness in the human data. For example, it was common for participants to make one, two or three fixations per group, whereas the models typically made just one, and people adopted a much wider variety of scanpaths. It remains to be seen which are the most important sources to incorporate into accurate predictive models.

CONCLUSION

This research combines eye tracking and cognitive modeling to provide an explanation of the detailed perceptual and oculomotor strategies used in a number of visual search tasks, including the search of visual hierarchies. The explanations provide a scientific basis for a number of graphic design and screen layout design guidelines. The methodology of carefully comparing the data and the models on a multitude of measures provides a thorough means of validating and refining the models. The models, built before eye movement data were collected, make numerous accurate *a priori* predictions of the observed eye movement data. This project demonstrates that cognitive modeling is advancing towards routine application in a purely predictive mode.

FUTURE WORK

An integrative theory and predictive model of visual search is being built and validated from the ground up. Future work will investigate how semantic content influences eye movements during hierarchical and nonhierarchical search. The work will examine, for example, if the gaze lingers slightly longer on links that are semantically similar to the target, and what degree of similarity is needed to halt a search. Modeling work will investigate how to predict and identify semantic matches.

REFERENCES

1. Boff, K. R., & Lincoln, J. E. (1988). Visual Search, *Engineering Data Compendium: Human Perception and Performance*.: AAMRL, Wright-Patterson AFB, OH, 1549-1602.

2. Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies*, 55, 41-84.

3. Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world task performance. *Human-Computer Interaction*, 8, 237-309.

4. Hornof, A. (2002). Cognitive strategies for the visual search of hierarchical computer displays. Department of CIS Technical Report 02-04, University of Oregon, ftp://ftp.cs.uoregon.edu/pub/hornof/Hierarchical.pdf.

5. 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.

6. Hornof, A. J., & Halverson, T. (2002). Cleaning up systematic error in eye tracking data by using required fixation locations. *Behavior Research Methods, Instruments, and Computers*, 34(4), 592-604.

7. Hornof, A. J., & Kieras, D. E. (1997). Cognitive modeling reveals menu search is both random and systematic. *Proceedings of ACM CHI 97: Conference on Human Factors in Computing Systems*, New York: ACM, 107-114.

8. Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section commentary). In J. Hyona, R. Radach, & H. Deubel (Eds.), *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*. Oxford: Elsevier Science.

9. 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.

10. Mullet, K., & Sano, D. (1995). *Designing Visual Interfaces: Communication Oriented Techniques*. Englewood Cliffs, New Jersey: Prentice Hall PTR.

11. Salvucci, D. D., & Anderson, J. R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction*, 16, 39-86.