

Visual Search and Mouse-Pointing in Labeled versus Unlabeled Two-Dimensional Visual Hierarchies

ANTHONY J. HORNOF

Department of Computer and Information Science, University of Oregon

An experiment investigates (1) how the physical structure of a computer screen layout affects visual search and (2) how people select a found target object with a mouse. Two structures are examined—labeled visual hierarchies (groups of objects with one label per group) and unlabeled visual hierarchies (groups without labels). Search and selection times were separated by imposing a point-completion deadline that discouraged participants from moving the mouse until they found the target. The observed search times indicate that labeled visual hierarchies can be searched much more efficiently than unlabeled visual hierarchies, and suggest that people use a fundamentally different strategy for each of the two structures. The results have implications for screen layout design and cognitive modeling of visual search. The observed mouse-pointing times suggest that people use a slower and more accurate speed-accuracy operating characteristic to select a target with a mouse when visual distractors are present, which suggests that Fitts' law coefficients derived from standard mouse-pointing experiments may under-predict mouse-pointing times for typical human-computer interactions. The observed mouse-pointing times also demonstrate that mouse movement times for a two-dimensional pointing task can be most-accurately predicted by setting the w in Fitts' law to the width of the target along the line of approach.

Categories and Subject Descriptors: H.1.2 [Information Systems]: Models and Principles—*human factors; human information processing*; H.5.2 [Information Interfaces and Presentation]: User Interfaces—*input devices and strategies; screen design*

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: Fitts' law, mouse pointing, screen layout design, visual search

1. INTRODUCTION

To improve the speed and quality of visual communication, screen layout design guidelines recommend that screen layouts be organized with a clear visual structure that allows the user to visually navigate the layout in a

This work was supported in part by the Office of Naval Research through grant N00014-92-J-1173 to the University of Michigan, David E. Kieras and David E. Meyer, principal investigators.

Author's address: Department of Computer and Information Science, 1202 University of Oregon, Eugene, OR 97403-1202, e-mail: hornof@acm.org.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works, requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept, ACM Inc., 1515 Broadway, New York, NY 10036 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2001 ACM 1073-0516/01/0900-0171 \$5.00

systematic and somewhat predictable manner. Though it might seem to be an obvious recommendation, many screen layouts in computer interfaces and Web sites do not impose a clear and useful structure, and thus make visual tasks more difficult than necessary. The work presented here demonstrates that one particular structure—a clearly labeled visual hierarchy—will significantly reduce visual search time and also motivate a more systematic search strategy.

Visual search of a computer screen is often followed by manual selection with a mouse. In order to develop more accurate models of visual search and manual selection, detailed and precise measurements of both subtasks are required. To that end, this article introduces a new paradigm for separating visual search time from mouse-pointing time for experiments that incorporate these two subtasks. As well, this article contributes to the ongoing refinement of Fitts' law as a means of predicting mouse-pointing times by providing evidence that (1) target width is best measured along the line of approach, and (2) typical pointing experiments may underestimate Fitts' law parameters, and (3) the Shannon formulation of Fitts' law does not fit data better than the Welford version.

2. LABELED VISUAL HIERARCHIES

A visual hierarchy is an arrangement of visual elements in which related items are physically grouped together. The physical structure of the layout guides the viewer to the information that he or she needs. Screen layout design guidelines recommend the use of a visual hierarchy:

By grouping similar elements together, the designer helps the user deal with a complex information display by reducing it to a manageable number of units. Higher level structures orient the user and help them establish a plan for moving the attention to some interesting portion of the display for a more detailed reading. [Mullet and Sano 1995, p. 94]

Reorganize the elements into logical groups defined by space (proximity), give the important elements the visual prominence they deserve (contrast), and unify the various elements by aligning them. Do you find that your eye now has a clearer path to follow? [Williams and Tollet 1998, p. 120]

A visual hierarchy creates perceptual grouping [Norman 1991, p.130; Treisman 1982] and facilitates visual navigation. The content and structure of a visual hierarchy, and the contents of each group, can be made even clearer by adding clear and useful labels to groups of related items, labels (or headings) that concisely characterize the relevance of each group. A labeled visual hierarchy allows the viewer to first find the group that would likely contain the desired item, and to then search only within that group.

Figure 1 shows a layout with an unlabeled visual hierarchy. Related visual elements are grouped together, but the groups are not given headings. Figure 2 shows the same layout but with headings—a labeled visual hierarchy. It would presumably take longer to find a map of the library when the group labels are absent (as in Figure 1) than when the group headings are present (as in Figure 2).

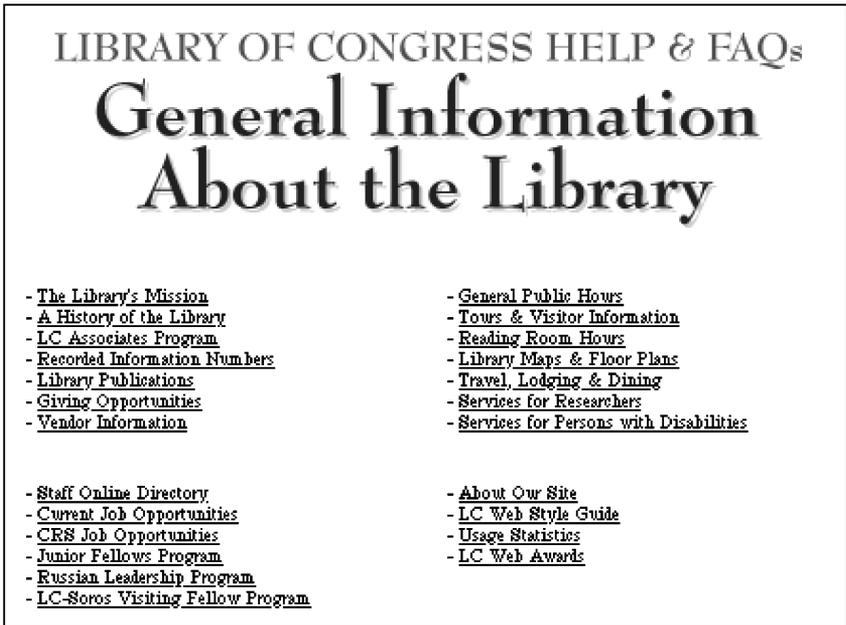


Fig. 1. A screen layout design that does not incorporate a labeled visual hierarchy, making it unnecessarily difficult and time-consuming to find a map of the library. (Adapted from www.loc.gov/help/, 8/14/00).

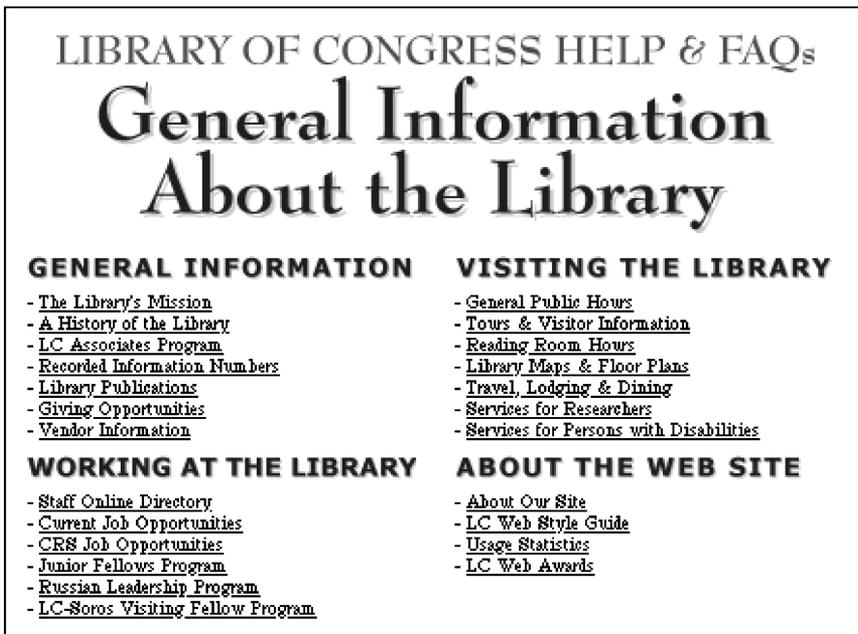


Fig. 2. A screen layout design that incorporates a clearly labeled visual hierarchy that would help a user with a search task such as finding a map of the library. (Adapted from www.loc.gov/help/, 8/14/00).

The problem can be considered in the context of *information foraging* [Pirolli and Card 1999], the process by which people search for information in online systems, moving from patch to patch (i.e., page to page) looking for information, much like birds fly from patch to patch looking for berries. In the context discussed here, each group is a patch. The *between-patch* cost of moving the eyes from one group to another is about the same with or without labels. But once the eyes arrive at a new group, the cost of evaluating the relevance or import of that group (the *within-patch* cost) can be decreased if a useful group label is present.

3. SEARCH OF VISUAL LAYOUTS

Though it might seem obvious that a labeled visual hierarchy will contribute to an efficient visual search of a screen layout, few if any studies verify or measure the effect. Reviews of screen layout design research (such as Galitz [1996] and Tullis [1997]) and of computer menu research (such as Norman [1991] and Paap et al. [1997]) do not cite studies that examine labeled versus unlabeled visual hierarchies. “Unfortunately, the empirical evidence directly relevant to the grouping of elements on a display is somewhat sparse” [Tullis 1997, p. 510]. Galitz [1996, p. 148] specifically recommends adding headings to groups of items on a screen, but does not provide empirical support for this recommendation.

Studies have examined various effects of grouping items in a screen layout. Card [1982] and McDonald et al. [1983] found that menu items grouped by category (but without group labels) can initially be searched more quickly than randomly ordered menus. Triesman [1982] found evidence that, when items are grouped, groups rather than individual items will be scanned serially for conjunctions of features. Tullis [1997] provides evidence that search time is fastest when groups are smaller than five degrees of visual angle across. Burns et al. [1986] found that visual search time of a space shuttle display was decreased by adding clear and useful group labels, but the effect of adding these labels was not studied independently of other visual layout improvements. Previous research shows the value of groups, but not of adding labels to those groups.

Computer menu research has studied the depth versus the breadth of hierarchical menu systems (see reviews by Norman [1991] and Paap et al. [1997]). But these are systems that display one menu on the screen at a time. The user searches the menus by manually flipping through screens, such as on an automated teller machine. This is a fundamentally different activity than scanning a single hierarchically-organized visual layout—the *between-patch* cost of changing a screen is much higher than the cost of moving the eyes to a new location.

Some visual layout guidelines advise against putting too much information on a single screen, explaining that “visual clutter” will make it too difficult for people to find what they need. Galitz [1996], for example, recommends that no more than 30% of a screen layout should be used to display information, and that the remaining 70% should remain blank. A number of experimental

studies (such as Thacker [1986] and Treisman [1982]) seem to support such a recommendation by describing cases in which increasing screen density—the proportion of the screen displaying information—increased search time. In actual practice, however, this is not always the case. For a hospital information system, Staggers [1993] found quite the opposite to be true—putting everything on a single screen improved both user performance and user satisfaction. Similarly, Burns [2000] observed that faults in a simulated power-plant control system were diagnosed more quickly and accurately when plant information was presented in a dense and spatially integrated format. In both cases, the increased density of the display obviated flipping through multiple screens to accomplish a task, and the displays were dense but well-organized, and thus easy to navigate visually.

A number of computational tools have been built to predict how long people will take to find things on a screen layout, but few predict that a labeled hierarchy decreases search time. Tullis' [1988] Display Analysis Program (DAP) predicts search times based on grouping, density, and layout complexity, but ignores visual semantic cues such as how labels can guide the viewer to the target. The screen layouts shown in Figures 1 and 2 were converted into the alphanumeric input required by DAP and evaluated by the program. DAP predicts a search time of 3800 msec for Figure 1 (unlabeled) and 3830 msec for Figure 2 (labeled). DAP does not predict that a labeled visual hierarchy speeds visual search.

In order to predict that a labeled visual hierarchy will speed search, predictive theories need to incorporate the strategies that people use to navigate visual layouts. Lohse's [1993] Understanding Cognitive Information Engineering (UCIE) is an example of a predictive tool that incorporates visual strategies that take advantage of the structure of a visual layout. UCIE accurately predicts the time required to glean information from line graphs, bar graphs, and tables. For such tools to be extended to predict aspects of visual search in visual hierarchies such as those shown in Figures 1 and 2, more data is needed to better understand the search strategies that people would use for such tasks, and to what extent a labeled visual hierarchy will speed the visual search.

4. IDENTIFYING VISUAL SEARCH STRATEGIES

Building computational models of the perceptual, cognitive, and motor processes involved in visual search tasks is an effective way to determine the strategies that people use to accomplish visual search tasks. Through cognitive modeling, some strategies can be ruled out as unlikely and implausible, and others can be promoted as likely candidates for the strategies that people actually use.

Fundamental aspects of visual search strategies can be revealed by building cognitive models that account for the task-completion times of relatively simple tasks, such as finding and selecting an item in a pull-down menu. Nilsen [1991] conducted a menu selection experiment in which targets and distractors were distributed on a single physical dimension (up or down), there were

very few distractors (two, five, or eight), and the visual objects were extremely simple (the numerical digits 1 through 9). The simplicity of the task helped to facilitate subsequent cognitive modeling of the visual search component of the task [Anderson et al. 1997; Byrne 2001; Hornof 1999; Hornof and Kieras 1997]. These models provide insight into the strategies that participants used when searching Nilsen's randomly ordered menus.

Hornof and Kieras [Hornof 1999; Hornof and Kieras 1997] dismissed visual search strategies in which participants considered each menu item one at a time by showing that models incorporating such strategies could not predict the observed task-completion times. Based on a better fit with the data, strategies in which participants considered more than one menu item at a time appear to be more plausible explanations. Likewise, based on a poor fit with the data, the notion that participants searched exclusively randomly or systematically was dismissed. Based on a good fit with the data, a combination of random and systematic search appears to provide a more plausible explanation. These models emphasize *global control*, in which eye movements are programmed based on the structure of the layout and the task.

A model by Anderson et al. [1997] incorporates a fundamentally different visual search strategy, in which attentional shifts are motivated by the features in the menu items, and less by the structure of the menu. The model explains Nilsen's data reasonably well, and provides a specific example of how a visual search strategy can use preattentive visual processing. The model emphasizes *local control*, in which eye movements are programmed in response to visual information that becomes available during task execution. Byrne [2001] demonstrates how both selection time and eye movement data for a menu selection task can be explained by a cognitive model with a visual search strategy that incorporates elements from the two previous modeling efforts, but still emphasizes local control.

An outstanding issue is to what extent visual search of computer screen layouts is based on global versus local control. A more general outstanding question is simply "What are the visual search strategies that people use to search a visual layout?" To better answer these questions, data from more complex visual search tasks needs to be collected and modeled.

5. SEPARATING VISUAL SEARCH TIME FROM MOUSE-POINTING TIME

In experiments in which participants search for a visual target and then select it with a mouse, visual search time and mouse movement time are typically recorded as a single value. When modeling such a task, it is not entirely clear how much of the total time should be attributed to each of the two subtasks. The menu models discussed in the previous section, for example, include different assumptions regarding how much of the total selection time should be attributed to visual search, and how much to mouse movement. In the models by Hornof and Kieras [1997] and by Byrne [2001], mouse movement does not start until visual search is completed. In the model by Anderson et al. [1997], however, the mouse moves in lockstep with the eyes, and thus does not add to the total processing time. Detailed mouse movement data [Byrne et al. 1999]

favors the former of these two assumptions, but also demonstrates that it is not always correct.

The experiment presented here introduces the *point-completion deadline* as an experimental technique for separating visual search time from mouse-movement time. With the empirical data for each of the two subtasks recorded separately, it will be clear how much time was spent on each subtask, and clearer how to model each subtask accurately.

The point-completion deadline works as follows: Once the mouse starts to move, the participant is given a limited amount of time to click on the target; if the participant does not click the mouse button before this point-completion deadline is reached, then the trial is interrupted and recorded as an error. Thus, participants are discouraged from moving the mouse until they have found the target, and the start of the mouse movement consistently marks the end of visual search.

To allow more time for more distant targets, the point-completion deadline in this experiment was computed using Welford's [1968] version of Fitts' law:

$$\text{Movement time} = a + b \log_2(d/w + 0.5) \text{ msec}, \quad (1)$$

where d is the distance to the target and w is the width of the target. The width was computed along the direct line of travel from the precue, as discussed by MacKenzie [1992]. Empirical studies of mouse-pointing times typically result in a and b values ranging from about 100 to 300 [Card et al. 1978; Epps 1986; Gillan et al. 1990; Han et al. 1990; Walker et al. 1993]. Hence, a and b were both set to 300, assumed to be large enough to accommodate the slower participants, but small enough to provide enough time for only a single aimed movement.

The Welford version of Fitts law is used here because it is the one originally used to predict mouse-movement times [Card et al. 1978] and because it has not been substantially improved. The "Shannon formulation" of Fitts' law [MacKenzie 1992; MacKenzie and Buxton 1992] has been introduced since, but this new formulation has the ill-effect of producing a and b parameters that cannot be meaningfully compared to the a and b parameters from previous mouse-movement studies that analyze data using the Welford [1968] version (such as Card et al. [1978] and Gillan et al. [1990]). There is little or no evidence that the Shannon formulation fits data better than the Welford version. MacKenzie [1991] provides an extensive comparison between predictions made by the Shannon formulation and the *original* version of Fitts' law [1954], and in only 6 out of 28 comparisons is there a significant difference between the correlation of each model to the predicted data. The correlation between the Welford version and the observed data is not provided for any of the studies, but is stated to be consistently between those of the Shannon formulation and the original version. There is little or no empirical evidence that the Shannon formulation fits data better than the Welford version, though this claim has been made (by MacKenzie [1992], for example).

The Model Human Processor [Card et al. 1983] and the cognitive architectures EPIC [Kieras and Meyer 1997] and ACT-R/PM [Anderson and Lebiere 1998; Byrne, 2001] use the Welford version to predict mouse-movement times.

EPIC and ACT-R/PM also impose a minimum predicted movement time of 100 msec, since discrete manual aimed movements of a shorter duration are rarely observed (for example, see the data in the mouse-pointing studies cited earlier and in Rosenbaum [1991]). It is not necessary to impose such a minimum time when computing the point-completion deadline, however, because a and b values of 300 insure a deadline greater than 100 msec.

Imposing a point-completion deadline prevents participants from using the cursor as a “visual anchor” as they search (as suggested by Sears and Shneiderman [1994, p. 34]); but until it is demonstrated that this technique improves search performance, there is no reason to believe that users with optimized search strategies actually use it.

6. EXPERIMENT

The experiment presented here is designed to both (a) collect visual search and mouse pointing data that can be used to extend existing visual search and mouse-pointing models to explain more complex tasks, and (b) to investigate how a labeled visual hierarchy influences visual search.

The experiment extends Nilsen’s menu-selection task to a more complex visual search task, and collects data that can be used to guide the construction of computational cognitive models that account for a wider range of visual search phenomena. Nilsen’s task is extended in three ways: (1) the targets and distractors are distributed in two dimensions on the computer screen; (2) there are many more distractors; (3) the targets and distractors are words and pseudo-words, instead of single numerical digits. One indicator of success will be simply if any of the trends seen in Nilsen’s data that helped to reveal the search strategies can also be seen in the data collected here.

The experiment tests two specific hypotheses. The first is that a labeled visual hierarchy can be searched more quickly than an unlabeled visual hierarchy. The second is that people will use a fundamentally different strategy to search a labeled visual hierarchy as opposed to an unlabeled visual hierarchy.

In the experiment, participants are presented with structured arrangements of up to 30 visual objects. Each unique arrangement is referred to as a *layout*. In all but one of the layouts, the visual objects are arranged in groups, with five objects per group. In some layouts, the groups have labels that guide the user to the group containing the target. In other layouts, there are no group labels. It is anticipated that the labeled layouts will improve search performance by allowing the participant to first find the group containing the target and then find the target within that group.

The number of groups that appear in each layout is varied. The menu-length effect in Nilsen’s [1991] data was particularly helpful in guiding the process of identifying search strategies. It is anticipated that a corresponding number-of-groups effect here would be equally helpful in identifying plausible search strategies.

Clearly, there are semantic issues in visual search [Card 1982; McDonald et al. 1983; Somberg and Picardi 1983]. When searching for a target in a visual layout, the semantic relationship between the precue and the target, and

between the target and the group labels, will affect the usefulness of the labels. Determining the best words to use for command and object names is a problem unto itself [Furnas et al. 1987]. To emphasize the fundamental mechanisms and strategies involved in visual search rather than semantic issues, this experiment assumes that participants will know the target and group label before starting their search. Semantics has specifically been removed as a confound.

7. EXPERIMENTAL METHOD

7.1 Stimuli and Apparatus

Eight different layouts were used. Each layout contained a *target* object that the participant had to find and then select with a mouse. Seven of the layouts also included *distractor* objects, which slowed down the search for the target. Layouts with distractors contained 5, 10, 20, or 30 objects. These objects were arranged in one to six groups, with five objects per group.

Layouts with distractors were either labeled or unlabeled. In *labeled* layouts, each group was assigned a heading that appeared at the top left of the group. In *unlabeled* layouts, no such headings appeared. The precue appeared above the first group of objects and showed the target object for the trial. For labeled layouts, the precue also showed the label of the group that would contain the target.

Figure 3 shows a sample layout with six labeled groups, and the number associated with each of the 30 fixed positions in which objects appeared.

Each target and distractor was a CVC (consonant-vowel-consonant) in a rectangular box. The CVC for every position in every trial was randomly chosen, with no duplication within a trial. CVCs that began with the letters *Q* or *X* were not used because they were deemed difficult to pronounce, and 10 potentially offensive CVCs were not used. All other consonant-vowel-consonant combinations were used, to normalize across word shape and letter frequency rather than meaningfulness.

Group labels were single numerical digits, flanked by *X*'s to inhibit parafoveal recognition. Group labels were randomly assigned to groups for every trial.

Visual stimuli were presented on an Apple Multiple Scan 17 Display set to 33 pixels per cm and to maximum brightness and contrast, which was driven by a 408 MHz G3 Apple Power Macintosh running Macintosh OS 8.5.1. The mouse was a new Apple Desktop Bus Mouse II. The mouse control panel was set to the third-fastest tracking setting. The mouse was placed on a mouse pad on a table 68 cm above the floor. The experimental software sampled the mouse position approximately once every 40 msec.¹

All text appeared in uppercase 18-point Helvetica font. Characters were 4-mm high, 3- to 4-mm wide, and with a 0.8-mm stroke width, in rectangles that were 8-mm high and 21-mm wide. At the eye-to-screen distance of 56 cm

¹An executable copy of the software configured as it was for the experiment is available via anonymous ftp at <ftp://ftp.cs.uoregon.edu/pub/hornof/>. The software can be reconfigured for similar visual search experiments.

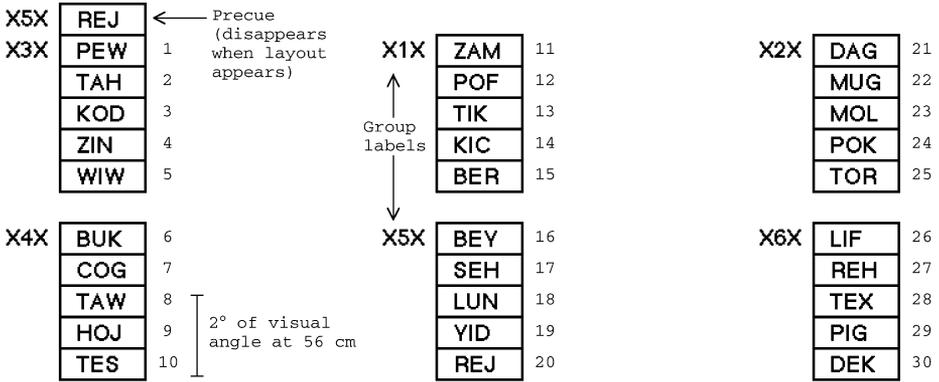


Fig. 3. A sample layout with six labeled groups, drawn to scale and annotated with details such as the number associated with each target position. The target REJ is in position 20. Two degrees of visual angle at a viewing distance of 56 cm are shown.

(22 in.) used in the experiment, character height subtended a visual angle of 0.4° and the height of the rectangles subtended a visual angle of 0.8° . There were 71-mm between the columns in the layout.

7.2 Design

The experiment was based on a 2×3 factorial design (label presence \times number of groups). Group labels were either present or absent. The layout included two, four, or six groups. Two additional layouts were also examined. The first contained one group with no label. Since adding a single label would not aid visual search, a layout with one labeled group was not examined. The second additional layout was the *target-only* layout. In this layout, the target object appeared in any of the 30 positions, with no distractors. In the target-only layout, the task was reduced to a simple pointing task. Figure 4 shows examples of the eight screen layouts used in the experiment.

Each participant completed eight blocks of trials, one block for each layout shown in Figure 4. The order of the blocks for each participant was determined by a randomized Latin square. Each block consisted of three trials for every possible target position in that layout. The target positions used in each block were randomly ordered. Any time that a participant made an error, an additional trial with the same target position was added to that block, and the remaining trials were randomly re-ordered. Every participant thus completed three error-free trials for every target position in each layout.

To separate search time from pointing time, participants were instructed to not move the mouse until they had found the target. To encourage compliance with this instruction, the experimental software imposed a point-completion deadline as discussed in Section 5. Once the participant moved the mouse more than five pixels in any direction, he or she had a limited amount of time to click on the target before the point-completion time elapsed. If the participant did not click the mouse button before the deadline was reached, then the trial

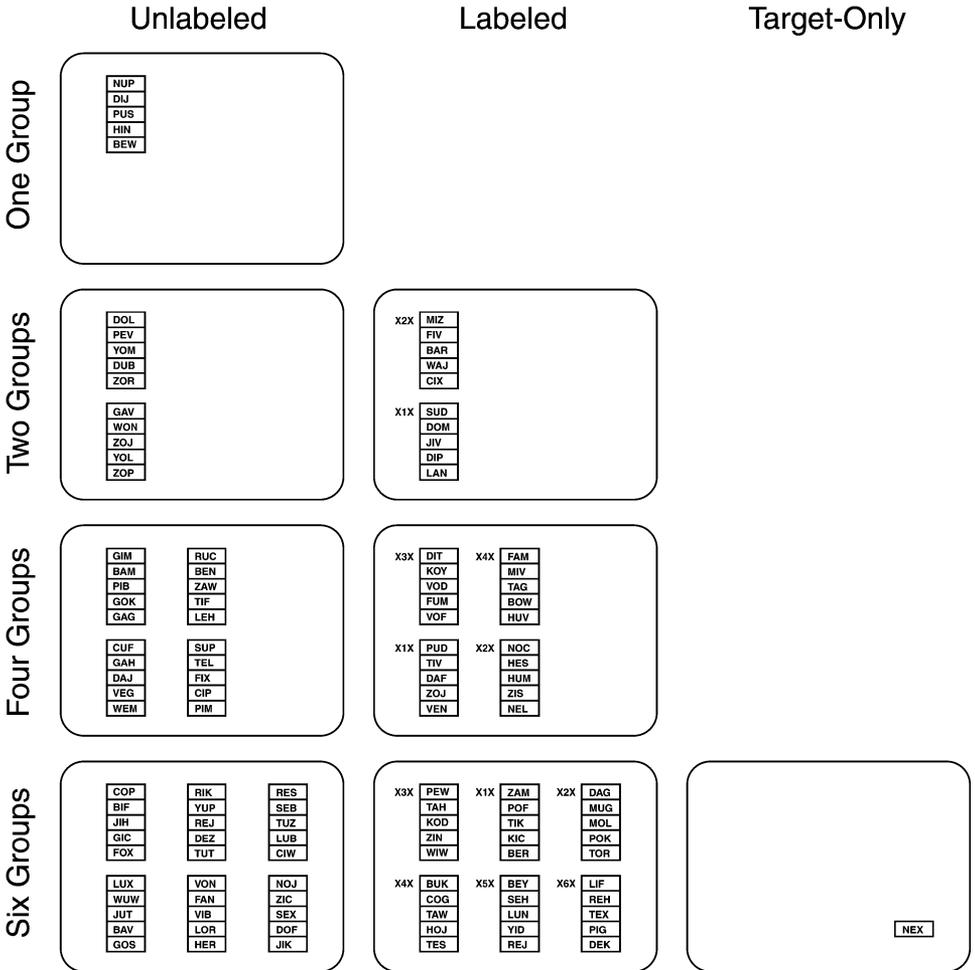


Fig. 4. Examples of the eight screen layouts used in the experiment. In the target-only layout, the target appeared in any of 30 positions. (The layouts are not drawn to scale—the horizontal distance between the columns is compressed.)

was interrupted, a buzzer sounded, an alert box appeared reminding the participant to find the target before moving the mouse, and the trial was recorded as an error.

7.3 Participants

There were 16 participants, 8 women and 8 men, ranging from 22 to 51 years of age, with a mean age of 38 years. Ten participants were recruited from a church in Ann Arbor, Michigan. Six were graduate students and support staff at the University of Michigan. All participants were experienced computer users and had no apparent visual or motor impairments. Participants received a base payment of \$10 plus a bonus (ranging from \$0 to \$7.59) based on their performance.

7.4 Procedure

Participants completed the eight blocks of trials in 30 to 50 minutes and then answered a series of questions about the experiment. The computer was fitted to the participant based on standard ergonomic recommendations, such as those distributed by Apple Computer [1999]. Participants sat in a nonreclining ergonomic chair with the seat and arm height adjusted to maintain feet flat on the floor and the forearms level with the worksurface. The room was dimly lit, and the monitor was positioned at an eye-to-screen distance of 56 cm (22 in.), and at a height so that the precue was level with the participant's eyes. A chin rest was not used because arm movements can be disrupted when the head is restricted [Biguer et al. 1985]. Instead, if a participant shifted their position relative to the monitor, which happened at most one time per participant, the experimenter re-adjusted the monitor position.

Before starting the first block of trials, participants were allowed an unlimited number of trials from the first one or two blocks to grow accustomed to the point-completion deadline. Once they indicated they were ready, the experimental software was reset and the timed trials started. At the start of each block, participants were also allowed five practice trials during which the word "Practice Trial" appeared next to the precue.

Participants were told that in some layouts, the groups would be labeled, and that the precue for these layouts would indicate both the target item as well as the label of the group that would contain the target item.

Each trial proceeded as follows: (1) the computer screen is blank except for the precue; (2) the participant moves the cursor to the precue and clicks (presses and releases) the mouse button; (3) the precue disappears and the screen layout appears; (4) the participant finds the target item; (5) the participant moves the cursor to and clicks on the target item; (6) the software provides immediate and brief feedback on the performance for that trial; (7) the precue for the next trial appears.

Performance feedback for each trial was provided as follows. If the participant successfully clicked on the target before the point-completion deadline, the target object flipped to reverse video for 170 msec and a pleasant sound of clinking coins sounded for the same duration. On all error trials, an annoying buzzer sounded for 350 msec.

Participants were instructed to find and select each target as quickly as possible, to not move the mouse until they had found the target, and to try to make no mistakes. They were informed that their bonus pay for good performance decreased at a constant slow rate within the timed portion of each trial, but that they lost everything for that trial and incurred an additional penalty if they missed the target or the trial was interrupted because they moved the mouse early.

Bonus pay was determined as follows: Each participant started each trial with a small bankroll for that trial, from four to seven cents, depending on the layout. As the timed portion of the trial passed, the bankroll diminished at a continuous rate of one cent per second until the target was selected. If the participant missed the target or missed the point-completion deadline, the bankroll

Table I. Search and Pointing Times for all Eight Layouts, Averaged Across Target Position and Participant

Layout Type	Number of Groups	Search Time		Pointing Time		Total Time	
		Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>
Unlabeled	1	583	156	448	82	1031	143
Unlabeled	2	1143	153	531	80	1674	181
Unlabeled	4	2113	343	580	78	2694	395
Unlabeled	6	3090	420	635	86	3725	443
Labeled	2	1046	173	554	81	1600	196
Labeled	4	1503	226	597	86	2101	267
Labeled	6	1740	199	661	88	2401	228
Target-only		208	27	609	81	817	85

All times are in msec; *SD* = standard deviation of the 16 participant means for each layout.

went to zero and a penalty of five cents accrued. The sum of all bankrolls and penalties was the participant's bonus pay for the experiment. The exact monetary values were adjusted slightly during the course of the experiment to keep the bonus between zero and five dollars; the values described here are the final values used.

8. RESULTS

The total time for each trial was divided into two parts: search time and pointing time. Search time began when the participant clicked on the precue and ended when the participant began moving the mouse. Pointing time began with the moving of the mouse, and ended with the click on the target.²

Table I shows the mean search time, pointing time, and total time for each layout, averaged across target position and participant, as well as the standard deviation of the 16 participant means for each layout. In the target-only layout, the target appeared equally often in all 30 positions. Hence, the target-only times should be compared to the times for the unlabeled-6-group and labeled-6-group layouts. Error trials are excluded from the timing analyses, but will be discussed later. Mean search and pointing times on a position-by-position basis are listed in the Appendices.

8.1 Search Time

As can be seen in Table I, mean search time increases as the number of groups increases, and labeled layouts are faster than unlabeled layouts. The search times for unlabeled and labeled layouts with 2, 4, or 6 groups were analyzed. A two-factor repeated measure ANOVA confirms a main effect of label presence, $F(1, 15) = 215$, $p < 0.0001$, and number of groups, $F(2, 30) = 467$, $p < 0.0001$. There is also a significant interaction between these two factors, $F(2, 30) = 117$, $p < 0.0001$.

²*Pointing-and-clicking time* is a more precise description of the second measurement, but *pointing time* is used for the sake of brevity.

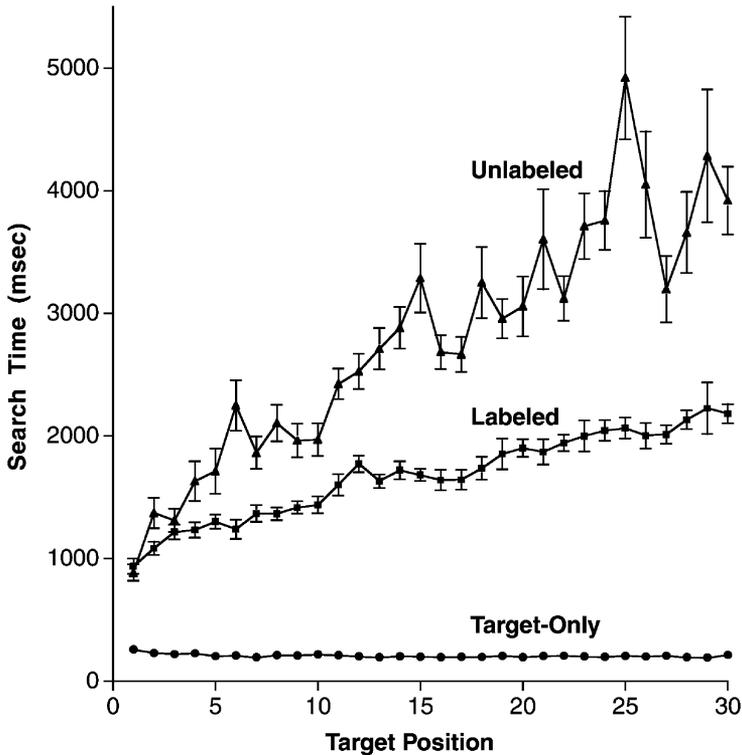


Fig. 5. Mean search time as a function of the target position, separated by layout type (unlabeled, labeled, and target-only). The standard error of each mean was computed using the standard deviation of the 16 participant means in the corresponding plot symbol. The target-only error bars are too small to appear. To facilitate a balanced comparison between the labeled and unlabeled data, search times from the one-group unlabeled layout are not included in the unlabeled means.

The target-only layout was searched much more quickly than labeled and unlabeled layouts. Target-only search times were compared to search times for the labeled-6-group layout (the faster of the two other layouts that used all 30 positions). A one-factor repeated measure ANOVA found a significant difference, $F(1, 15) = 903$, $p < 0.0001$.

8.1.1 Target Position Effect. There is a strong target position effect for both labeled and unlabeled layouts. Recall that target positions 1 through 10 are in the first column, target positions 11 through 20 are in the second column, and target positions 21 through 30 are in the third column. As can be seen in Figure 5, the search times for both labeled and unlabeled layouts appear to be a somewhat linear function of the target position. Linear functions provide a reasonably good fit with the data for both layout types. For labeled layouts, $\text{Time} = 1076 + 39 \cdot \text{Position}$, $r^2 = 0.95$. For unlabeled layouts, $\text{Time} = 1204 + 102 \cdot \text{Position}$, $r^2 = 0.87$.

The variability in the unlabeled layouts is much greater than in the labeled layouts. This is illustrated in Figure 5, both by the larger standard errors for the unlabeled layouts and the lack of smoothness in the unlabeled layout data

curve, though the increased variability for the target positions further to the right results in part because fewer layouts use those positions. Calculated using the individual trial responses, the unlabeled layout data has a coefficient of variation of 0.78, whereas the labeled layout data has a coefficient of variation of 0.44. The coefficient of variation is the standard deviation divided by the mean, and measures the variability relative to the magnitude of the data. It provides a good comparison of the relative variability in search times for the different layout types.

There is little or no target position effect for the target-only layout. The mean search times for the target-only layout is very fast (208 msec) and is relatively constant across all 30 positions. There is also little variability. The coefficient of variation, calculated using the individual trial responses, is 0.27.

8.1.2 Number-of-Groups Effect. Figure 6 shows search time as a function of the target position for each of the seven unlabeled and labeled layouts. In the two graphs in Figure 6, search times are averaged by group (i.e., positions 1–5, 6–10, etc.) in order to smooth out the noise and emphasize the *number-of-groups effect*. The number-of-groups-effect is the difference in search time to the same target position in two layouts that differ only in the number of groups in the layout. The effect is illustrated by the shaded areas in Figure 6.

The number-of-groups effect is larger for unlabeled layouts. The average number-of-groups effect for unlabeled layouts is 618 msec and for labeled layouts is 149 msec. These averages were computed by finding the mean difference in search times for positions 1–10 from the 2-group and 4-group layouts, and for positions 1–20 from the 4-group and 6-group layouts. These differences span the shaded areas in Figure 6.

To determine if the number-of-groups effect is significantly larger for unlabeled layouts, two repeated-measure ANOVAs were conducted to test for an interaction between label presence and number-of-groups. Since the various experimental conditions include different numbers of target positions, all observations could not be included in a balanced-design analysis. Hence, two balanced subsets of observations were analyzed. First, search times for positions 1–10 from the 2-group and 4-group layouts were analyzed. A two-factor (label presence versus number-of-groups) repeated-measures ANOVA confirmed a significant interaction between these two factors, $F(1, 15) = 11, p < 0.005$. Second, positions 1–20 from the 4-group and 6-group layouts were analyzed. A similar two-factor ANOVA also confirmed a significant interaction between the two factors, $F(1, 15) = 59, p < 0.0001$. These two analyses confirm that the number-of-groups effect is significantly larger for the unlabeled layouts.

An additional feature in the data that can be seen in Figure 6 is that the shape of the data plots for unlabeled layouts tends to be curvilinear, but for labeled layouts the data plots are rather straight.

8.2 Pointing Time

Figure 7 shows the distribution of pointing times to each target position across all trials, including error trials. Figure 7 also shows the point-completion deadline, which was computed using Fitts' law (see Eq. 1) with $a = 300$ and $b = 300$.

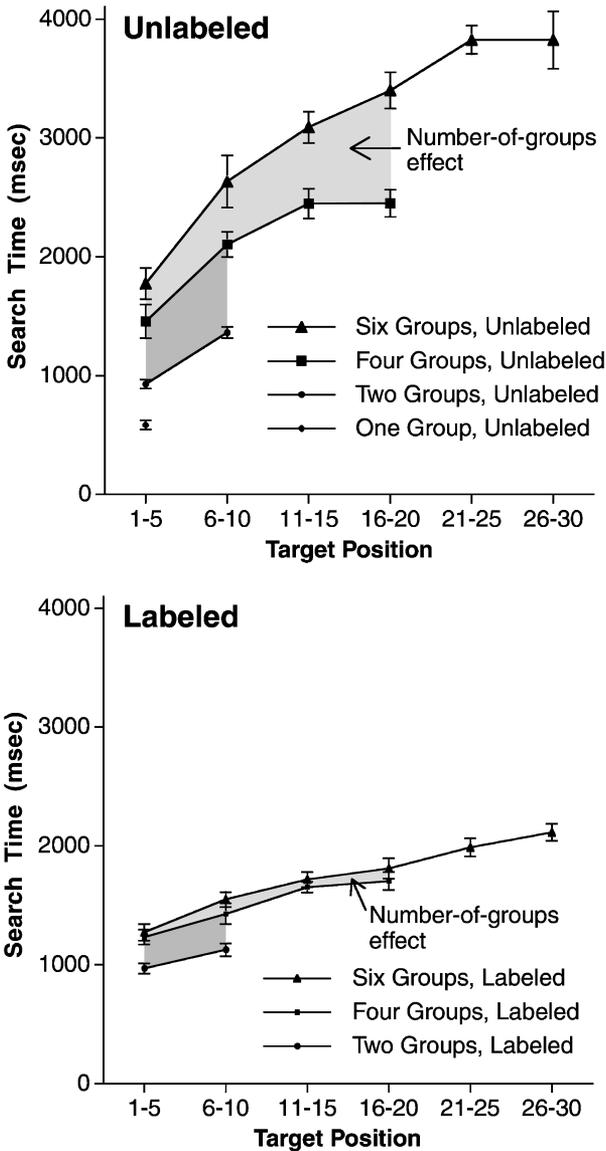


Fig. 6. Mean search time as a function of the target position for each of the seven layouts with distractors. The number-of-groups effect (the shaded areas between the curves) is larger for the unlabeled layouts (shown in the top graph). Search times are averaged within each 5-position group to reduce the noise and to emphasize the number-of-groups effect. The bars above and below each plot symbol shows the standard error.

As can be seen in the figure, nearly all pointing times are well below the deadline. Most of the outliers that appear above the deadline are trials in which the experimental software would have interrupted the trial; a few are trials in which the deadline was extended slightly due to the mouse-position sampling rate.

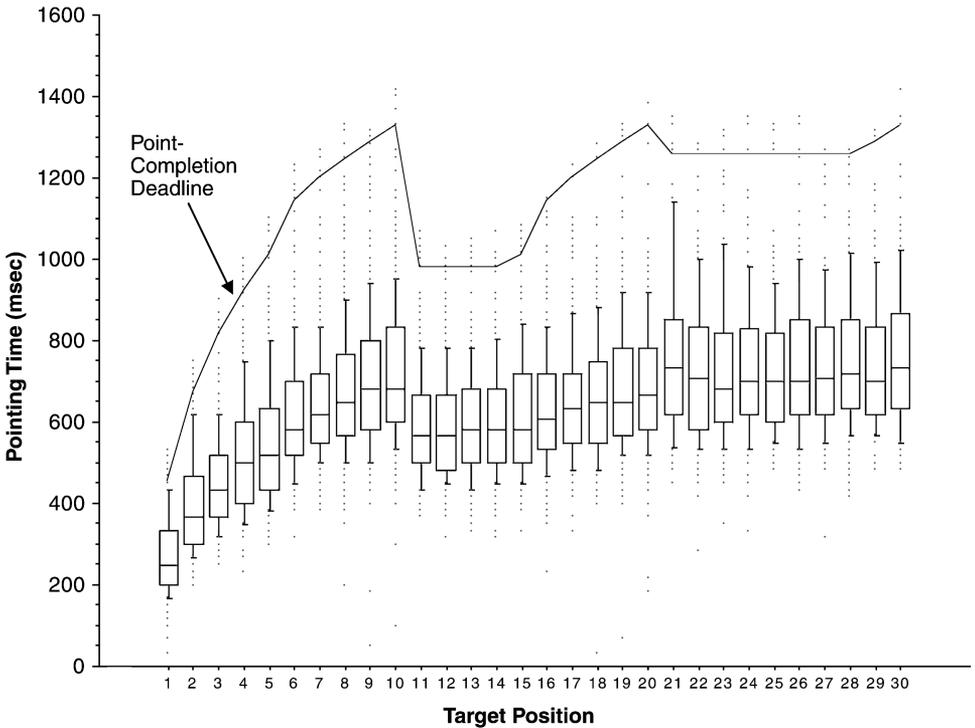


Fig. 7. Pointing times as a function of the target position, from all trials. Each box plot shows the 10th, 25th, 50th, 75th, and 90th percentile. Also shown: the point-completion deadline imposed by the experimental software. Most of the outliers that appear above the deadline are error trials in which the experimental software interrupted the trial.

In Figure 7, there is a sharp drop from position 10 to position 11 both in the point-completion deadline and the observed pointing times. The drop from position 10 to position 11 occurs because position 10 is at the bottom of the first column and position 11 is at the top of the second column. Fitts' law (Eq. 1) predicts pointing time as a function of the index of difficulty (ID), which is

$$ID = \log_2(d/w + 0.5) \quad (2)$$

where d is the distance to the target and w is the width of the target. Positions 10 and 11 are roughly the same distance from the precue, but the width of position 10 is much smaller than the width of position 11 if the widths are measured along the path of travel from the precue, as when calculating the point-completion deadline. Thus, the ID of position 10 is larger than that of position 11 (3.42 as opposed to 2.26), pointing time to position 10 is greater than that to position 11, and there is a sharp drop in the graph from position 10 to 11. The slight drop in the point-completion deadline from position 20 to 21 can be similarly explained, though the drop is somewhat lost in the noise of the observed data.

Mean pointing times vary slightly for each of the three layout types (labeled, unlabeled, and target-only). Target-only layouts are slightly (25 msec) faster

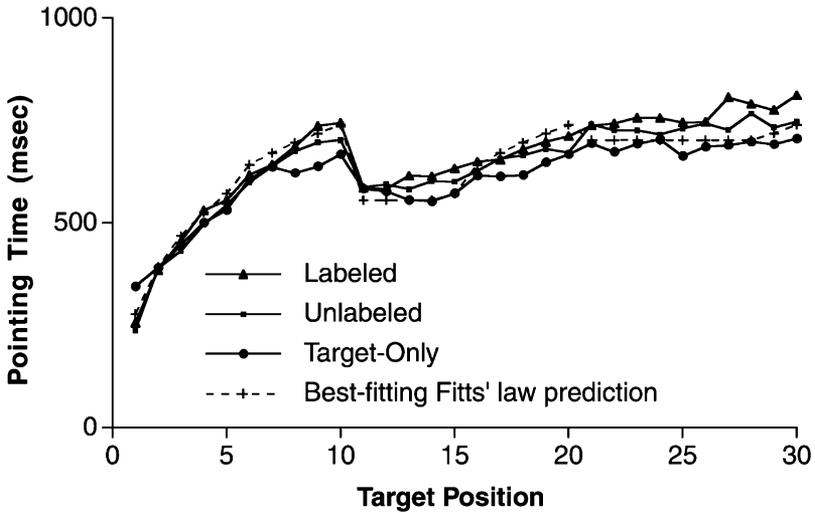


Fig. 8. Mean pointing time as a function of target position, from all nonerror trials, separated by layout type (labeled, unlabeled, and target-only). Also shown is the overall best-fitting Fitts' law prediction ($a = 196$, $b = 159$). Note that the ordinate is expanded in comparison to the search time graphs that appeared earlier.

than unlabeled layouts, and unlabeled layouts are slightly (22 msec) faster than labeled layouts. Two single-factor repeated measures ANOVAs (using the 16 participant means for each layout type, giving equal weight to each target position) confirmed that these differences, although small, are statistically reliable. There is a significant difference between target-only and unlabeled layouts, $F(1, 15) = 9.9$, $p < 0.01$, and between unlabeled and labeled layouts, $F(1, 15) = 5.0$, $p < 0.05$.

Figure 8 shows the mean pointing time as a function of the target position, with each of the three layout types shown separately. All nonerror trials for each position in each layout type are given equal weight in computing each mean. As can be seen in the figure, there are slight differences in pointing times for each of the three layout types, but these differences are slightly exaggerated in the graph because the ordinate is expanded in comparison to the search time graphs shown earlier.

As can be seen in Figure 8, the pointing times for each layout type conform reasonably well to the overall best-fitting Fitts' law prediction (with w set to the width of the target along the line of approach, and with $a = 196$ and $b = 159$). The prediction explains the overall mean observed pointing times with an r^2 of 0.94. Table II shows the best-fitting Fitts' law parameters for each layout type and for all nonerror trials. Curiously, the parameters for the target-only layout are quite different than those for the labeled and unlabeled layouts. These differences will be discussed later.

8.3 Errors

Two types of errors occurred. An *interruption* occurred when the participant moved the mouse but did not click on anything before the point-completion

Table II. The Best-Fitting Fitts' Law Parameters for the Observed Pointing Times

Layout Type	a	b	r^2
Labeled	179	173	0.92
Unlabeled	183	164	0.92
Target-only	260	127	0.92
All	196	159	0.94

Welford version of Fitts' law: Movement time = $a + b \log_2(d/w + 5)$ msec.

Table III. Error Rates for Each Layout Type

Layout Type	Interruption Rate		Miss Rate	
	Mean	SD	Mean	SD
Labeled	3.4%	3.5%	1.9%	1.6%
Unlabeled	3.5%	2.6%	1.7%	1.1%
Target-only	1.6%	1.9%	2.8%	2.7%

SD = standard deviation of the 16 participant error rates for each layout type.

deadline. A *miss* occurred when the participant clicked before the deadline, but not on the target. Table III shows the mean and standard deviation of both error rates for the three layout types. The labeled and unlabeled layouts have similar error rates, which suggests that participants used roughly the same speed-accuracy operating characteristic for labeled and unlabeled layouts, and that the reaction time differences between the two layouts should be attributed to task demands and not to a speed-accuracy tradeoff. Note, however, that layouts with distractors (labeled and unlabeled) have different error rates than those of the target-only layout.

9. DISCUSSION

The search time, pointing time, and error results have implications for understanding the fundamental human processing involved in (a) the visual search of a labeled visual hierarchy and (b) selecting an item on a computer screen with a mouse. The implications for mouse-pointing are discussed first.

9.1 Mouse-Pointing

It appears as if participants used a slightly different strategic integration of perceptual and motor processing, or *microstrategy* [Gray and Boehm-Davis 2000], to move the cursor to the target when distractors were present (in the labeled and unlabeled layouts) as opposed to when distractors were not present (in the target-only layout). It appears as if participants made slower and more accurate mouse movements when distractors were present. This is demonstrated by (a) slightly slower pointing times in layouts with distractors; (b) higher b parameters in the best-fitting Fitts' predictions for layouts with distractors; and (c) higher interruption rates in layouts with distractors.

Participants were evidently more careful to hit the target when other visual objects were near the target. This is a plausible strategic decision, given that all participants were experienced users of graphical user interfaces and, when using such an interface, the cost of undoing a wrong action (such as clicking on the wrong button) is generally greater than the cost of retrying a failed action (such as clicking on blank space next to the right button).

Another possible explanation for the faster pointing times in the target-only layout is simply that, since no visual search was required, only a single eye movement was made from the precue to the target. The ocular-motor command and the proprioceptive feedback from the muscle movements required for this eye movement may have helped to generate a more efficient cursor movement along this trajectory because it would have been the same trajectory as the just-completed eye movement. Abrams et al. [1990] discuss this sort of eye-hand coordination.

Faster pointing times in the target-only layout may also have resulted because, as proposed by Meegan and Tipper [1999], the motor system seems to automatically prepare movements to all visual objects that appear in a scene, regardless of a conscious intention to act on that information. When distractors are present, less preparation is dedicated to the target item, and the movement to the target is slower.

These three possible explanations all suggest that Fitts' parameters derived from experiments in which the target appears without distractors (including most Fitts' law experiments, such as those discussed earlier) will under-predict mouse-pointing time for real-world pointing tasks in which other visual objects are on the screen near the target and in which mouse movements follow visual search.

The observed movement times affirm that Fitts' law can be applied to accurately predict movement times in a two-dimensional pointing task—that is, when the angle of the pointing movement is diagonal and when the target is a rectangle. Researchers have investigated which physical dimension of a target object should be used as the width w in Fitts' law when the target is approached from different angles [Gillan et al. 1990; MacKenzie 1992; MacKenzie and Buxton 1992].

The movement times reported here are best predicted by Fitts' law if w is set to the width of the target along the line of approach, as discussed by MacKenzie [1992]. None of the other measurements proposed for w fit the data as well. Gillan et al. [1990] investigated using the object's height, width, height + width, and height · width. But using each of these as the w in Fitts' law, to explain the movement times reported here, results in best-fitting Fitts' law functions with r^2 values of, respectively, 0.87, 0.86, 0.84, and 0.72. Whereas setting w to the width along the line of approach produces an r^2 of 0.94. Mackenzie and Buxton [1992] conducted a two-dimensional mouse-pointing experiment and concluded that the data could be fit equally well by setting w to either (a) the smaller of the two physical dimensions (in this case, the height); or (b) the width along the line of approach. The data collected here, however, is best explained by using the latter of the two (resulting in an r^2 of 0.94, as opposed to 0.87). Based on the performance of participants in this experiment, it appears as if setting w

to the width along the line of approach contributes to the most accurate Fitts' prediction in a two-dimensional pointing task.

These analyses were also conducted using the Shannon formulation of Fitts' law [MacKenzie 1992; MacKenzie and Buxton 1992], setting w to the height, width, height + width, and height · width, and the width along the line of approach. The best-fitting functions resulted in r^2 values of, respectively, 0.86, 0.83, 0.82, 0.70, and 0.93. Again, setting w to the width along the line of approach again results in a substantially better fit. Comparing these correlations to those for the Welford version, it is clear that the Shannon formulation does not fit the data better.

Though there are slight differences in mean pointing times for the different layout types, these differences are very small when compared to the overall task times. The similarity in pointing times across all trials, which can be seen in Figures 7 and 8, suggests that participants had only enough time for one mouse movement in each trial, and that they spent a rather consistent amount of time on that one movement. It appears as if the point-completion deadline successfully discouraged mouse movements from being made until the target was found, and that the experiment successfully separated search time from pointing time.

9.2 Visual Search

The much faster search times for labeled layouts clearly support the first hypothesis—that a labeled visual hierarchy can be searched more quickly than an unlabeled visual hierarchy. When labels are present, participants use them to find the target much more quickly. This supports the claim made by screen layout design guidelines that a visual hierarchy is a powerful design tool. A higher level structure of grouping and useful labels can quickly direct the viewer's attention to the required item on the display. People can quickly find their way through a lot of information on a computer screen if it is well-organized.

The second hypothesis, that people use a fundamentally different strategy to search a labeled visual hierarchy, as opposed to an unlabeled visual hierarchy, is supported by differences in the search time data for labeled and unlabeled layouts. Different search strategies create different trends in the data. Some of these trends will become evident if the data is examined on a position-by-position basis, as is done here.

Before discussing the differences in search time data for labeled and unlabeled layouts, there are a couple of similarities in the data that suggest similarities in search strategies. The target position effect for labeled and unlabeled layouts suggests that, in both cases, search is ordered and self-terminating (nonexhaustive). The target position effect suggests that participants tended to start their search at the top of the first column, search that column from top to bottom, move to the second column (if it was there), search it from top to bottom, move to the third column (if it was there), and search it from top to bottom.

The obvious difference in the strategies evidently used for labeled versus unlabeled layouts is that, in the labeled layouts, participants seem to have first

limited their search to the labels until the group containing the target was found, and then confined their search to that group. In the unlabeled layouts, however, participants would have examined the individual menu items.

The more subtle difference between the strategies evidently used for labeled versus unlabeled layouts is that participants appear to have searched the labeled layouts in a much more consistent and systematic manner. The smaller number-of-groups effect suggests as much. When positions are examined in the same systematic order regardless of how many items appear, each target position requires roughly the same amount of time regardless of how many items appear. There is little or no number-of-groups effect, as is the case with the labeled layouts. In a random search, however, the search time for each target position goes up when there are more items in the layout to compete for the next randomly-chosen consideration, and there is a larger number-of-groups effect, as is the case in the unlabeled layouts. The larger coefficient of variation and standard errors for unlabeled layout search times also suggest that participants used a wider variety of search orders for the unlabeled layouts.

A factor that may have contributed to the more varied performance with unlabeled layouts is that the unlabeled layouts relied more heavily on lexical search. The targets and distractors had a wide range of lexical and semantic properties, which may have led to some being recognized and identified more quickly than others. The more salient semantic properties of a target such as *FOX*, for example, might help it to be found more quickly than a target such as *LEL*. This lexical and semantic variability would have contributed more to the variability in search times for unlabeled layouts than for labeled layouts. An interesting follow-up study would examine how search times vary as a function of the meaningfulness of the target and distractor CVCs.

The one result not yet discussed is the search time for the target-only layout. The very fast and nearly constant search time for the target-only layout suggests that, when no distractors are present, the target's location was recognized in about 200 msec, regardless of where it appeared on the screen. Evidently, little or no search was required.

9.3 Implications for Visual Search Models

The screen layouts in this experiment were converted into alphanumeric screenshots and input into Tullis' DAP. Figure 9 shows DAP's predictions and the mean search time observed for each layout. As can be seen, DAP predicts the upper bound of the observed search times, but not the dramatic speed improvements realized with smaller layouts and labeled layouts.

In order to predict that labeled layouts will speed visual search, a visual search model needs to incorporate some form of a visual search strategy—the procedure used to find a target given details of the task and layout. Visual search strategies are an integral part of Lohse's [1993] UCIE system, as well as the menu-selection models built by Anderson et al. [1997], Byrne [2001], and Hornof and Kieras [1997; 1999]. Group labels are a small and subtle modification to the physical layout, but they create a hierarchical visual structure that motivates a different and more efficient strategy. Comprehensive

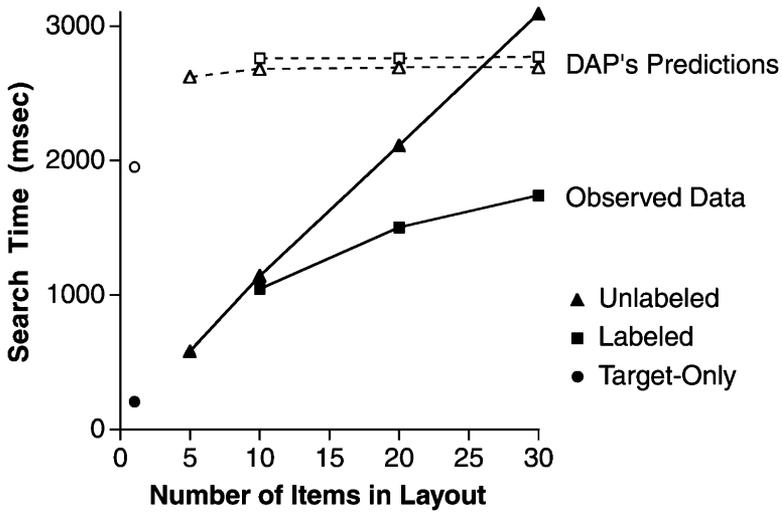


Fig. 9. Search times predicted by Tullis' DAP (dashed lines, unfilled symbols) and observed (solid lines, filled symbols) for the eight layouts, as a function of the number of items in the layout. DAP does not predict the dramatic speed increases observed with smaller layouts and with labeled layouts.

visual search models need to incorporate strategies that respond to such minor layout modifications.

9.4 Implications for Screen Layout Design

This experiment supports the claim made by screen layout designers that people establish a plan for how to search a visual layout based on the organization and visual structure of the layout. A labeled visual hierarchy can greatly speed visual search. The work presented here also demonstrates that a screen layout with a large number of items to choose from will not necessarily result in a long search time. Search time will increase as a function of the number of items on the screen, but at a much slower rate when items are arranged in groups with useful headings that aid the information forager. These results do not support screen layout guidelines that set arbitrary limits to screen density (such as proposed by Galitz [1996]) or the number of items in a screen layout. If the layout is well-organized, search times will not increase dramatically as the number of items on the screen is increased.

In this experiment, examining labeled versus unlabeled layouts, the two extremes were examined. More work is needed to determine the effects when groups are unlabeled, but grouped based on functional or other similarities that can be determined by examining one or two of the items. Card [1982] and McDonald et al. [1983] have looked at such issues. Clearly, all screen layouts will not be as rigidly structured as those used in this experiment, but these results suggest that well-structured layouts will be searched in a predictable manner, and that a screen layout designer has some control over how a layout will be searched.

APPENDIX A

Table A. Mean Observed Search Times for Each Layout

Target Position	Unlabeled				Labeled			Target-Only
	6 Groups	4 Groups	2 Groups	1 Group	6 Groups	4 Groups	2 Groups	
1	882	1051	728	560	1052	1047	719	259
2	2159	1159	803	514	1205	1123	927	231
3	1295	1662	982	600	1332	1217	1108	222
4	2361	1460	1079	606	1326	1420	956	228
5	2164	1940	1039	635	1440	1344	1125	205
6	2964	2441	1347		1389	1373	961	210
7	2308	2014	1271		1526	1416	1164	196
8	2867	2010	1443		1607	1403	1090	213
9	2628	1945	1322		1603	1413	1238	210
10	2393	2104	1417		1615	1532	1171	219
11	2614	2238			1568	1639		213
12	2758	2300			1803	1743		203
13	2950	2474			1690	1578		196
14	3357	2411			1747	1698		203
15	3765	2811			1769	1596		200
16	3068	2303			1681	1602		197
17	3088	2247			1700	1588		199
18	4197	2308			1771	1704		199
19	3505	2411			1983	1726		208
20	3136	2982			1906	1900		197
21	3607				1871			206
22	3122				1946			209
23	3713				2001			202
24	3758				2046			199
25	4923				2066			207
26	4052				2004			202
27	3198				2014			208
28	3660				2134			196
29	4287				2228			193
30	3921				2183			215

Times are in msec.

APPENDIX B

Table B. Mean Observed Pointing Time for Each Target Position and Layout Type, for all Nonerror Trials. (Also, the distance from the center of the precue to the center of each target, and the width of each target along the direct path of travel from the precue)

Target Position	Target Distance	Target Width	Layout Type			All Trials
			Labeled	Unlabeled	Target-Only	
1	25	27	255	238	345	257
2	50	27	384	390	390	388
3	75	27	455	431	446	442
4	100	27	530	496	500	509
5	125	27	555	543	531	546
6	175	27	617	600	607	608
7	200	27	642	638	636	640
8	225	27	686	675	622	672
9	250	27	737	696	638	705
10	275	27	744	702	667	715
11	301	70.2	585	587	583	586
12	304.1	71	583	594	577	586
13	309.2	72.2	615	582	555	590
14	316.2	73.8	614	601	552	597
15	325	70.2	632	600	572	607
16	347.3	53.6	649	627	615	634
17	360.6	48.7	655	657	614	647
18	375	45	679	665	616	661
19	390.5	42.2	698	680	648	681
20	407	40	711	671	667	687
21	600.5	70.1	737	740	694	724
22	602.1	70.2	743	726	673	714
23	604.7	70.5	757	725	693	725
24	608.3	71	756	716	703	725
25	612.9	71.5	745	730	663	713
26	625	72.9	746	743	685	725
27	632.5	73.8	806	727	689	741
28	640.8	74.8	790	767	698	752
29	650	70.2	775	733	691	733
30	660	64.8	811	747	706	754
30	660	64.8	811	747	706	754

Distance and width are in pixels. Times are in msec.

ACKNOWLEDGMENTS

The author would like to thank Judy Olson for motivating this project; David Kieras for extensive feedback and guidance; David Meyer for assisting with the experimental design; Wayne Gray, Sarah Douglas, Arthur “Ted” Kirkpatrick, Alistair Sutcliffe, and four anonymous reviewers for feedback on various drafts; and Tom Tullis for providing a copy of DAP.

REFERENCES

- ABRAMS, R. A., MEYER, D. E., AND KORNBUM, S. 1990. Eye-hand coordination: Oculomotor control in rapid aimed limb movements. *J. Exper. Psychol.: Human Perception Perform.* 16, 2, 248–267.
- ANDERSON, J. R. AND LEBIERE, C. (Eds.). 1998. *The Atomic Components of Thought*. Erlbaum, Hillsdale, NJ.
- ANDERSON, J. R., MATESSA, M., AND LEBIERE, C. 1997. ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Comput. Interaction*, 12, 4, 439–462.
- Apple Computer INC. 1999. *iMac User's Guide*. Apple Computer, Inc., Cupertino, CA.
- BIGUER, B., JANNEROD, M., AND PRABLANC, C. 1985. The role of position of gaze in movement accuracy. In *Attention and Performance XI*. M. I. Posner and O. S. M. Marin (Eds.), Erlbaum, Hillsdale, NJ, 407–424.
- BURNS, C. M. 2000. Navigation strategies with ecological displays. *Int. J. Human-Comput. Stud.* 52, 111–129.
- BURNS, M. J., WARREN, D. L., AND RUDISILL, M. 1986. Formatting space-related displays to optimize expert and nonexpert user performance. In *Proceedings of the CHI '86 Conference*, ACM, New York, NY.
- BYRNE, M. D. 2001. ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *Int. J. Human-Comput. Stud.* 55, 41–84.
- BYRNE, M. D., ANDERSON, J. R., DOUGLASS, S., AND MATESSA, M. 1999. Eye tracking the visual search of click-down menus. In *Proceedings of the CHI 99 Conference*, ACM, New York, NY, 402–409.
- CARD, S. K. 1982. User perceptual mechanisms in the search of computer command menus. In *Proceedings of the CHI '82 Conference*, ACM, New York, NY, 190–196.
- CARD, S. K., ENGLISH, W. K., AND BURR, B. J. 1978. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics* 21, 601–613.
- CARD, S. K., MORAN, T. P., AND NEWELL, A. 1983. *The Psychology of Human-Computer Interaction*. Erlbaum, Hillsdale, NJ.
- EPPS, B. W. 1986. Comparison of six cursor control devices based on Fitts' law models. In *Proceedings of the Human Factors Society 30th Annual Meeting*. Human Factors Society, Santa Monica, CA, 327–331.
- FITTS, P. M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exper. Psychol.* 47, 6, 381–391.
- FURNAS, G. W., LANDAUER, T. K., GOMEZ, L. M., AND DUMAIS, S. T. 1987. The vocabulary problem in human-system communication. *Commun. ACM* 30 (Nov.), 964–971.
- GALITZ, W. O. 1996. *The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques*. John Wiley, New York.
- GILLAN, D. J., HOLDEN, K., ADAM, S., RUDISILL, M., AND MAGEE, L. 1990. How does Fitts' law fit pointing and dragging? In *Proceedings of the CHI '90 Conference*, ACM, New York, NY, 227–234.
- GRAY, W. D. AND BOEHM-DAVIS, D. A. 2000. Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *J. Exper. Psychol. Appl.* 6, 4, 322–335.
- HAN, S. H., JORNA, G. C., MILLER, R. H., AND TAN, K. C. 1990. A comparison of four input devices for the Macintosh interface. In *Proceedings of the Human Factors Society 34th Annual Meeting*, Human Factors Society, Santa Monica, CA, 267–271.
- HORNOF, A. J. 1999. Computational models of the perceptual, cognitive, and motor processes involved in the visual search of pull-down menus and computer screens. Ph.D. dissertation, University of Michigan, Ann Arbor.

- HORNOF, A. J. AND KIERAS, D. E. 1997. Cognitive modeling reveals menu search is both random and systematic. In *Proceedings of the ACM CHI 97: Conference on Human Factors in Computing Systems*, ACM, New York, NY, 107–114.
- HORNOF, A. J. AND KIERAS, D. E. 1999. Cognitive modeling demonstrates how people use anticipated location knowledge of menu items. In *Proceedings of ACM CHI 99: Conference on Human Factors in Computing Systems*, ACM, New York, NY, 410–417.
- KIERAS, D. E. AND MEYER, D. E. 1997. An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Comput. Interaction* 12, 4, 391–438.
- LOHSE, G. L. 1993. A cognitive model for understanding graphical perception. *Human-Comput. Interaction* 8, 353–388.
- MACKENZIE, I. S. 1991. Fitts' law as a performance model in human-computer interaction. Ph.D. dissertation, University of Toronto. <http://www.yorku.ca/faculty/academic/mack/phd.html>.
- MACKENZIE, I. S. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-Comput. Interaction* 7, 91–139.
- MACKENZIE, I. S. AND BUXTON, W. 1992. Extending Fitts' law to two-dimensional tasks. In *Proceedings of the CHI '92 Conference*, ACM, New York, NY, 219–226.
- MCDONALD, J. E., STONE, J. D., AND LIEBELT, L. S. 1983. Searching for items in menus: The effects of organization and type of target. In *Proceedings of the Human Factors Society 27th Annual Meeting*, Human Factors Society, Santa Monica, CA, 834–837.
- MEEGAN, D. V. AND TIPPER, S. P. 1999. Visual search and target-directed action. *J. Exper. Psychol. Human Perception Perform.* 25, 5, 1347–1362.
- MULLET, K. AND SANO, D. 1995. *Designing Visual Interfaces: Communication Oriented Techniques*. Prentice Hall, Englewood Cliffs, NJ.
- NILSEN, E. L. 1991. Perceptual-motor control in human-computer interaction. Tech. Rep. and Ph.D. dissertation 37, Univ. of Michigan, Ann Arbor.
- NORMAN, K. L. 1991. *The Psychology of Menu Selection: Designing Cognitive Control of the Human/Computer Interface*. Ablex, Norwood, NJ.
- PAAP, K. R. AND COOKE, N. J. 1997. Design of menus. In *Handbook of Human Computer Interaction* (2nd ed.), M. Helander, T. K. Landauer, and P. Prabhu (Eds.), North-Holland, Amsterdam, 533–572.
- PIROLI, P. AND CARD, S. K. 1999. Information foraging. *Psychol. Rev.* 106, 643–675.
- ROSENBAUM, D. A. 1991. *Human Motor Control*. Academic Press, New York, NY.
- SEARS, A. AND SHNEIDERMAN, B. 1994. Split menus: Effectively using selection frequency to organize menus. *ACM Trans. Comput. Human Interaction* 1, 1, 27–51.
- SOMBERG, B. L. AND PICARDI, M. C. 1983. Locus of the information familiarity effect in the search of computer menus. In *Proceedings of the Human Factors Society 27th Annual Meeting*, Human Factors Society, Santa Monica, CA, 826–830.
- STAGGERS, N. 1993. Impact of screen density on clinical nurses' computer task performance and subjective screen satisfaction. *Int. J. Man-Mach. Stud.* 39, 775–792.
- THACKER, P. P. 1986. Tabular displays: A human factors study. Ph.D. dissertation. Rice Univ., Houston, TX.
- TREISMAN, A. 1982. Perceptual grouping and attention in visual search for features and for objects. *J. Exper. Psychol. Human Perception Perform.* 8, 2, 194–214.
- TULLIS, T. S. 1988. A system for evaluating screen formats: research and application. In *Advances in Human-Computer Interaction* (Vol. 2), R. Hartson and D. Hix (Eds.), Ablex, Norwood, NJ, 214–286.
- TULLIS, T. S. 1997. Screen design. In *Handbook of Human-Computer Interaction* (2nd ed.), M. Helander, T. K. Landauer, and P. Prabhu (Eds.), North-Holland, Amsterdam, 503–531.
- WALKER, N., MEYER, D. E., AND SMELLICER, J. B. 1993. Spatial and temporal characteristics of rapid cursor-positioning movements with electromechanical mice in human-computer interaction. *Human Factors* 35 3, 431–458.
- WELFORD, A. T. 1968. *Fundamentals of Skill*. Methuen, London.
- WILLIAMS, R. AND TOLLET, J. 1998. *The Non-Designer's Web Book*. Peachpit Press, Berkeley, CA.

Received October 2000; revised January 2001; accepted March 2001