Cognitive Modeling Reveals
Menu Search is Both Random and Systematic

Anthony J. Hornof and David E. Kieras

Artificial Intelligence Laboratory
Electrical Engineering & Computer Science Department
University of Michigan
1101 Beal Avenue, Ann Arbor, MI 48109-2110
+1 313 763 6985
hornof@umich.edu, kieras@eecs.umich.edu

ABSTRACT

To understand how people search for a known target item in an unordered pull-down menu, this research presents cognitive models that vary serial versus parallel processing of menu items, random versus systematic search, and different numbers of menu items fitting into the fovea simultaneously. Varying these conditions, models were constructed and run using the EPIC cognitive architecture. The selection times predicted by the models are compared with selection times of human subjects performing the same menu task. Comparing the predicted and observed times, the models reveal that 1) people process more than one menu item at a time, and 2) people search menus using both random and systematic search strategies.

KEYWORDS

Cognitive models, menu selection, visual search.

INTRODUCTION

Models of human performance permit aspects of user interfaces to be evaluated for usability by making predictions based on task analysis and established principles of human performance [4, 5]. Though much previous research (including [3, 9, 12, 14]) has investigated menu selection, there are no empirically validated models of the low-level perceptual, cognitive, and motor processing that people use when they select a known target item from an unordered pull-down menu.

Researchers have proposed theories about the low-level strategies that people use to find a known item in an unordered menu. Norman [12] and Vandierendonck, Van Hoe, and De Soete [14] suggested that people process one menu item at a time. But they did not validate this low-level assumption empirically. There have also been conflicting theories. Card [3] proposed that people randomly choose which item to examine next, while Lee and MacGregor [9] provided evidence that people search systematically from top to bottom. The research presented here examines the plausibility of these theories by providing an empirically validated model of the low-level

Appears in the ACM CHI'97 Conference Proceedings.

perceptual, cognitive, and motor processing that people use in a menu selection task.

THE EPIC COGNITIVE ARCHITECTURE

The EPIC (Executive Process Interactive Control) cognitive architecture [6, 7] provides a general framework for simulating a human interacting with their environment to accomplish a task, and is well-suited to model a menu selection task. EPIC resembles the Model Human Processor [4], but differs in that EPIC is a precise computational model, has a programmable production-rule cognitive processor, and incorporates more specific constraints synthesized from human performance literature.

EPIC consists of a production-rule cognitive processor and perceptual-motor peripherals. To model human performance aspects of accomplishing a task, a cognitive strategy and perceptual-motor processing parameters must be specified. A cognitive strategy is represented as a set of production rules, much the same way that CCT [2], ACT-R [1], and SOAR [8] represent procedural knowledge. The simulation is driven by a description of the task environment that specifies aspects of the environment that would be directly observable to a human, such as what objects appear at what times, and how the environment changes based on EPIC's motor movements. EPIC computational models are *generative* in that the production rules only represent general procedural knowledge of the task, and when EPIC interacts with the task environment, EPIC generates a specific sequence of perceptual, cognitive, and motor activities required to perform each specific instance of the task.

EPIC takes as its input:
- The cognitive strategy for accomplishing a task.
- Availability of object features, to represent human perceptual limitations.
- Details of the task environment.

EPIC generates as output:
- The time required to execute the task.
- A trace of the modeled human processing.

As shown in Figure 1, information flows from sense organs, through perceptual processors, to a cognitive processor (consisting of a production rule interpreter and a working memory), and finally to motor processors that

control effector organs. All processors run independently and in parallel.
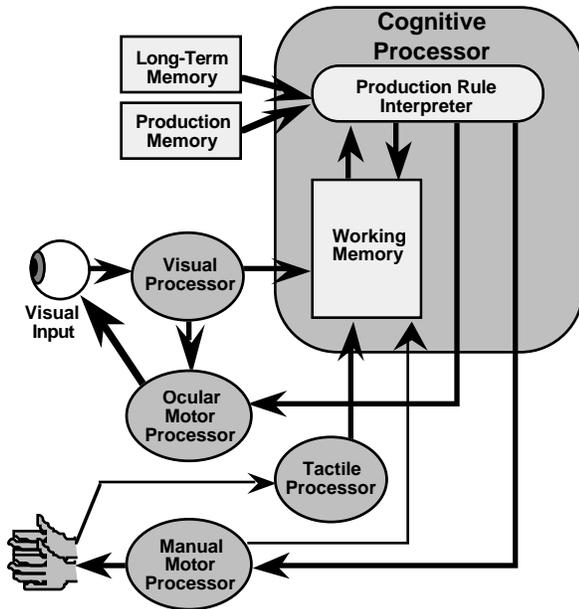
Figure 1. Subset of EPIC architecture, showing flow of information and control. The processors run independently and in parallel. Not shown: Auditory and vocal motor processors, task environment.

A single stimulus in the task environment can produce multiple outputs from a perceptual processor to be deposited in working memory at different times. First the detection of a perceptual event is sent, followed later by features that describe the event. The perceptual processors are "pipelined." If an object's features begin moving to working memory, the arrival of those features will not be delayed by any other processing. Working memory contains these items deposited by perceptual processors, as well as control information such as the current task goal. At the end of each simulated 50 msec cycle, EPIC fires all of the production rules whose conditions match the current contents of working memory. EPIC allows for parallel execution of production rules in the cognitive processor, and some parallelism in each motor processor.

In short, EPIC is applied to a task as follows: The production-rule strategy directs the eyes to objects in the environment. The eyes have a resolving power which determines the processing time required for different object features, such as location and text. When information needed to determine the next motor movement arrives in working memory, the strategy instructs the ocular motor and manual motor processors to move the eyes and hands.

Information processing and motor movement times are held constant across modeling efforts, and are based on human performance literature. Manual movement times, for example, are determined by Fitts' law (see [4], Ch. 2). For lack of space, EPIC cannot be described in full detail here. A more thorough description is presented in [6, 7].

THE TASK

The specific pull-down menu task modeled in this paper is based on a menu selection task used by Nilsen in an experiment with human subjects (Experiment 2 in [11]). Nilsen used menus that had three, six, and nine menu items. Each menu item was a single numerical digit. Menu items were randomly reordered for each trial. Subjects were experienced mouse (and thus presumably menu) users and were financially motivated to perform each trial as quickly as possible. Nilsen ran eight subjects, each with six trials for every possible combination of menu length and target position. The distance between menu items was roughly 0.2 inches. The distance from eye to screen was neither controlled nor measured.

As shown in Figure 2, each trial consisted of the following steps: Using the mouse, move the cursor to the GO box which causes the precue of the target item to appear above the GO box. Commit the precue to memory. Click on the GO box. The GO box and precue disappear, the menu appears, and the clock starts. As quickly as possible, click on the target item in the menu. The clock stops.
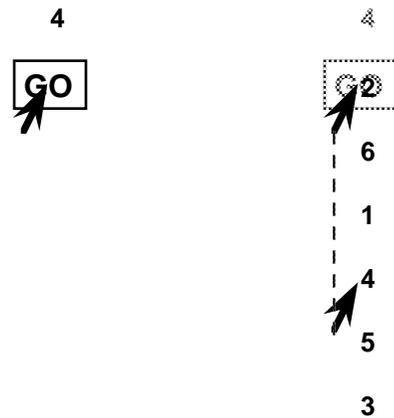
Figure 2. Nilsen's task with six items in the menu.

This task isolates a subset of the processes required in a "real world" menu task. It is thus particularly well-suited for studying the low-level perceptual-motor processes of visual search and response selection. The task is not confounded with more complex processes of reading, comprehension, judgment, decision making, and problem solving. Though Nilsen mostly used the data to examine motor control, this modeling effort focuses on visual search. The data is particularly useful for modeling visual search of menus because Nilsen varied menu length and reported selection time as a function of the serial position of the target menu item. Few researchers have reported such data. As will be shown, this combination is critical for revealing search strategy.

THE OBSERVED DATA

Figure 3 shows Nilsen's observed data, averaged across subjects and blocks, as well as the time required to move the mouse to each position as predicted by Fitts' law (Welford's form of Fitts' law, see [4], Ch. 2).
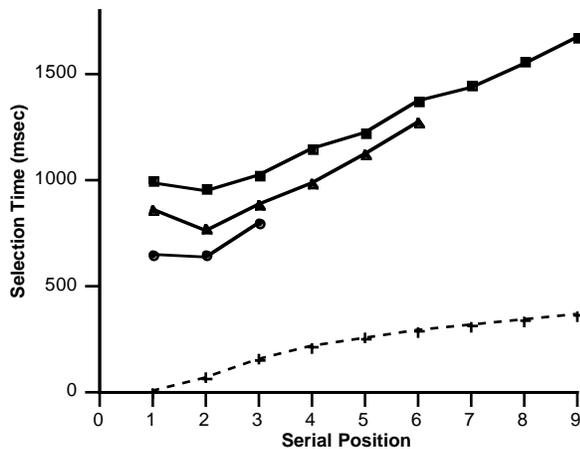
Figure 3. Nilsen's observed data (solid lines). Mean selection times as a function of serial position of target item, for menus with three, six, or nine items. Also: Time required to move the mouse to each target position as predicted by Fitts' law (dashed line).

There are several key features to note in the observed data:

- When the target item is in the same position across menus of different lengths, shorter menus are faster.
- Selection time increases with a fairly linear slope of about 100 msec per item for each menu length (excluding serial position 1). As can be seen in the graph, the mouse movement time predicted by Fitts' law cannot entirely account for this slope.
- The selection time for serial position 1 is a little higher than the selection time for serial position 2.

The EPIC models that follow are all evaluated with respect to how well they match these trends in Nilsen's observed data.

THE MODELS

This section presents six models that result from varying two strategic dimensions: Serial versus parallel processing of menu items, and random versus systematic search. In the parallel processing models, the eye-to-screen distance is varied (8 and 20 inches) to result in one or three items being visible in the fovea simultaneously. The fovea is fixed as the circular region within 1° of visual angle from the center of the gaze. It is assumed that recognition of digits is only possible in the fovea.

The discussion of each model includes a flowchart that summarizes the production rules written in EPIC to represent that model. Production rules were written to maximize performance within the constraints imposed by EPIC, and to be as simple as possible. EPIC was otherwise used 'as is' for all models. Details and parameters such as the availability of object features were

established and validated in other modeling projects in different task domains, and are discussed in [6, 7].

Serial Processing Models

The serial processing models represent a belief that people move their gaze to an item, visually process it, decide if it is the target, click on the item if it is, or go on to the next item if it is not. Figure 4 represents a serial processing model proposed by Norman [12]. Since the model proposed in Figure 4 does not specify the search strategy used to find the next item, two separate sets of production rules were built in EPIC to represent two possible models, one with random search and the other with systematic top-to-bottom search.
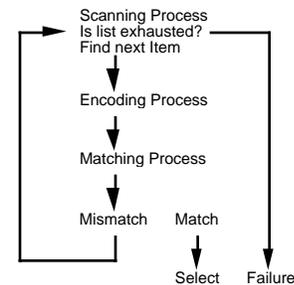
Figure 4. Norman's [12] information processing model for search of an explicitly known target.

Both serial processing models were only run with an eye-to-screen distance of 8 inches so that only one item would fit into the fovea, insuring the serial encoding process specified by Norman. At greater distances, more than one item would fit into the fovea simultaneously, and parallel encoding would ensue.

Serial Processing Random Search Model

The results from running the Serial Processing Random Search model are shown in Figure 5. Each predicted selection time is averaged from 300 trials run for that menu length and serial position combination.
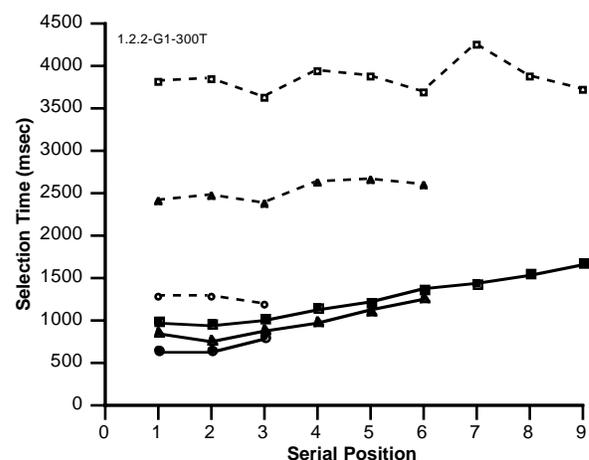
Figure 5. Selection times observed (solid lines) and predicted (dashed lines) by the Serial Processing Random Search model run with one item fitting into the fovea.

The results in Figure 5 (on preceding page) suggest that the Serial Processing Random Search model is wrong. The only feature in the observed data that this model accounts for is that shorter menus are faster than longer menus. Otherwise, the model does not fit the observed data. Selection times are much too high overall. Slopes are very small because every item takes on average the same amount of time to find and select; any slope that appears is due to the mouse movement. A higher selection time for serial position 1 is not predicted. This model does not account for the observed data.

Serial Processing Systematic Search Model
The results from running the Serial Processing Systematic Search model are shown in Figure 6.
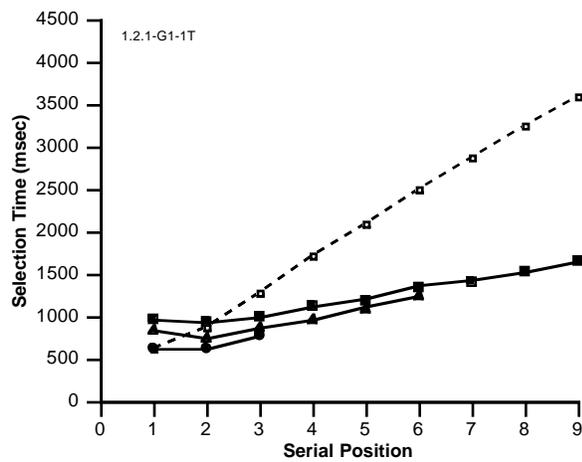
Figure 6. Selection times observed (solid lines) and predicted (dashed lines) by the Serial Processing Systematic Search model run with one item fitting into the fovea. The predicted times for the same serial position in different menu lengths are the same and are thus superimposed.

The results in Figure 6 suggest that this model is also wrong. The only feature in the observed data that this model accounts for is a positive slope greater than that of the predicted Fitts movement time. The model accounts for no other features in the observed data. Shorter menus are not faster. The slope of the predicted data is too steep. The selection time for serial position 1 is not higher than for serial position 2. This model does not account for the observed data.

The prediction has a slope resulting from more than just the mouse movement, but the predicted slope is too steep, about 380 msec per item as opposed to about 100 msec per item in the observed data. The discrepancy between the predicted and observed data results from all of the processing that must take place before moving the gaze to the next menu item. The slope of approximately 380 msec results because this is the time required for EPIC to move the eye, perceptually process a menu item, move the features to working memory, and decide on an item. Serially processing each item cannot produce a slope of 100 msec

per item. Only by processing multiple items at once can a model produce such a small slope.

The results provided by the serial processing models provide strong evidence that, when scanning a menu, people process more than one menu item at a time. The serial processing models asserted by Norman [12] and Vandierendonck, Van Hoe, and De Soete [14] are highly implausible. Menu selection models should take this human capability into consideration. The remaining models presented in this paper utilize parallel processing of menu items.

Parallel Processing Models
The parallel processing models represent a belief that people move their gaze across the menu as quickly as their perceptual-cognitive-motor processes allow, process the features of all objects that appear in the fovea in parallel using a "pipeline" facility to continue recognition even after the gaze has shifted away, and at the same time continually check working memory to see if the target item has yet been seen. As soon as the target item has been located, the person moves their gaze to it and clicks on it. In one of the parallel processing models, people search randomly for the target; in the other, they start at the top and scan down the menu.

Both parallel processing models were run with different eye-to-screen distances that resulted in one and three items fitting into the fovea simultaneously. When more than one item is visible in the fovea, all of those objects' features are sent to working memory in parallel. To prevent a random eye "movement" to essentially the same location while searching, both models choose the next item to look at from outside the fovea.

Parallel Processing Random Search Model
The Parallel Processing Random Search model was inspired by Card [3], who proposed that a random search model could account for menu selection times observed in an experimental task. Card concluded that people randomly decide which item in the menu to examine next. But note that Card's task was not a search task in which subjects are precued with the target item before timing starts. Rather, in Card's task the target item appeared above the menu at the same time that the menu itself appeared and at the same time that the clock started, perhaps combining a matching task with a search task. In Nilsen's task, modeled here, subjects were precued and could commit the target item to memory before initiating the timed portion of the trial. These are arguably different tasks, with Nilsen's more closely resembling a menu task in which the user knows the target item before opening the menu.

Figure 7 shows a flowchart that represents the production rules built in EPIC to investigate the possibility that subjects used a Parallel Processing Random Search strategy. To prevent a random eye "movement" to essentially the same location, the model chooses the next item from among all items currently outside the fovea.

Look at the precue.
Click the mouse on GO box to show menu and move eyes to a random location on menu.

These steps take place in parallel
- As soon as it is determined which items are not in the fovea, randomly choose one of the items and move eyes to it.
- Quit searching when target item appears in working memory.

Move cursor and eyes to item.
Click mouse.
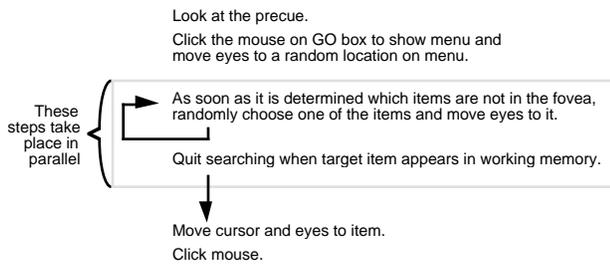
Figure 7. Parallel Processing Random Search model.

The results from running the Parallel Processing Random Search model are shown in Figure 8. Each predicted selection time is averaged from 300 trials run for that menu length and serial position combination.
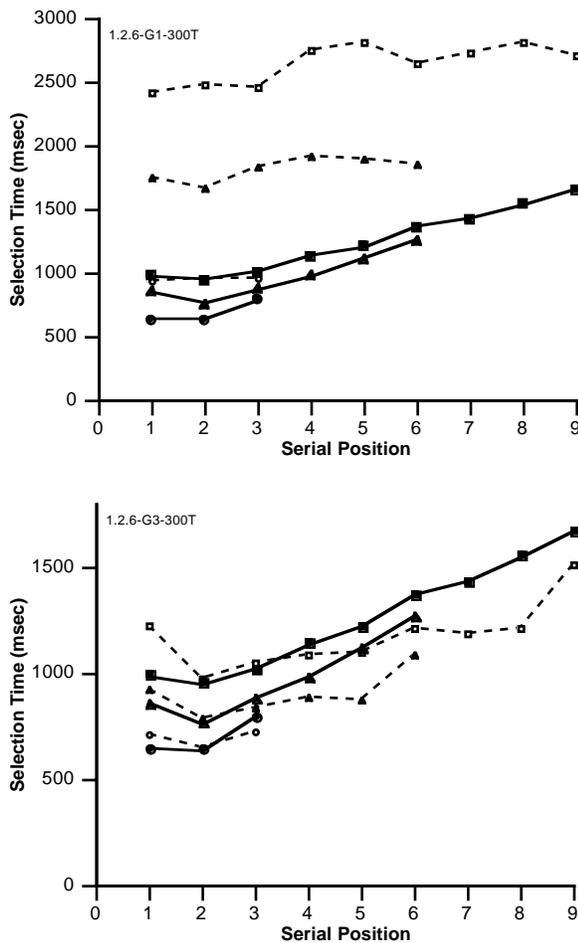
Figure 8. Selection times observed (solid lines) and predicted (dashed lines) by the Parallel Processing Random Search model run with one item (top graph) and three items (bottom graph) fitting into the fovea.

The predictions from the Parallel Processing Random Search model have some features that correspond to the observed data, but also have some problems.

As can be seen in Figure 8 (top graph), when one item at a time is visible in the fovea, the model accounts for shorter menus being faster, but no other features of the observed

data. The overall predicted times are, however, significantly lower than in the Serial Processing Random Search model discussed above.

As can be seen in Figure 8 (bottom graph), when three items are visible in the fovea simultaneously, the model can account for some features of the observed data Shorter menus are faster, and about the right amount faster, as is shown by the distance between the predicted lines approximating the distance between the observed lines. The predicted values fall entirely within the range of the observed values. Most importantly, this model accounts for serial position 1 being higher than serial position 2. However, the overall slope is still too small.

In Figure 8 (bottom graph), both the first and last serial positions are higher because the model combines random search with three menu items fitting into the fovea. Items at both ends of the menu have a lower probability of being in the fovea after any random fixation. Any of the middle menu items can be foveated by moving the eye to that item, or to either of the two adjacent items. But the first and last items only have one adjacent item. This might explain serial position 1 being higher than serial position 2 in the observed data.

The predictions from the Parallel Processing Random Search model suggest that the model is partly correct, and partly incorrect.

Parallel Processing Systematic Search Model
Figure 9 is a flowchart that represents the production rules built in EPIC to investigate the possibility that subjects used a Parallel Processing Systematic Search strategy. Though other systematic searches are possible, top-to-bottom is the one most commonly proposed.

In this model, the first eye movement is made to any of the items that are within one foveal radius from the topmost item (to insure the first gaze captures the topmost item). Each subsequent movement is made to an item one foveal diameter below the center of the current fixation. These details represent the belief that, when using a systematic search strategy, people attempt to maximize the foveal coverage with a minimum number of eye movements.

Look at the precue.
Click the mouse on the GO box to show the menu and move eyes to where one of the top items will appear.

These steps take place in parallel.
- Determine the item one foveal diameter below gaze.
- Move eyes to that item.
- Quit searching when target item appears in working memory.

Move mouse and gaze to item.
Click mouse.
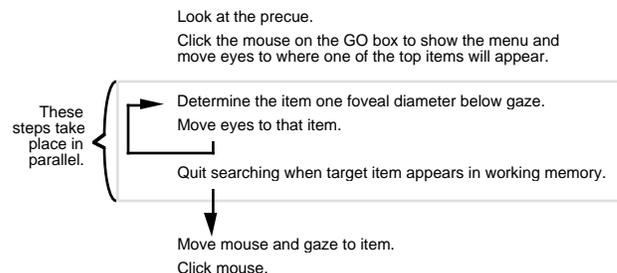
Figure 9. Parallel Processing Systematic Search model.

The results from running the Parallel Processing Systematic Search model are shown in Figure 10. Each predicted selection time is averaged from one trial run for each possible combination of menu length, serial position, and first eye movement.

1.2.7-G1-1T

(Top graph: Selection Time (msec) vs Serial Position)

1.2.7-G3-1T-S1/2

(Bottom graph: Selection Time (msec) vs Serial Position)
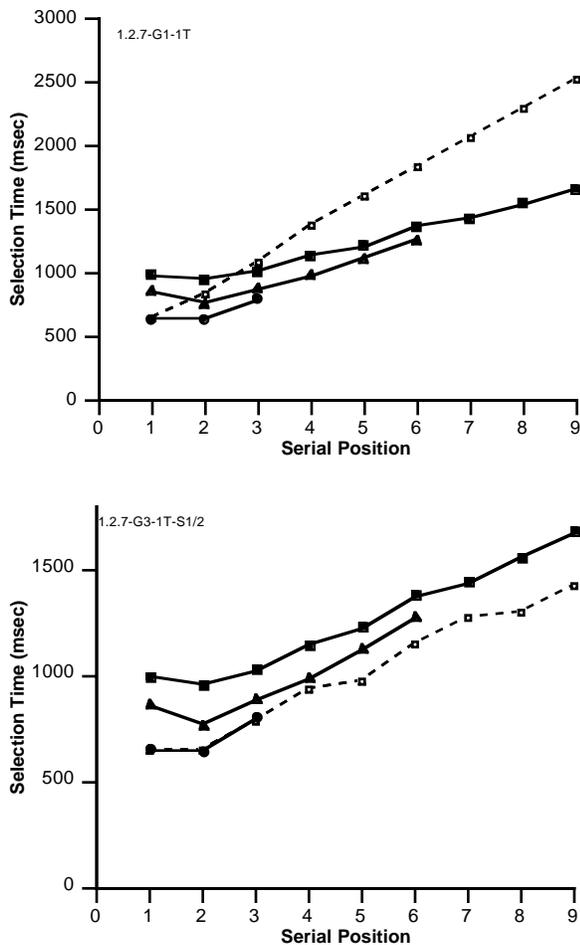
Figure 10. Selection times observed (solid lines) and predicted (dashed lines) by the Parallel Processing Systematic Search model run with one item (top graph) and three items (bottom graph) fitting into the fovea. In each graph, the predicted times for the same serial position in different length menus are the same and are thus superimposed.

The predictions from the Parallel Processing Systematic Search model have some features that correspond to the observed, but also have some problems.

As can be seen in Figure 10 (top graph), when one item at a time is visible in the fovea, the model only accounts for a positive slope. The model does not predict that shorter menus will be faster, the slope is too steep, and serial position 1 is not higher.

As can be seen in Figure 10 (bottom graph), when three items are visible in the fovea simultaneously, the model can account for important features of the data. The slope is correct and the predicted values fall entirely within the range of the observed values. But again, the model does not account for shorter menus being faster, and serial position 1 is not higher.

These results show that the Parallel Processing Systematic Search model can partially explain how the subjects

accomplished the task, but not account for all aspects of the observed data.

None of the models presented thus far can account for all of the features in the observed data. The serial processing models account for essentially none of the features of the observed data. But all features of the observed data are accounted for by at least one of the various parallel processing models, as shown in Figure 11.

Search Strategy

		Random	Systematic
	One	Figure 8 (top graph) **+** Shorter menus are faster. **—** Slope is too small. **—** Serial position 1 is not higher.	Figure 10 (top graph) **—** Shorter menus are not faster. **—** Slope is too big. **—** Serial position 1 is not higher.
	Three	Figure 8 (bottom graph) **+** Shorter menus are faster. **—** Slope is too small. **+** Serial position 1 is higher.	Figure 10 (bottom graph) **—** Shorter menus are not faster. **+** Slope is correct. **—** Serial position 1 is not higher.

(Left vertical axis label: **Number of Items That Fit Into Fovea**)

Figure 11. Summary of how the parallel processing models account for (+) and do not account for (−) features in the observed data.

Hybrid Models

The hybrid models represent a belief that, when Nilsen ran his experiment, 1) subjects used both random *and* systematic search, and 2) screen-to-eye distance varied across trials.

These models were motivated by observing, as shown in Figure 11, that all of the features in the observed data are accounted for by at least one of the parallel processing models when run one or three items fitting into the fovea. The random search model accounts for faster selection times in shorter menus. When three items fit into the fovea, the random search model also accounts for serial position 1 being higher. The systematic search model accounts for the correct slope when three items fit into the fovea.

Dual Strategy Hybrid Model

The Dual Strategy Hybrid model represents the belief that subjects processed menu items in parallel in all of the observed trials, but that subjects searched randomly in half of the trials and systematically in half of the trials. Such a model could accurately account for the observed data if 1) some subjects searched randomly and others systematically, or 2) subjects varied their search strategy from trial to trial. Since the observed data were averaged across subjects and blocks, either scenario would produce the same results.

Predictions from this hybrid model can be obtained in two ways. The first is to build a set of EPIC production rules that contain the rules from both the Parallel Processing Random Search strategy and the Parallel Processing Systematic Search strategy; the strategy would randomly choose which search strategy to use at the start of each trial. The second is to average the predicted values produced by running the two models independently. Since both approaches would produce the same predictions, the second approach was chosen for expedience. Figure 12 shows the results of this model, as determined by taking an unweighted average of the results shown in Figure 8 and Figure 10.
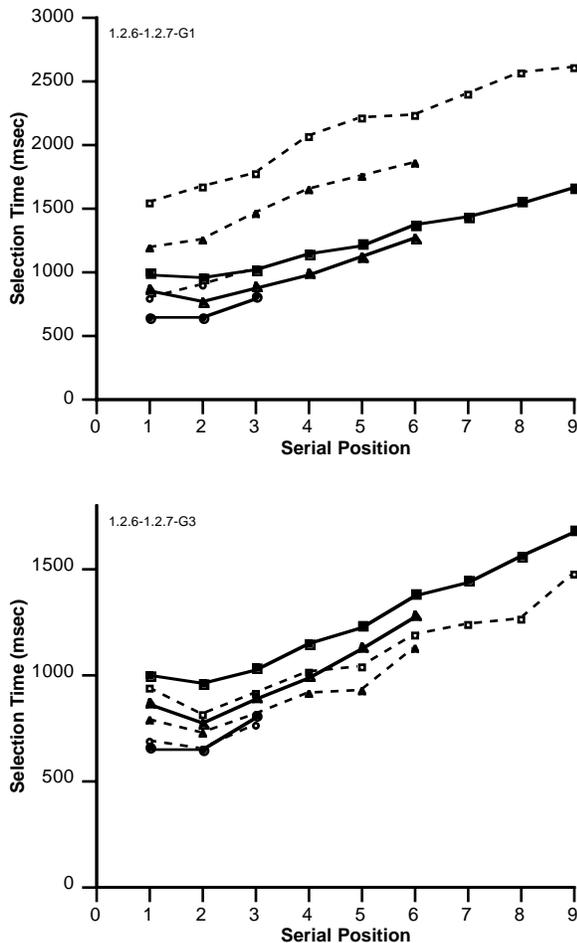
Figure 12. Selection times observed (solid lines) and predicted (dashed lines) by the Dual Strategy Hybrid model, with one item (top graph) and three items (bottom graph) fitting into the fovea.

The predictions from the Dual Strategy Hybrid model can account for most of the features in the observed data, but do not fit the observed values perfectly.

As can be seen in Figure 12 (top graph), when one item fits into the fovea, the model accounts for faster selection times in shorter menus and produces a near-perfect slope. But the model does not account for the higher selection time in serial position 1, and overall the predicted values are higher than the observed values.

As can be seen in Figure 12 (bottom graph), when three items fit into the fovea, the model accounts for faster selection times in shorter menus, produces a comparable slope, accounts for the higher selection time in serial position 1, and predicts values that are in range of the observed data. The only shortcoming of this model is that the predicted values do not exactly match the observed values.

The predictions from the Dual Strategy Hybrid model suggest that the model is almost correct.

Dual Strategy Varying Distance Hybrid Model
The Dual Strategy Varying Distance Hybrid model represents a belief that subjects performed the menu selection task as asserted by the Dual Strategy Hybrid model and that the screen-to-eye distance varied across trials. Since this distance was not controlled or measured during the experiment, it is very likely that some subjects sat closer to the computer screen than others, and that subjects moved nearer to and further from the screen during the course of the experiment.

Predictions from this hybrid model can be obtained in two ways. The first is to build a task environment that varies the screen distance from trial to trial, and to run a set of production rules developed for the Dual Strategy Hybrid model using this task environment. The second is to average the predicted values produced by running the Dual Strategy Hybrid model in two task environments, each with a fixed screen-to-eye distance. Since both approaches would produce the same predictions, the second approach was chosen for expedience. Figure 13 shows the results of this model, as determined by taking a weighted average of the results shown in the two graphs in Figure 12, with 15% from the top graph (one item in fovea) and 85% from the bottom graph (three items in fovea).
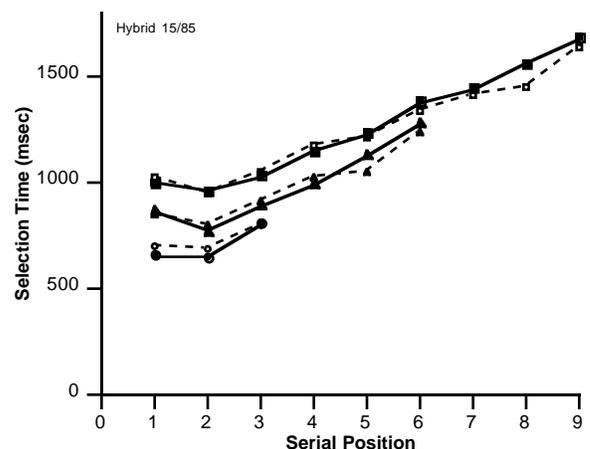
Figure 13. Selection times observed (solid lines) and predicted (dashed lines) with a Dual Strategy Varying Distance Hybrid model, with 15% of the trials at a one-item-in-fovea distance, and 85% of the trials at a three-items-in-fovea distance.

The Dual Strategy Varying Distance Hybrid model accounts for all of the features in the observed data. As can be seen in Figure 13, the model predicts the observed values very well ($r^2 = 0.99$). Matching the observed values, the Dual Strategy Varying Distance Hybrid model offers a highly plausible explanation of the task environment and strategies used by subjects in Nilsen's experiment.

CONCLUSION

The models presented here provide evidence that 1) people do not stop and decide on menu items individually, but rather process many items in parallel, and 2) people search menus using both systematic top-to-bottom and random visual search strategies. The models presented here provide a plausible explanation of the low-level perceptual, cognitive, and motor processing that people use when they select an item from a menu. Having validated aspects of these models with empirical data, these models provide strong evidence that previously asserted serial processing theories [12, 14] are incorrect, and that there is an element of truth to *both* previously asserted random search models (such as in [3]) *and* previously asserted systematic search models (such as in [9]). Perhaps unmeasured or unreported factors in the experiments biased subjects in one experiment towards random search and in another experiment towards systematic search, thus giving rise to these conflicting theories.

FUTURE WORK

Future plans include to attempt to explain Nilsen's observed data for ordered menus with an ordered menu selection model.

Also looking to the future, successfully modeling menu search provides evidence that a general purpose tool for evaluating the efficiency of visual aspects interfaces might be feasible. The tool would take as its input a definition of a screen layout and a task. The tool would provide as output a prediction of the time required for the user to execute the task. Previous researchers have set a precedent that such a tool can be built [10, 13]. Such a tool would analyze screen layouts and predict the cognitive effort required by a user to extract the information needed to accomplish a task.

ACKNOWLEDGMENTS

REFERENCES

1. Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

2. Bovair, S., Kieras, D. E., & Polson, P. G. (1990). The acquisition and performance of text editing skill: A cognitive complexity analysis. *Human-Computer Interaction*, 5, 1-48.

3. Card, S. K. (1984). Visual search of computer command menus. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of Language Processes,* (pp. 97-108). London: Lawrence Erlbaum Associates, Publishers.

4. Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

5. John, B. E., & Kieras, D. E. (1994). The GOMS family of analysis techniques: Tools for design and evaluation (Technical Report No. CMU-CS-94-181): Carnegie Mellon University School of Computer Science.

6. Kieras, D. E., & Meyer, D. E. (1995). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction (EPIC Tech. Rep. No. 5, TR-95/ONR-EPIC-5). Ann Arbor, Michigan: Department of Electrical Engineering and Computer Science.

7. Kieras, D. E., & Meyer, D. E. (in press). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction.*

8. Laird, J., Rosenbloom, P., & Newell, A. (1986). *Universal subgoaling and chunking*. Boston: Kluwer Academic Publishers.

9. Lee, E., & MacGregor, J. (1985). Minimizing user search time in menu retrieval systems. *Human Factors*, 27(2), 157-162.

10. Lohse, J. (1991). A cognitive model for the perception and understanding of graphs. In *Proceedings of CHI '91*, New Orleans, Louisiana. New York: ACM.

11. Nilsen, E. L. (1991). Perceptual-motor control in human-computer interaction (Tech. Rep. No. 37). Ann Arbor, Michigan: The Cognitive Science and Machine Intelligence Laboratory, The University of Michigan.

12. Norman, K. L. (1991). *The Psychology of Menu Selection: Designing Cognitive Control of the Human/Computer Interface*. Norwood, N. J.: Ablex.

13. Sears, A. (1993). Layout appropriateness: A metric for evaluating user interface widget layout. *IEEE Transactions on Software Engineering*, 19(7).

14. Vandierendonck, A., Van Hoe, R., & De Soete, G. (1988). Menu search as a function of menu organization, categorization and experience. *Acta Psychologica*, 69(3), 231-248.