Exploring the Divide Between Two Unified Theories of Cognition: Modeling Visual Attention in Menu Selection

Erik Nilsen
Lewis & Clark College
Portland, OR 97219, USA
+1 503 768 7657
nilsen@lclark.edu

Jake Evans
Lewis & Clark College
Portland, OR 97219, USA
+1 503 768 5136
jce@lclark.edu

ABSTRACT
Two cognitive modeling efforts (EPIC & ACT-R) have proposed computational models of a simple menu selection task involving searching for a single digit in an unordered, pull down menu. This paper presents an empirical study which extends the menu selection task in two dimensions (distance between menu items and whether the items are digits or words). Each of these manipulations should make a difference in selection time according to one of the models but not the other. An analysis of response times reveals that both factors produce significant differences in the direction predicted by the cognitive models. The magnitude of these differences, however, are smaller than predicted (7% for distance and only 3% for word vs. digit). Implications for future modeling of visual attention is briefly addressed.

KEYWORDS: Menus, cognitive models, visual search.

INTRODUCTION
Cognitive modeling is alive and well in the field of human computer interaction. Two modern computational cognitive architectures are addressing issues in user interface design as well as contributing to psychological theory development as Unified Theories of Cognition (UTC).

EPIC (Executive Process Interactive Control) [5] and ACT-R (Adaptive Control of Thought - Rational) [1] have much in common. They are both production-rule based computational models. Both are committed to developing their programming architecture to have psychological plausibility in line with the best available theoretical understanding. Both use basic and applied empirical task domains to validate and extend the coverage of their models.

One HCI task domain that both ACT-R [1] and EPIC [3] have studied is menu selection. In fact, they have both modeled the same data [6] involving selection from randomly ordered, pull down menus. Two papers presented at this years CHI conference [2,4] attest to their continuing interest in this task domain.

EPIC and ACT-R models of this menu selection task differ markedly in the details of the visual search process. ACT-R posits a systematic, top to bottom search with eye fixations on menu items sharing features with the target item [1]. EPIC posits a 50/50 blend of random and systematic search with eye fixations determined by the number of menu items visible in the fovea [3] (defined as a one degree radius around the fixation point).

This paper presents an empirical study which extends the menu selection task in two dimensions. Each of these is predicted to have a significant impact on both the selection time and search process by one of the models.

Increasing the distance between the menu items should have a large impact according to EPIC since the number of menu items located in the fovea is central to their models. The item distance in this study place either 1 or 3 menu items within EPIC's fovea. Selection times with one item in the fovea should take 66% longer than with three items [3].

Changing the menu items from single digits to words should impact selection time and visual search according to ACT-R. Unfortunately, precise predictions are not possible since ACT-R is ambiguous concerning what features are attended to in the word condition. However, it is most certainly the case that ACT-R predicts that the feature set for words is more complex than for digits, thus taking more time to process.

METHOD
Subjects
Eight computer literate subjects from the local college community were paid 15 - 20 dollars for participation (based on performance) in a ninety minute experimental session.

Equipment
The tasks were performed using two PC's. A 386 computer ran the experimental software presenting the menus and collecting the response times. Another computer controlled eyetracking hardware and software. Due to space limitations the eye movements will not be discussed here. Subjects sat 58 cm away from a 15 inch monitor (640 x 480). Subjects used a Microsoft InPort mouse to select from the menus.

Task
Subjects made a total of 1280 rapid menu selections from eight different menu types. All of the menus consisted of six items. Figure 1 shows two of the menu types containing the manipulations used in the experiment. The first dimension reflects the organization of the menus. Half of the menus were presented in a fixed numeric order while the rest were presented in a different random order on each trial. Another manipulation was the menu item content. Half of the menus contained the single digits 1-6 while the remainder contained these same digits spelled out as words.

Finally, the vertical distance between menu items was manipulated. Short menus had items separated by 1 cm while tall menu items were separated by 2 cm.

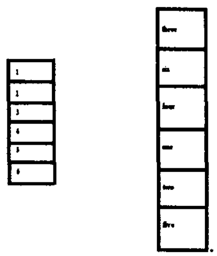

Figure 1. Two of the eight menu types in the experiment: Short fixed order digits and tall random order words

For every menu selection, the target item was displayed by itself. After memorizing it, subjects initiated the trial by clicking in a "go" box. A six item menu appeared directly below the cursor position. Selection time was recorded from the time the subject clicked on the go box until they clicked on a menu item. Blocks of 20 trials at a time were run including 2 practice trials and 3 selections at each of six menu positions. Correctness feedback was given after every trial and average times and error rates were shown after each block. Each block presented a single menu type.

RESULTS AND DISCUSSION

The analyses are based on the selection times for 1008 menu selections at asymptotic performance. Repeated Measures ANOVA shows that the strongest effect was the organization manipulation $F(1,7) = 196.6$, $p<.0001$. Fixed order menus (652 msec.) were much faster to select from than random order menus (1060 msec.). The distance between menu items impacted selection time $F(1,7) = 69.5$, $p<.0001$. Short menus (825 msec.) were faster to select from than tall menus (887 msec.). The main effect of menu content was also significant $F(1,7) = 12.1$, $p<.01$. Digit menus (842 msec.) were faster to select from than word menus (870 msec.). The EPIC and ACT-R models [3,1] pertain to the data in Figure 2 which shows the random ordered menu data separated out by serial position.

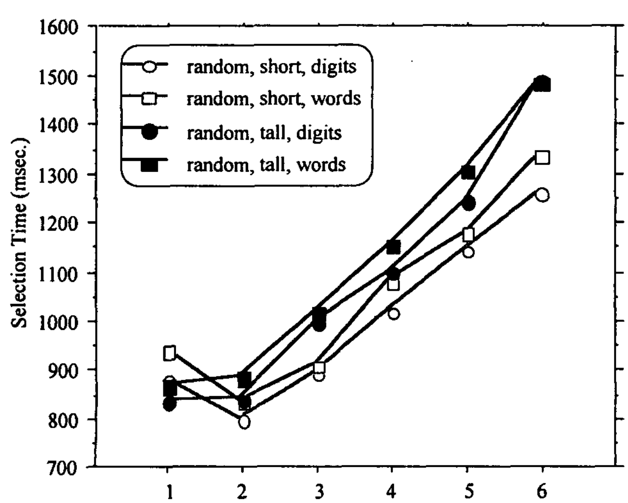

Figure 2. Serial position graph for the random order menus.

There is a strong linear trend in the data for positions 2 - 6 for all menu types. The linear trend is modeled by both EPIC and ACT-R, the divergence of menu item 1 is only handled by EPIC. Selection from the tall menus (filled symbols) is consistently slower than from short menus (unfilled symbols). However the magnitude of the difference is only 7%. Apriori EPIC models predict approximately a 66% difference. One implication of this data is that menu items can be processed in parallel beyond the 1 degree radius assumed by EPIC. Future work can address the extent of visual attention and the steepness of the gradient of information loss in parafoveal vision.

The difference between words (squares) and digits (circles) is small and varies by serial position. While ACT-R predictions are not readily available, it seems reasonable to posit that the higher visual complexity and greater feature overlap among words should lead to a steeper slope for the menus containing words. This is not seen in the data.

Both EPIC and ACT-R are at the early stage of model development. Implications for user interface await future extensions and validation on a greater range of tasks. This study provides new data to model, while challenging assumptions made by the first round of modeling menu selection. EPIC's fovea might need to extend over a larger area. ACT-R needs to address how menus with such different feature sets can lead to such similar selection times. Future work should make the menu items more realistic, and place the menu selection task in the context of achieving some procedural goal.

REFERENCES

1. Anderson, J.R., Matessa, M., and Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction, 12*,311-343.

2. Byrne, M. D., Anderson, J. R., Douglass, S., & Matessa, M. (1999). Eye tracking the visual search of click-down menus. To appear in *Proceedings of CHI 99.*

3. Hornof, A. J., & Kieras, D. E. (1997). Cognitive modeling reveals menu search is both random and systematic. *Proceedings of CHI '97*, New York: ACM, 107-114.

4. Hornof, A. J., & Kieras, D. E. (1999). Cognitive modeling demonstrates how people use anticipated location knowledge of menu items. To appear in *Proceedings of CHI 99.*

5. Kieras, D. E. & Meyer, D.E.(1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction, 12*,311-343.

6. Nilsen, E. (1991) *Perceptual-motor control in human-computer interaction.* (Technical Report 37). Ann Arbor: University of Michigan, Cognitive Science and Machine Intelligence Laboratory.