A New Model of Menu Search

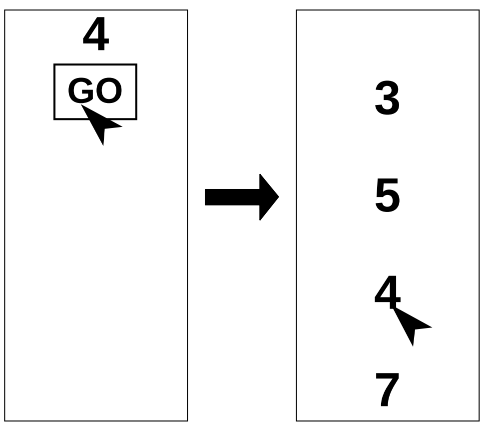
Michael D. Byrne
Department of Psychology
Rice University, MS-25
Houston, TX 77005
byrne@acm.org

John R. Anderson
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

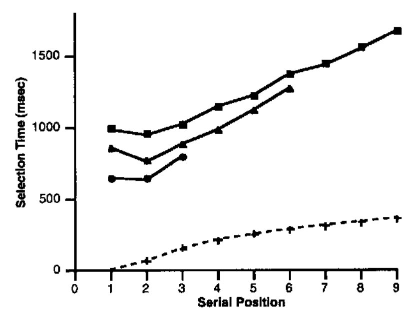
Overview

- Nilsen experiment
- Old models
  - ACT-R with Visual Interface
  - EPIC
- New data
  - Nilsen & Evans (1998)
  - Byrne, Anderson, Douglass, & Matessa (1999)
- Old Models (yes, again)
- New model (ACT-R/PM)
- Other issues

Menu Selection (Nilsen, 91)



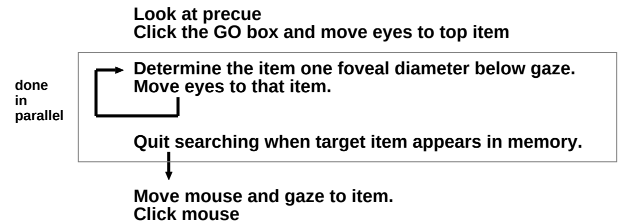
Menu Selection (Nilsen, 91)



EPIC Model (Hornof & Kieras, 97)

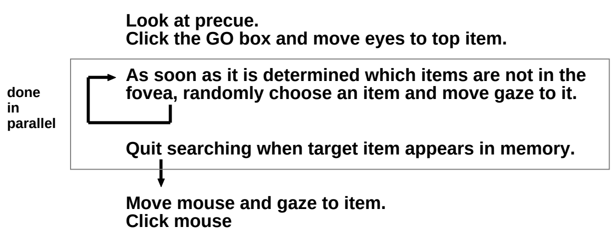
- "Parallel Processing Dual Strategy Varying Distance Hybrid Model"
- Assumptions:
  - 50/50 split between systematic top-down search and random search
  - Mix of one item in fovea and three items in fovea
  - Multiple items examined in parallel
  - No mouse movement until target has been detected, one aimed movement made

EPIC: Systematic Search



- Always starts at item 1
- Never skips items
- Gaze will usually overshoot

EPIC: Random Search



- Initial fixation equally likely to be to any menu item
- Up and down saccades equally likely

ACT-R Model (Anderson, Matessa, & Lebiere 97)

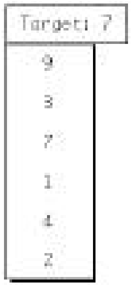
- Assumptions:
  - Model selects basic feature from target character
  - Attention shifts strictly top-to-bottom
  - Skips characters that do not share feature
  - Mouse moves with attention
- Predictions:
  - Initial fixation on first item with shared feature
  - Some items are skipped
  - Eye never overshoots target
  - Also predicts effects of letter vs digit search with letter vs. digit distractors

Nilsen & Evans 99

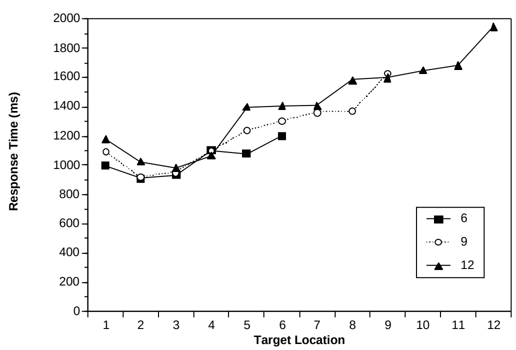
- New experiment to evaluate EPIC and ACT-R models
- EPIC predicts large effects of menu spacing
- ACT-R predicts large effects of features
- New experiment:
  - Large vs. small spacing
  - Search for (and amongst) single characters vs. words
- Results:
  - Reliable but small effect of spacing
  - Reliable but small effect of character vs. word search

Eye Tracking (byrne, Anderson, Matessa, & Douglass 99)

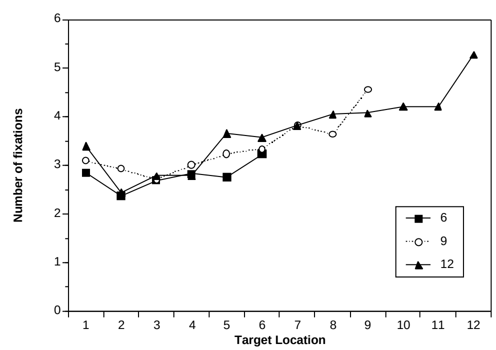
- 6, 9, and 12 menu lengths
- 11 participants, 108 trials after 36 practice trials
- ISCAN RK726/RK520 Pupil/CR tracker with magnetic head tracking
- Recorded point-of-regard (POR) and mouse position at 120 Hz (every ~8 ms)
- Characters ~0.34° high, ~0.69° apart
- Fixations identified on the basis of velocity



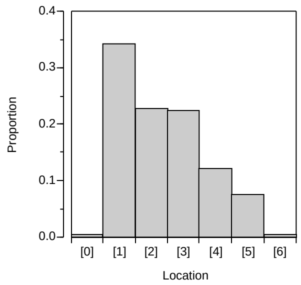
Results: Response Time

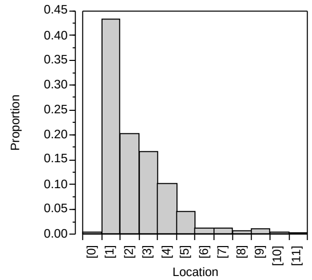


Results: Number of Fixations

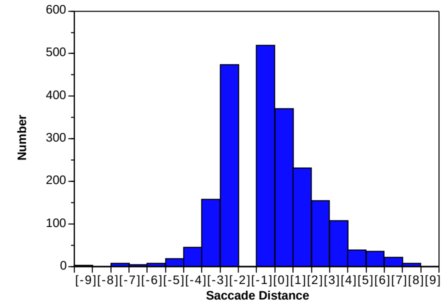


Resu l ts: Initial Fixation





Resu l ts: Non-initial Saccades



The Mode l s, Revisited

- Despite good fits to original Nilsen RT data, neither model fares well with respect to the new data
- What should the “right” model look like?
  - Initial fixation to one of the first three items, usually item 1
  - Mostly, but not exclusively, top-to-bottom
  - Some items are skipped
  - Most bottom-to-top is a result of the target item being skipped
  - Would still predict number vs. letter effects

New ACT-R/PM Mode l (in progress)

- Select a feature in the target item
- Click the menu open
- Make an initial attention shift
  - Not sensitive to feature content
  - Productions that encode the first five locations compete
  - Conflict resolution preference mirrors initial fixation preferences in the data
- Further attention shifts again a mix of competing productions
  - Matching item, up and nearby
  - Matching item, down and nearby
  - Matching item, anywhere down
  - Random item in any direction (more on this later)

Issue: Visual Features

- Current ACT-R/PM uses the same Rumelhart & McClelland feature set as the old Visual Interface
- LED-style representation of characters from 16 features:

```
 - -
|\|/|
 - -
|/|\|
 - -
```

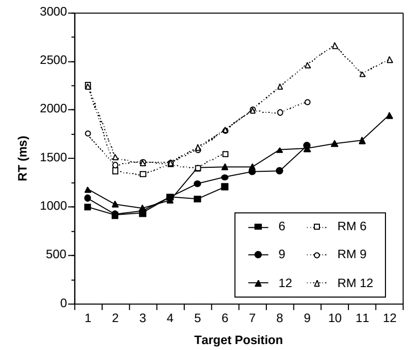
- So, for example, the letter "E":

```
 - -
|
 - -
|
 - -
```

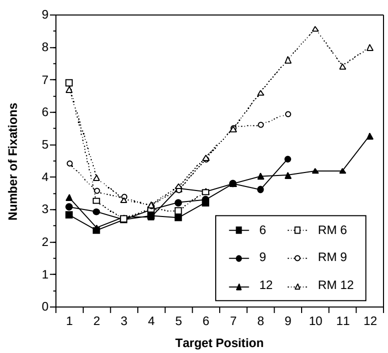
- Old Visual Interface could accurately search in peripheral vision for any one of the 16
- Problems:
  - ►Characters don't actually look like this (curvature seems important)
  - ►Discriminations seem too fine

First fit

- Represented only part of the features (horizontal vs. vertical vs. diagonal)
- Slightly simpler production set (nearest up vs. nearest down only)



Why So Bad? Too Many Misses

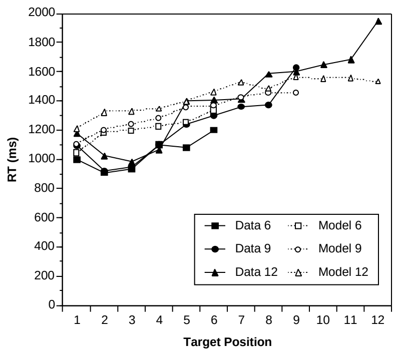


- Also wasn't capturing right distribution of saccades (not enough backtracking)

New Feature Set (Briggs & Hocevar 75)

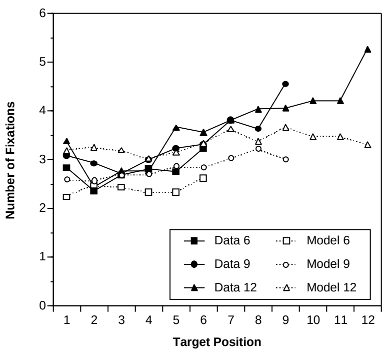
- Also 14 features
  - ►Horizontal (top, center, bottom)
  - ►Vertical (single and double)
  - ►Angle (open top, down, & horizontal)
  - ►Small curve (convex right, bottom, & left)
  - ►Large curve
  - ►Closed curve
  - ►Continuous curve
- Examples:
  - ►N: double vertical, angle down, angle top
  - ►O: large curve, continuous curve, closed curve
  - ►5: small curve convex left, horizontal top

Current Fit



- Somewhat better
- Still too slow in places, too fast in others
- Not capturing slower initial item

Current Fit (2)



- Not enough fixations for late positions, long menus
- Too much menu length effect

Other Unresolved Issues

- Mouse control
  - EPIC model is wait until target is found
  - Old ACT-R model says move mouse with eyes
  - Neither predicts mouse overshoots
  - Data say something in between, but with overshoots
- Are people really sensitive to feature matches?
  - Marginal evidence in eye-tracking data
  - Number vs. letter search effect is robust, however
  - What's the "right" feature set?
- How to best use eye tracking data to improve fits?
  - Are fixation durations necessary?
  - Relationship between fixations and attention shifts?
  - Dario's HMM parser may be the key