Predicting Situation Awareness Errors Using Cognitive Modeling

Troy Kelley
Army Research Laboratory
ACT-R Summer School

Human Error Modeling

- Potentially a very important area for cognitive modeling
- Typical analyses in the past were descriptive taxonomies of human error with little predictive value
  - Types of errors don't help much with prediction
- Can force simple hypothesis
  - So simple might not want to present to ACT-R workshop

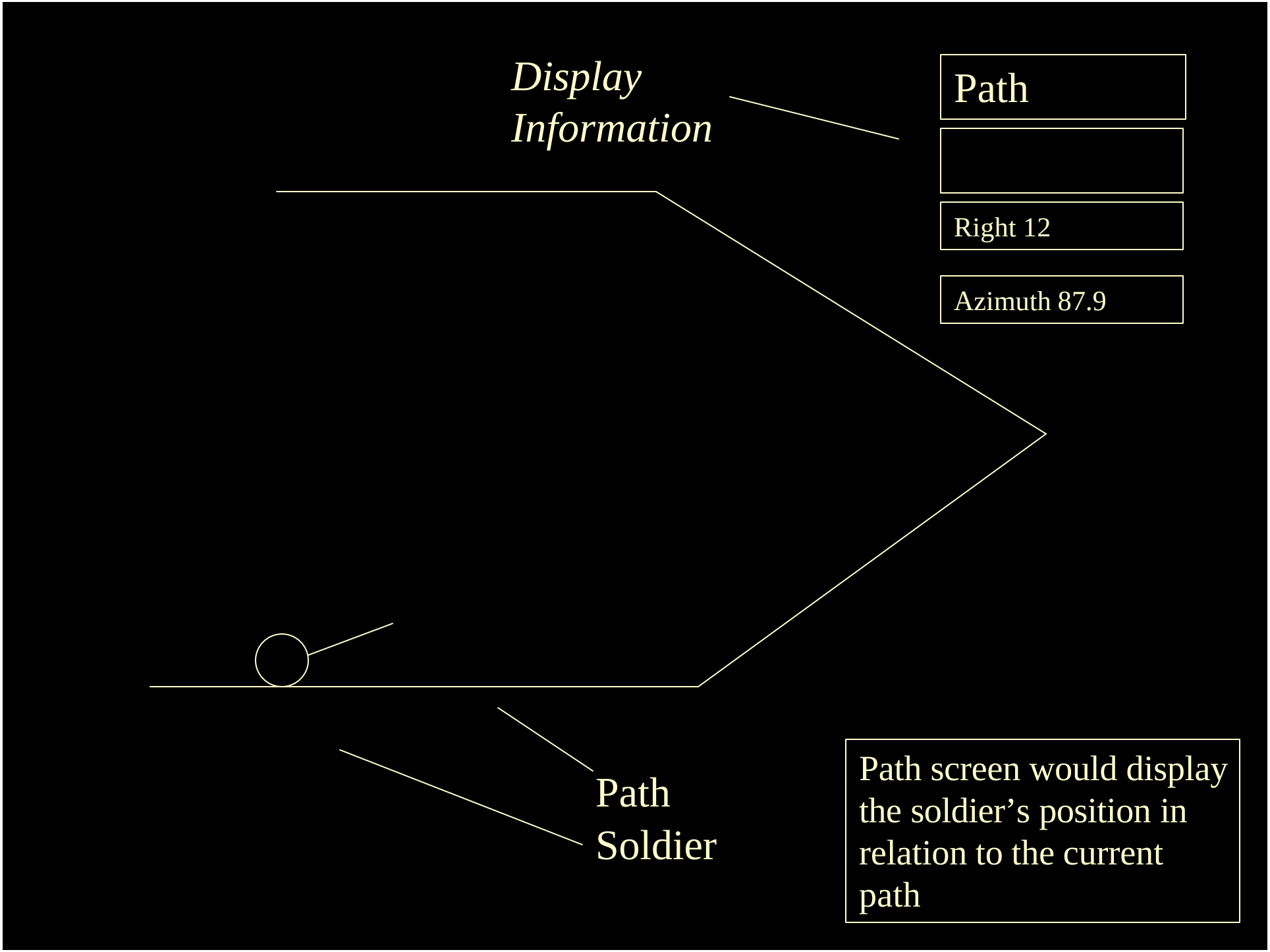
Model of a Navigation Task

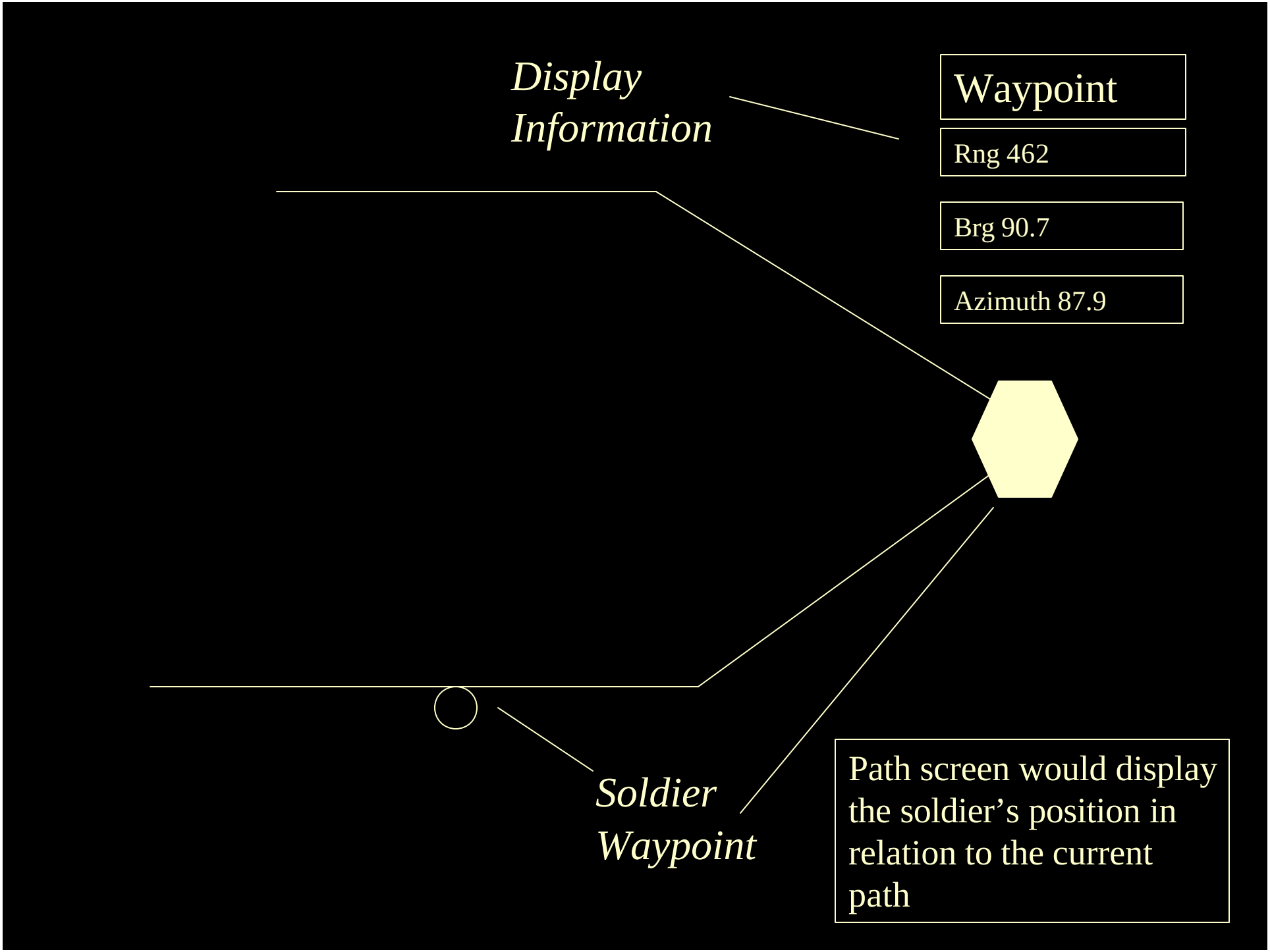
- Task: Navigate to waypoints and identify targets with as little deviation from a path as possible and answer situation awareness probe questions
- Equipment: Helmet mounted display (HMD) showing waypoint, target, unit and path information
- Model: Error data from probe questions*

*Data taken from: Glumm, M. M., Marshak, W. P., Branscome, T. A., Wesler, M. M., Patton, D. J., Mullins, L. L. (1998). A Comparison of Soldier Performance Using Current Land Navigation Equipment with Information Integrated on a Helmet-Mounted Display, (Technical Report ARL-TR-1604). Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Land Warrior System


TREKKER


Display Information
Path
Right 12
Azimuth 87.9
Path
Soldier
Path screen would display the soldier's position in relation to the current path


Display Information
Waypoint
Rng 462
Brg 90.7
Azimuth 87.9
Soldier
Waypoint
Path screen would display the soldier's position in relation to the current path

Probe Questions

- Attempting to predict the likelihood of a correct answer
- 20 Questions total (one question was thrown out)
- Screens were blanked before answering the each probe question
- Yes/No format
  - Are you within 50 meters of your next target?
  - Are there friendly units only to the left of your path?

Format of Probe Questions

- Yes or no format of questions made additional analysis difficult
- Additional analysis might pertain to what factors might influence right or wrong answers
- Difficulty of being a data parasite

Hypothesis

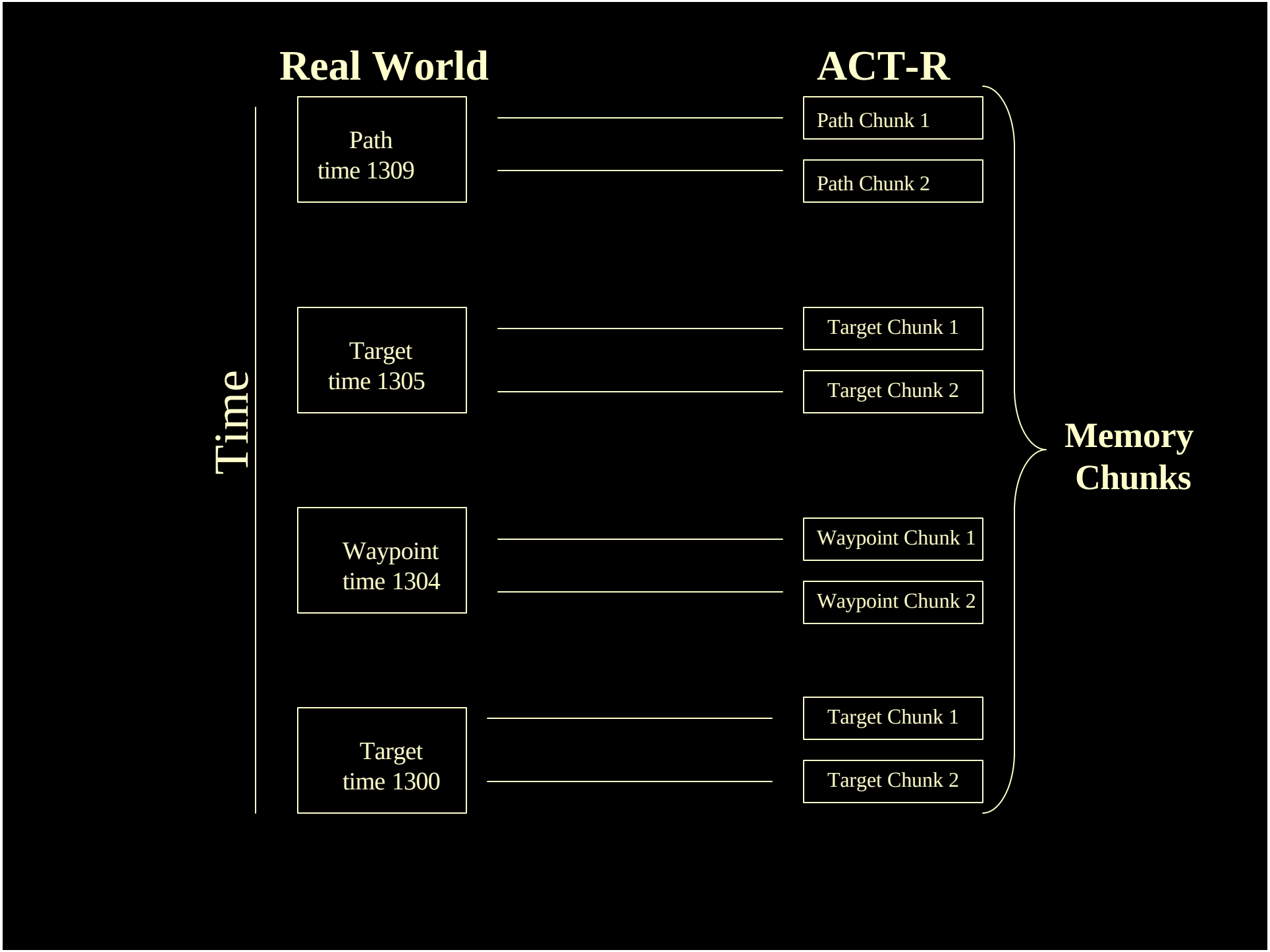
- Memory for HMD screens would affect current situation awareness
- Decay of memory for HMD screens would cause errors in situation awareness and errors on probe questions

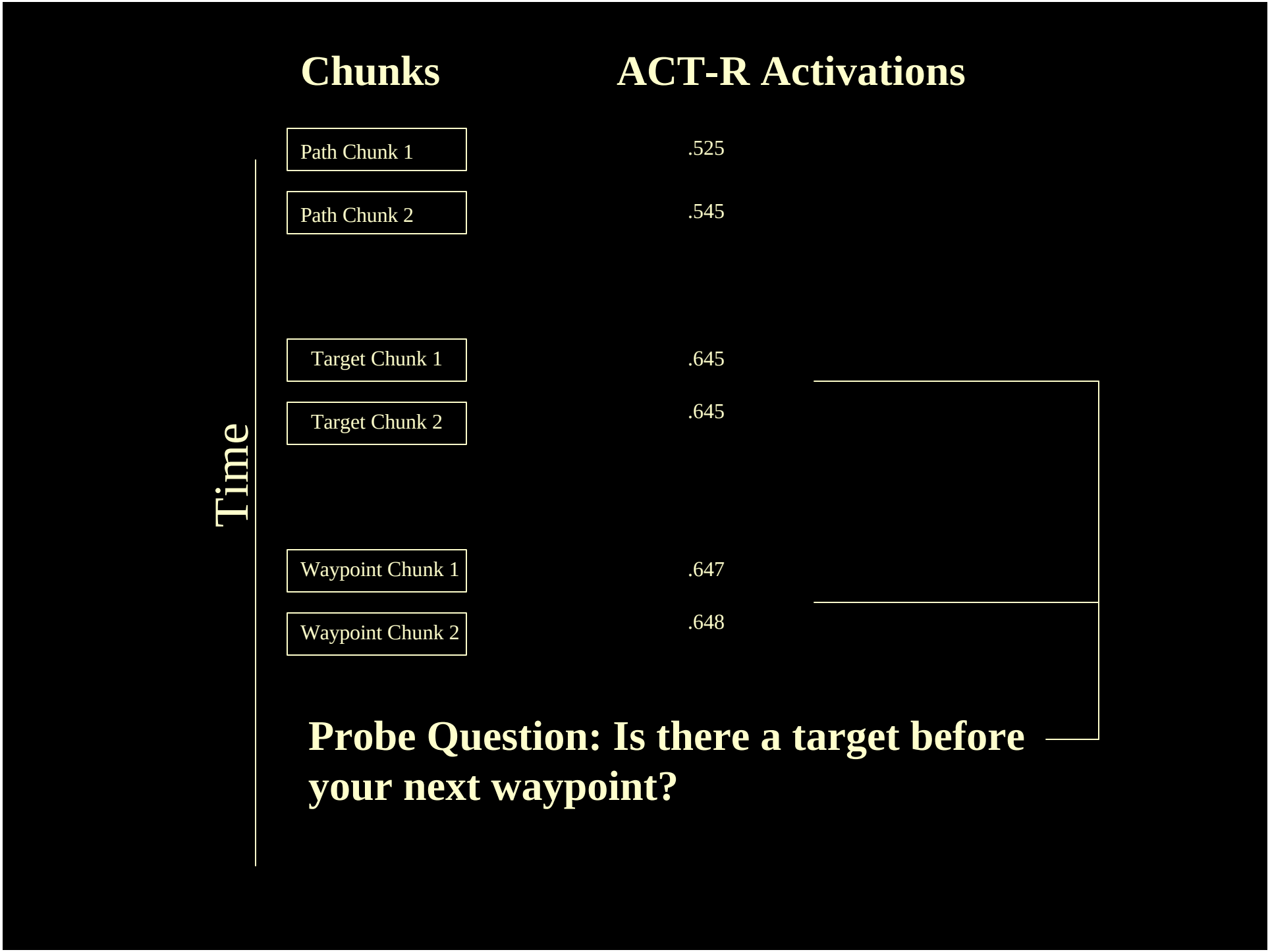
Memory Elements for ACT-R Model

- Screens
  - Path, Target, Waypoints, Units
- Unit
  - Separate memory elements for friendly and enemy
  - Specific unit information: (e.g., location)
- Pace Count
  - Technique used by infantry soldiers to calculate distances based on number of steps they have taken
  - A constantly updated memory element in model
- Auditory information pertaining to troop movements

ACT-R Memory Elements

- Strength of Memory (Activation Levels)
  - Subject to decay
  - Subject to spreading activation
    - Relationship between related memory elements
      - (Friendly and enemy unit information are related)


Real World
ACT-R
Time
Path
time 1309
Path Chunk 1
Path Chunk 2
Target
time 1305
Target Chunk 1
Target Chunk 2
Waypoint
time 1304
Waypoint Chunk 1
Waypoint Chunk 2
Target
time 1300
Target Chunk 1
Target Chunk 2
Memory
Chunks


Chunks
ACT-R Activations
Path Chunk 1
.525
Path Chunk 2
.545
Target Chunk 1
.645
Target Chunk 2
.645
Waypoint Chunk 1
.647
Waypoint Chunk 2
.648
Time
Probe Question: Is there a target before your next waypoint?

Data Analysis

- Models of *each* question for *each* subject (140 models) were run 40 times
- This produced retrievals of multiple memory chunks for each question, these were averaged.
  - This was done because the individual experience of each soldier was different, since they could retrieve one of the 4 screens any time they wanted
- Activations from 40 runs were averaged across runs, across memory chunks and across subjects to yield a single activation level for each question.

	Sub 1	Sub 2…….. Sub 20	Average	Average2
Question 1				
Target	.770[1]	.740 .781	.760	
Pace	.540	.531 .530	.531	.645[2]
Question 2				
Friendly	.441	.440 .431	.440	
Direction	.320	.352 .330	.332	
Unit-movement	.470	.440 .451	.450	.406

1 Average for 40 runs of the simulation

2 Average across subjects and across questions

Results

- One tailed Pearson's product moment correlation
- There was a significant negative correlation between activation levels and the percentage of errors for each question
- Higher activations levels resulted in a lower percentage of errors
- .$r(19) = -.43$, $p < .03$

Conclusions

- Difficult to predict human performance
  - Multiple interacting parameters
    - (Why not turn everything on?)
  - Set values for variables
- The multiple models needed to predict individual performance was cumbersome, however individual models has been suggested by ACT-R researchers
- Difficulty of modeling acquired data
  - Yes/No answers to questions presented some difficulties

Conclusions

- ACT-R activation levels can be used to predict likelihood of errors in situation awareness tasks
- Cognitive modeling can be used in a predictive manner instead of using as a "curve-fitting" simulation
- Future research needs to be done on using predictions to generate interface design analysis