Visualization of Tomorrow

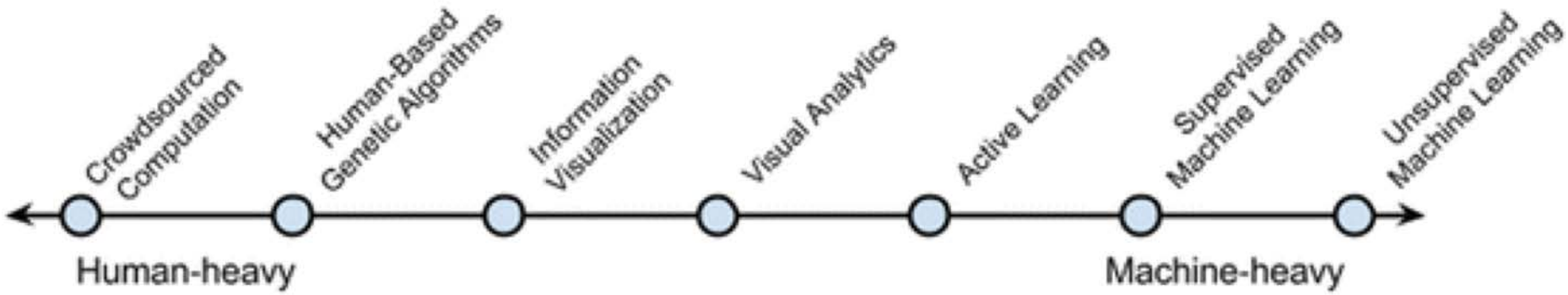
Alvitta Ottley

Washington University in St. Louis

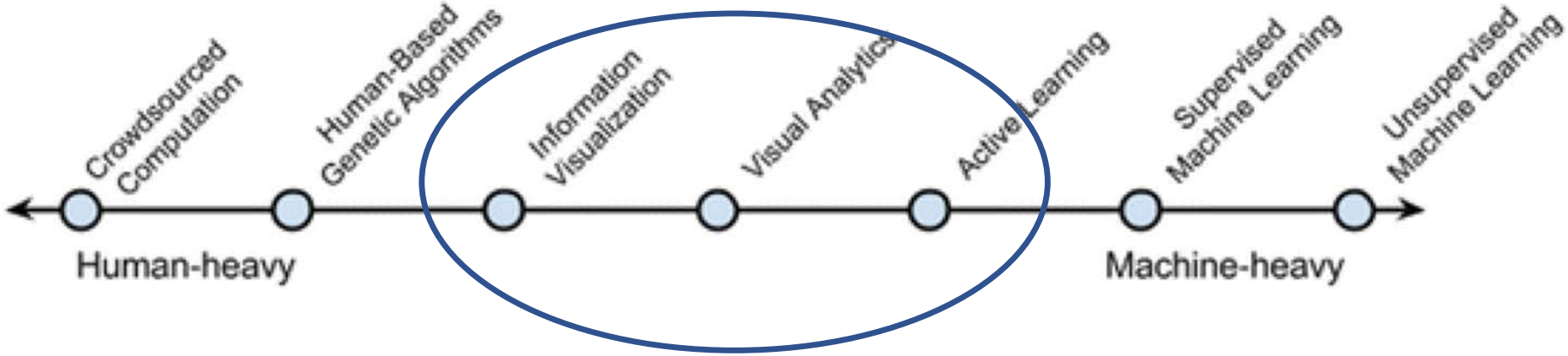
Visual Data Analysis Group

What does visualization mean to you?

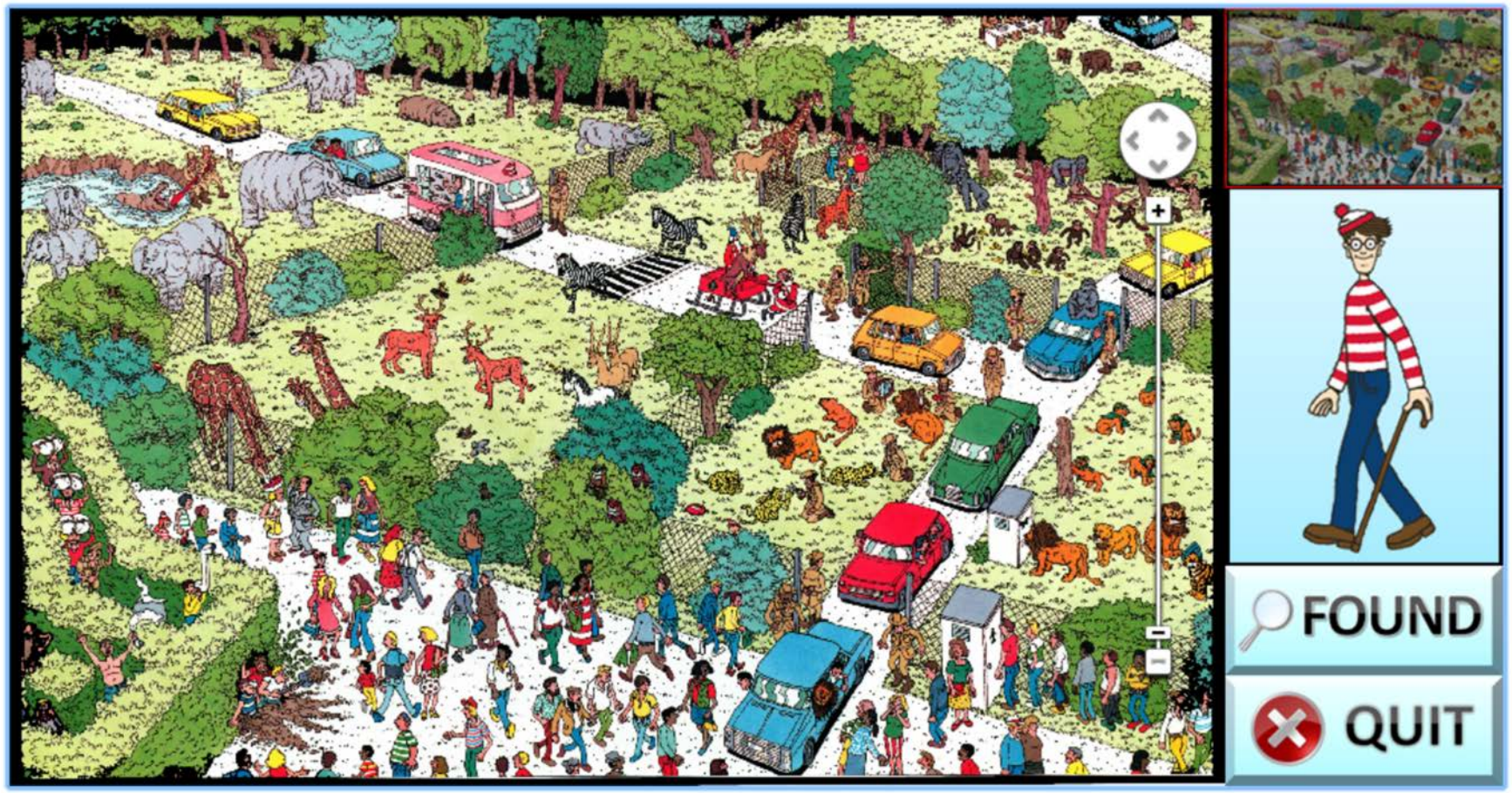
Human-machine collaboration continuum



Human-machine collaboration continuum



https://classes.engineering.wustl.edu/cse557/waldo



What are the possibilities for tomorrow?

The machine must know what you are trying to do

Research questions:

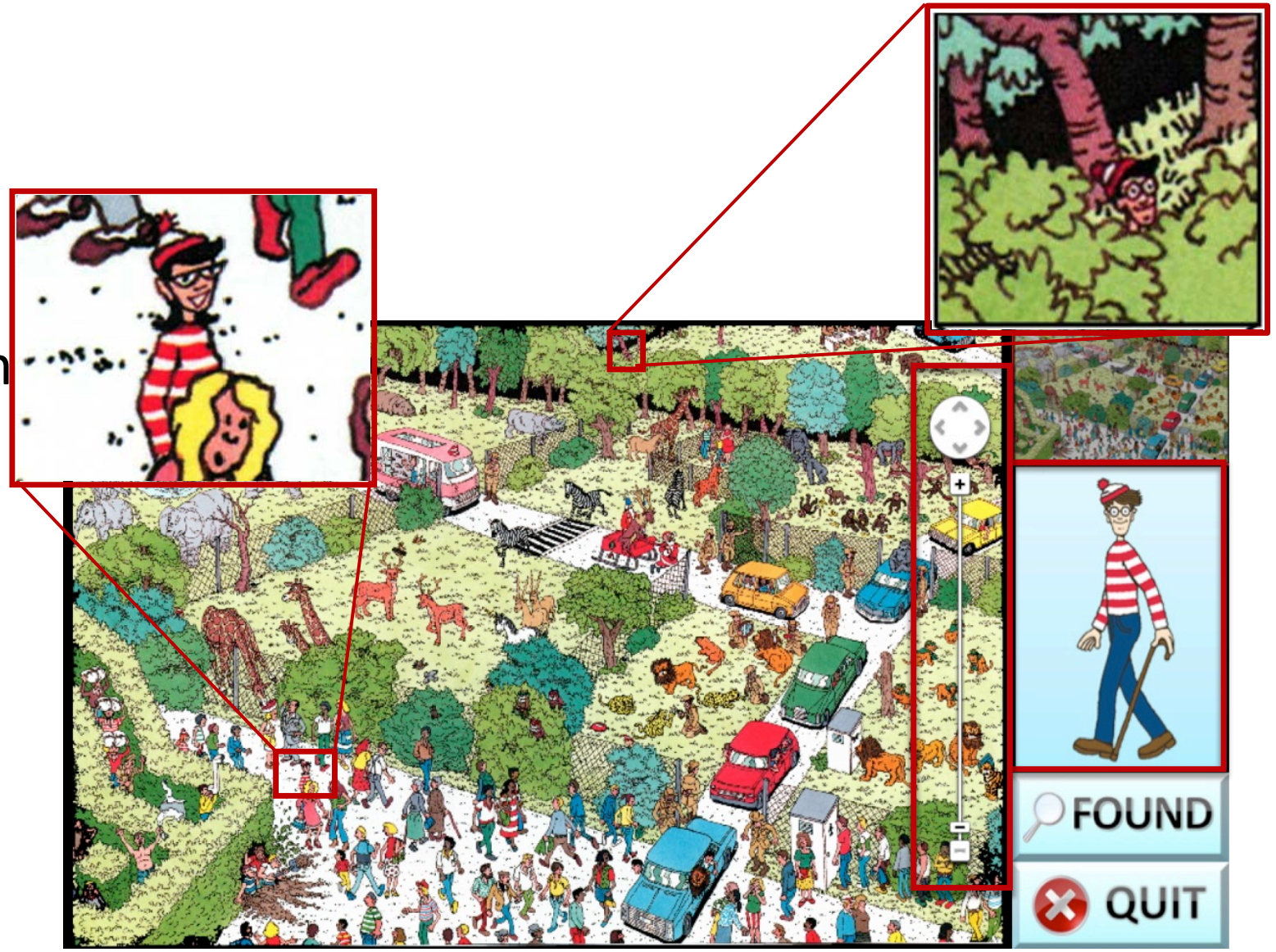
What can we infer through observing interactions?

Finding Waldo

7 control buttons:

Up, Down, Left, Right, Zoom In, Zoom Out, Found

We logged interactions as users search for Waldo



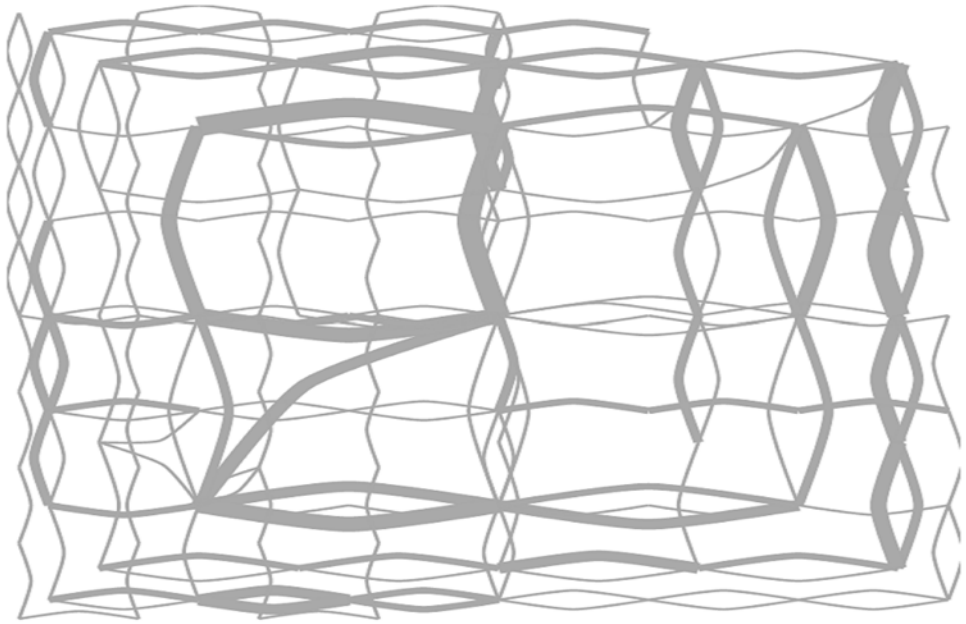
Finding Waldo: Learning about users from their interactions.

E. Brown, **A. Ottley**, J. Zhao, Q. Lin, R. Souvenir, A. Endert, and R. Chang, VAST 2015.

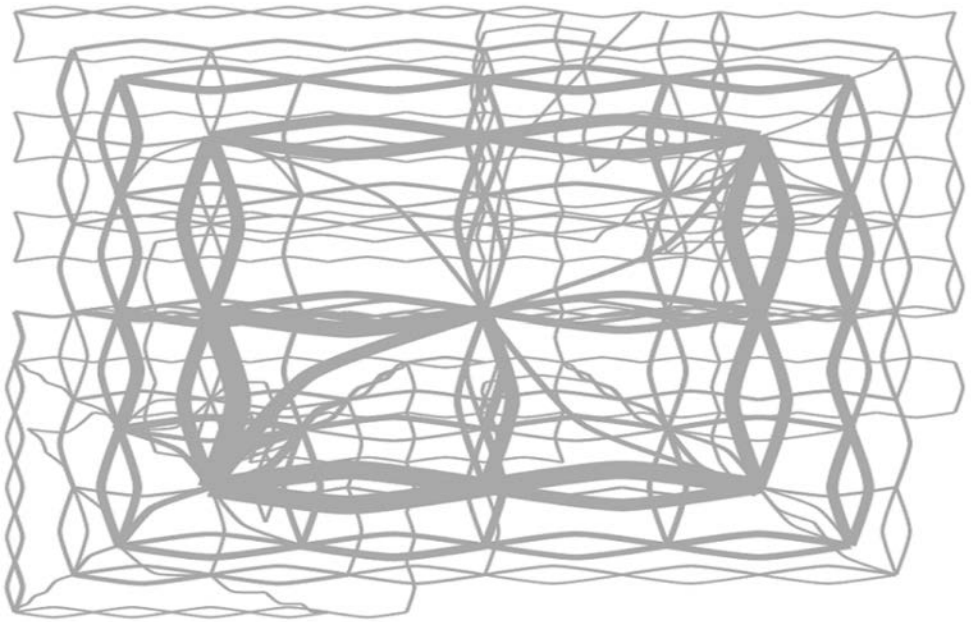
Experiment Design

- 90 online participants
- Mean completion time was 469.5 seconds (stdev = 351.9)
- We captured:
  - Mouse clicks and mouse moves
  - Collected personality data

Visualizing Interactions



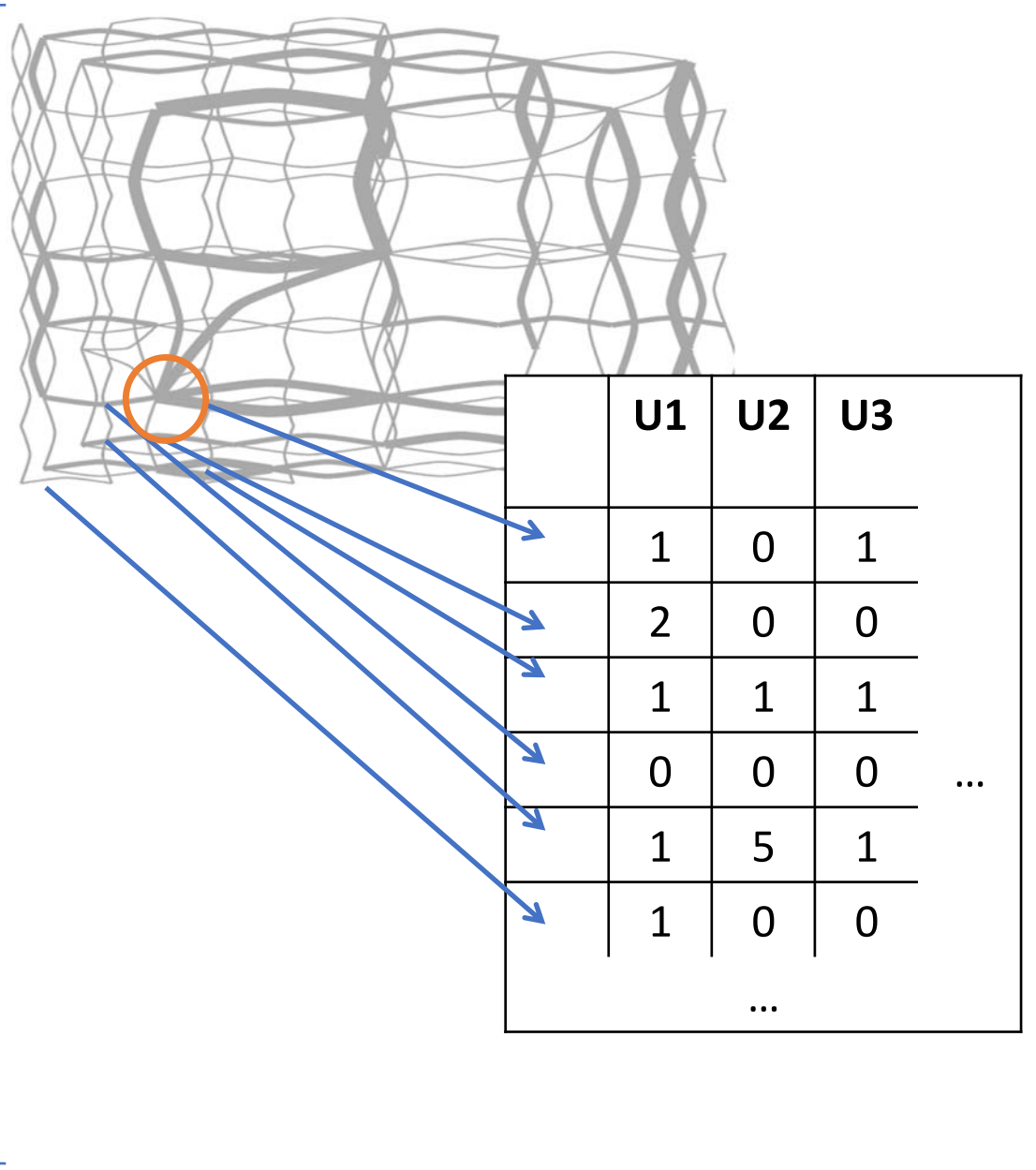
Quick completion time



Slow completion time

F,1,0,0,852,401,0,380826
C,1,0,0,234,494,0,382768
C,1,0,0,409,286,0,384276
C,1,0,0,88,120,0,400513
C,1,0,0,199,486,0,411208
C,1,0,0,595,183,0,444478
F,1,0,0,860,406,0,468280
C,1,0,0,652,288,0,469906
C,1,0,0,500,323,0,473823
F,1,0,0,876,398,0,497403
C,1,0,0,98,237,0,503475

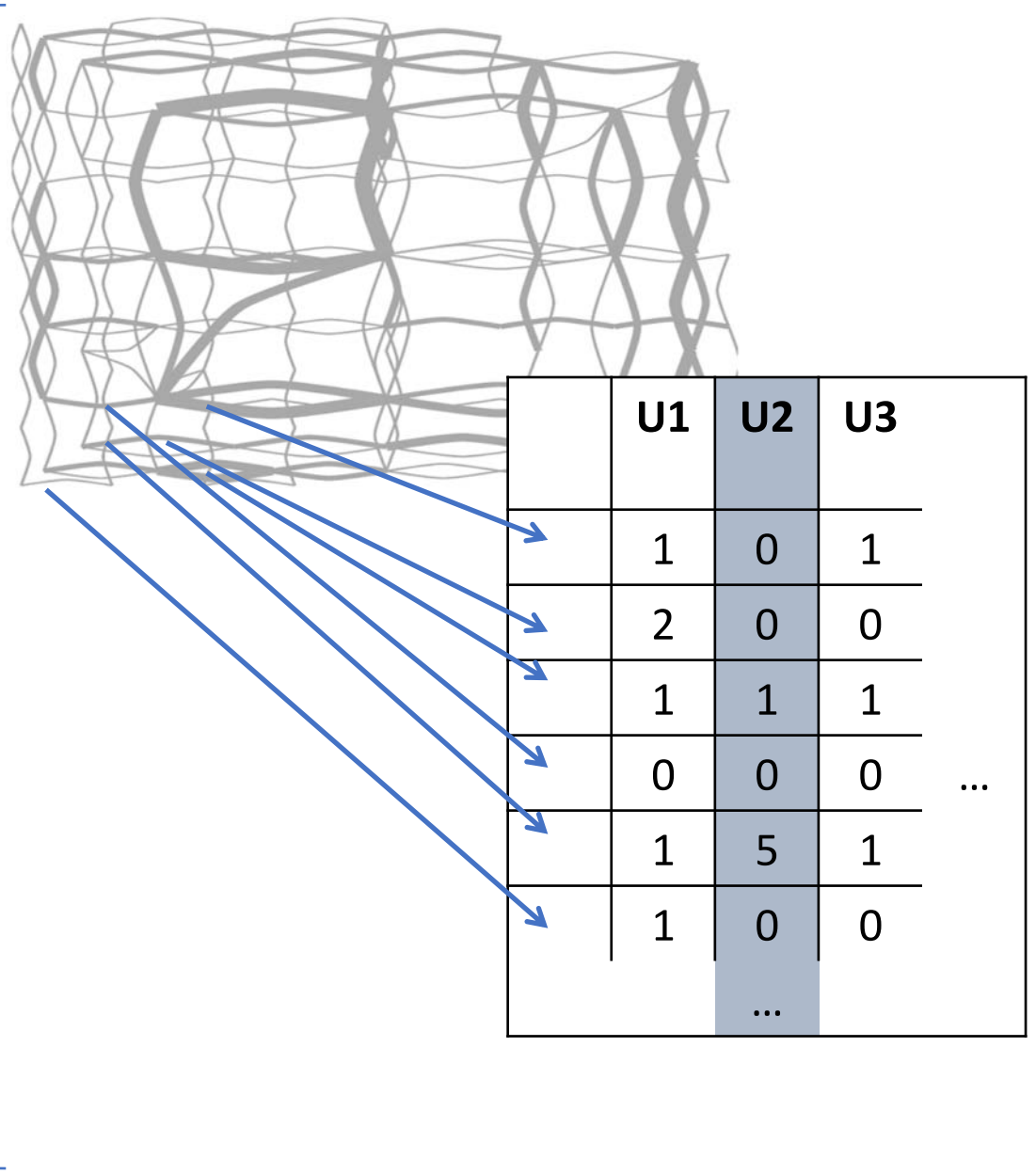
State space



	U1	U2	U3	
	1	0	1	
	2	0	0	
	1	1	1	
	0	0	0	…
	1	5	1	
	1	0	0	
…				

F,1,0,0,852,401,0,380826
C,1,0,0,234,494,0,382768
C,1,0,0,409,286,0,384276
C,1,0,0,88,120,0,400513
C,1,0,0,199,486,0,411208
C,1,0,0,595,183,0,444478
F,1,0,0,860,406,0,468280
C,1,0,0,652,288,0,469906
C,1,0,0,500,323,0,473823
F,1,0,0,876,398,0,497403
C,1,0,0,98,237,0,503475

State space

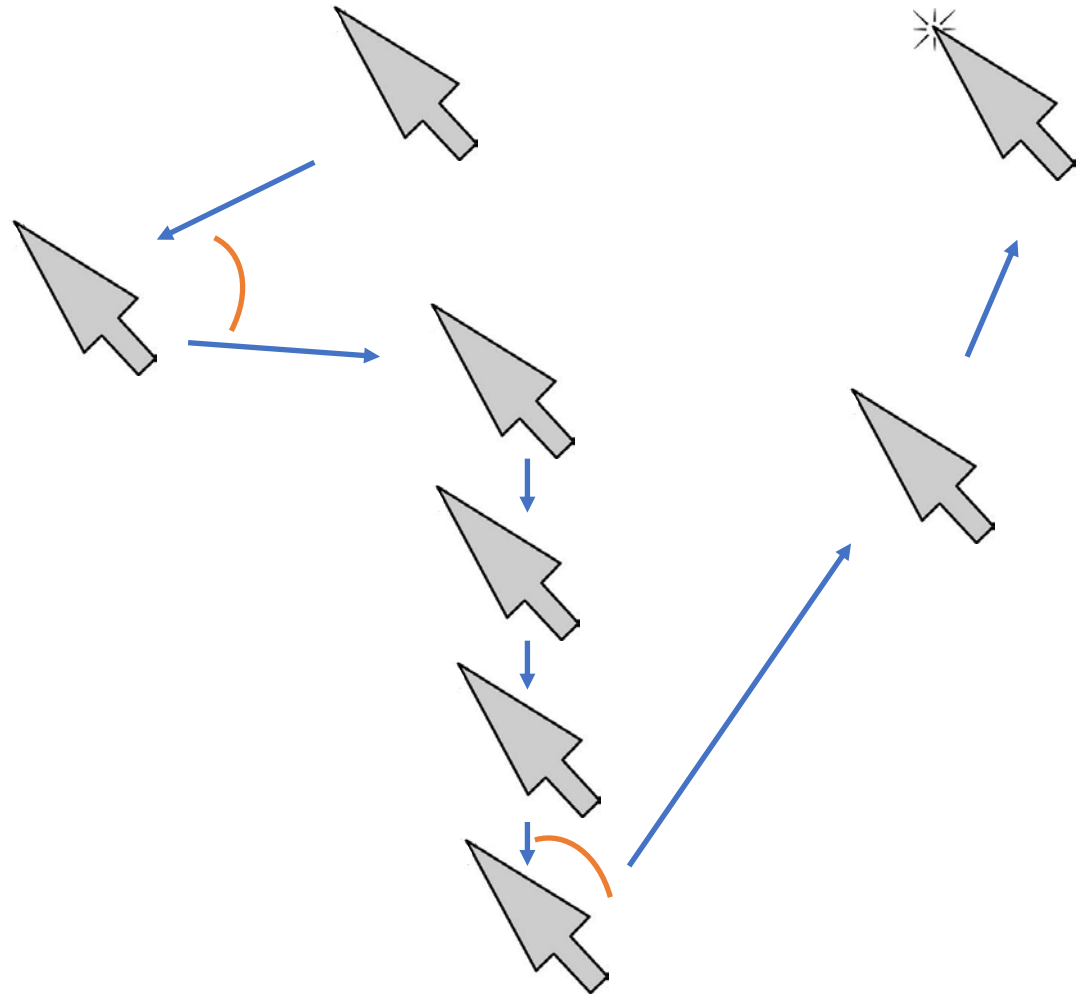


	U1	U2	U3	
	1	0	1	
	2	0	0	
	1	1	1	
	0	0	0	…
	1	5	1	
	1	0	0	
		…		

F,1,0,0,852,401,0,380826
C,1,0,0,234,494,0,382768
C,1,0,0,409,286,0,384276
C,1,0,0,88,120,0,400513
C,1,0,0,199,486,0,411208
C,1,0,0,595,183,0,444478
F,1,0,0,860,406,0,468280
C,1,0,0,652,288,0,469906
C,1,0,0,500,323,0,473823
F,1,0,0,876,398,0,497403
C,1,0,0,98,237,0,503475

State space

Mouse event



F,1,0,0,852,401,0,380826
C,1,0,0,234,494,0,382768
C,1,0,0,409,286,0,384276
C,1,0,0,88,120,0,400513
C,1,0,0,199,486,0,411208
C,1,0,0,595,183,0,444478
F,1,0,0,860,406,0,468280
C,1,0,0,652,288,0,469906
C,1,0,0,500,323,0,473823
F,1,0,0,876,398,0,497403
C,1,0,0,98,237,0,503475

State space

Mouse event

Sequences

U1: LIIDLURRUUIRRDD…
U2: URDIIDDLLOLO…
U3: LLIIDDORRRURIII…

…

	U1	U2	U3	
LI	1	0	2	
LII	1	0	0	
IID	3	1	1	
IDI	0	0	0	…
DD	1	2	1	
IRR	1	0	0	
		…		

L = Left, R = Right, U = Up, D = Down,
I = Zoom In, O = Zoom Out

F,1,0,0,852,401,0,380826
C,1,0,0,234,494,0,382768
C,1,0,0,409,286,0,384276
C,1,0,0,88,120,0,400513
C,1,0,0,199,486,0,411208
C,1,0,0,595,183,0,444478
F,1,0,0,860,406,0,468280
C,1,0,0,652,288,0,469906
C,1,0,0,500,323,0,473823
F,1,0,0,876,398,0,497403
C,1,0,0,98,237,0,503475

State space

Mouse event

Sequences

U1: LIIDLURRUUIRRDD…
U2: URDIIDDLLOLO…
U3: LLIIDDORRRURIII…

…

	U1	U2	U3	
LI	1	0	2	
LII	1	0	0	
IID	3	1	1	
IDI	0	0	0	…
DD	1	2	1	
IRR	1	0	0	
		…		

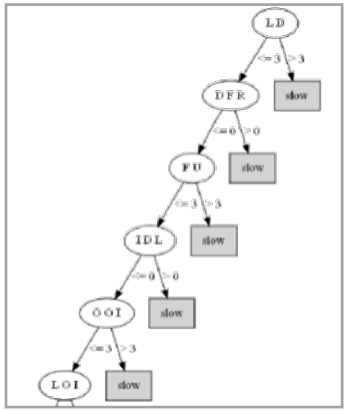
L = Left, R = Right, U = Up, D = Down,
I = Zoom In, O = Zoom Out

F,1,0,0,852,401,0,380826
C,1,0,0,234,494,0,382768
C,1,0,0,409,286,0,384276
C,1,0,0,88,120,0,400513
C,1,0,0,199,486,0,411208
C,1,0,0,595,183,0,444478
F,1,0,0,860,406,0,468280
C,1,0,0,652,288,0,469906
C,1,0,0,500,323,0,473823
F,1,0,0,876,398,0,497403
C,1,0,0,98,237,0,503475

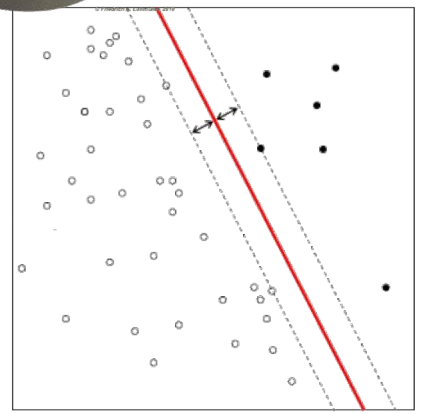
State space

Mouse event

Sequences



decision tree



SVM

Performance Prediction Accuracy

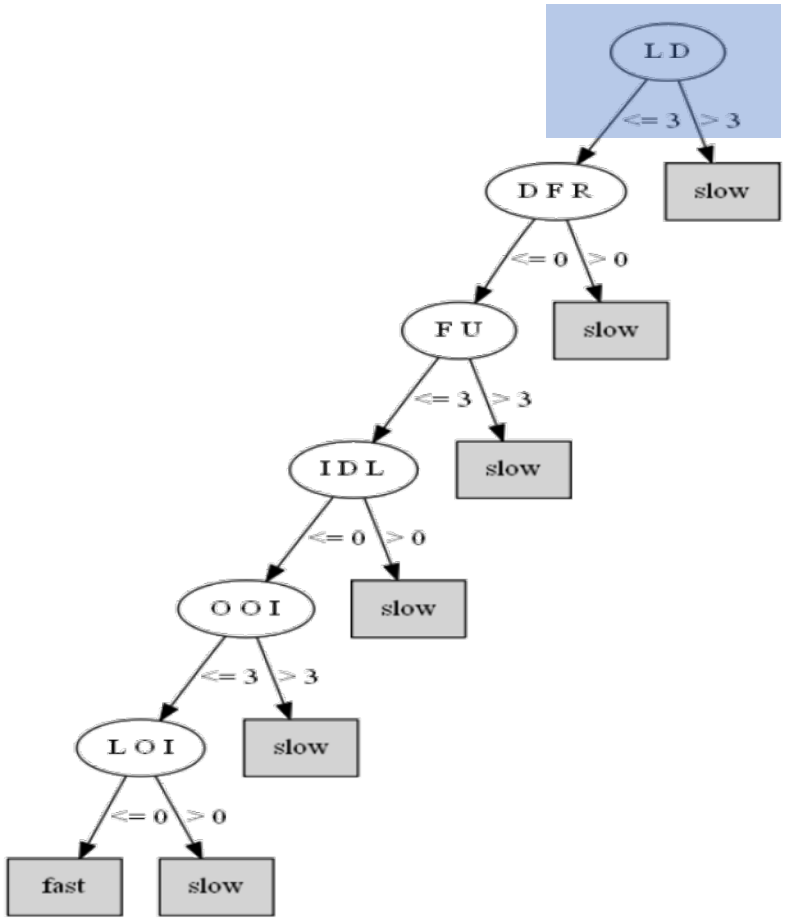
state space (SVM)	83 %
mouse events (SVM)	79 %
sequences (decision tree)	79 %

Interpreting Decision Trees

Decision tree on sequences encoding

Left, Down too much => slow

Waldo is in upper right



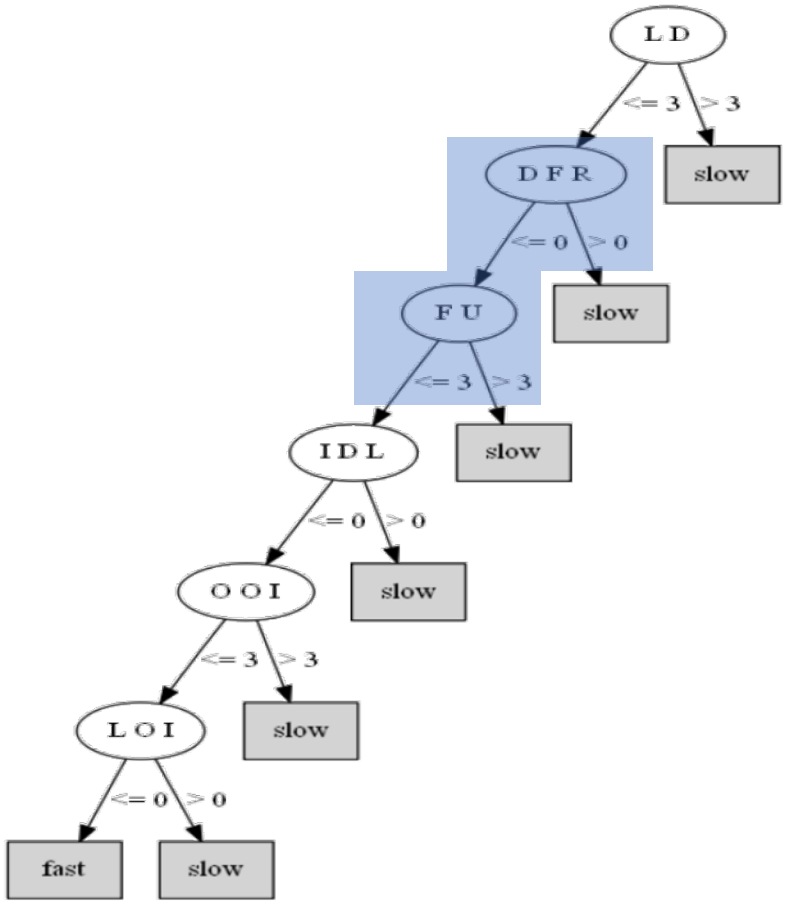
Interpreting Decision Trees

Decision tree on sequences encoding

Found, Up and

Down, *Found*, Right => slow

Clicking Found incorrectly



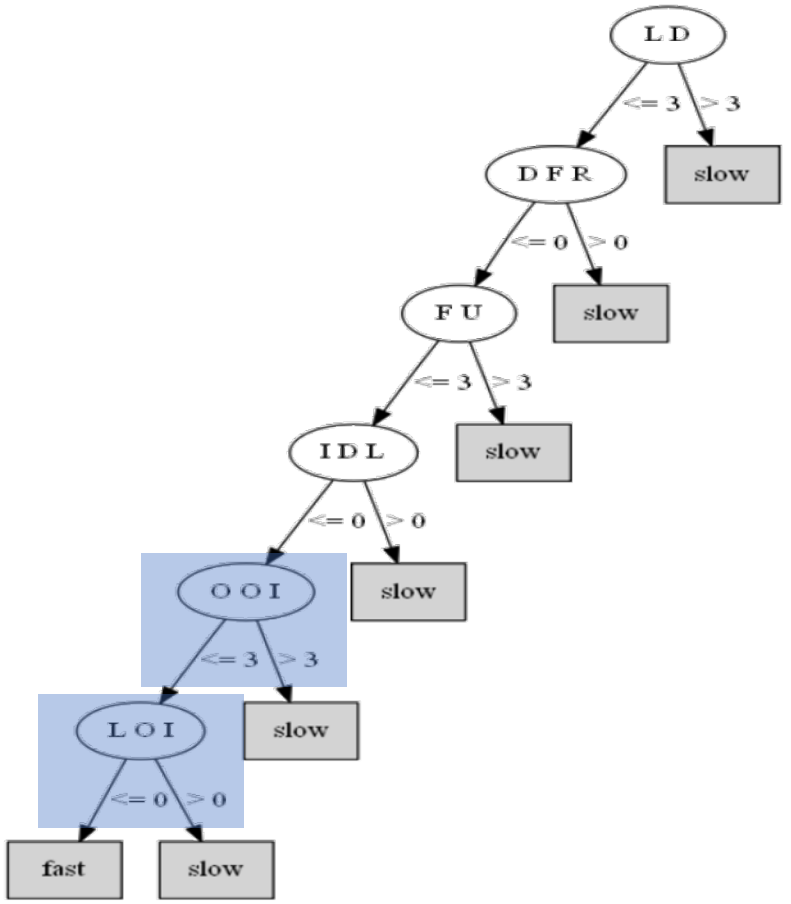
Interpreting Decision Trees

Decision tree on sequences encoding

Out, Out, In and

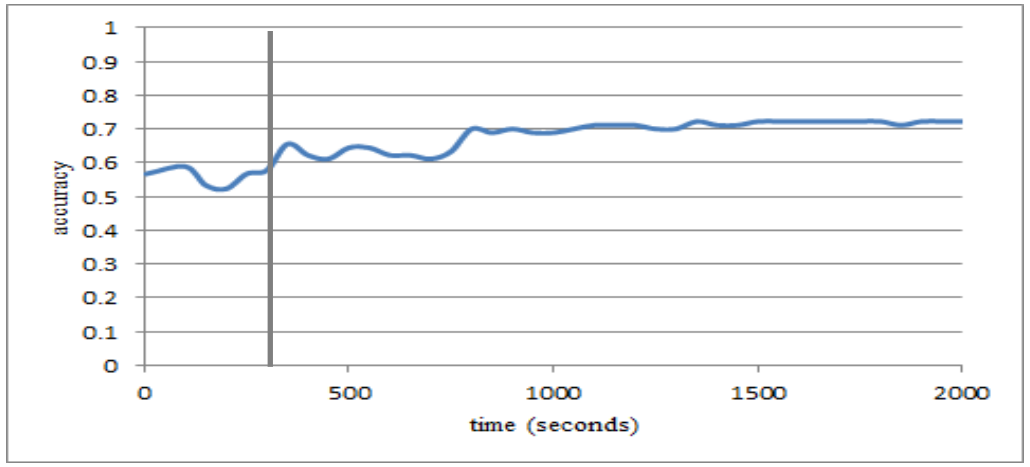
Left, Out, In => slow

Out then in, null move?



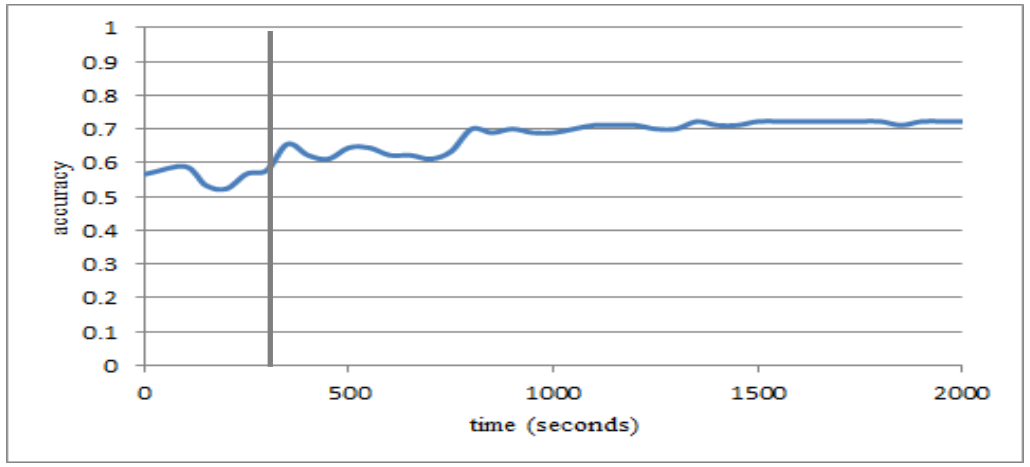
Can we do this real time?

Can we do this real time?

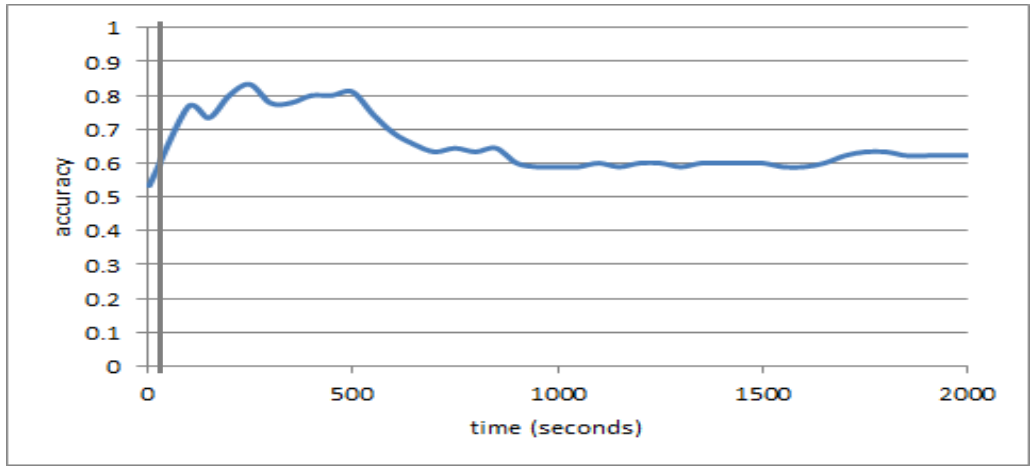


State Space

Can we do this real time?

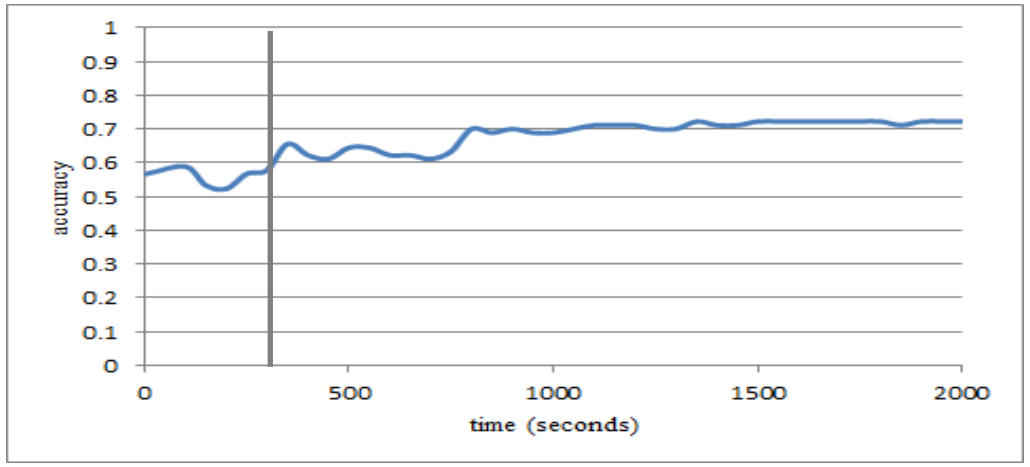


State Space

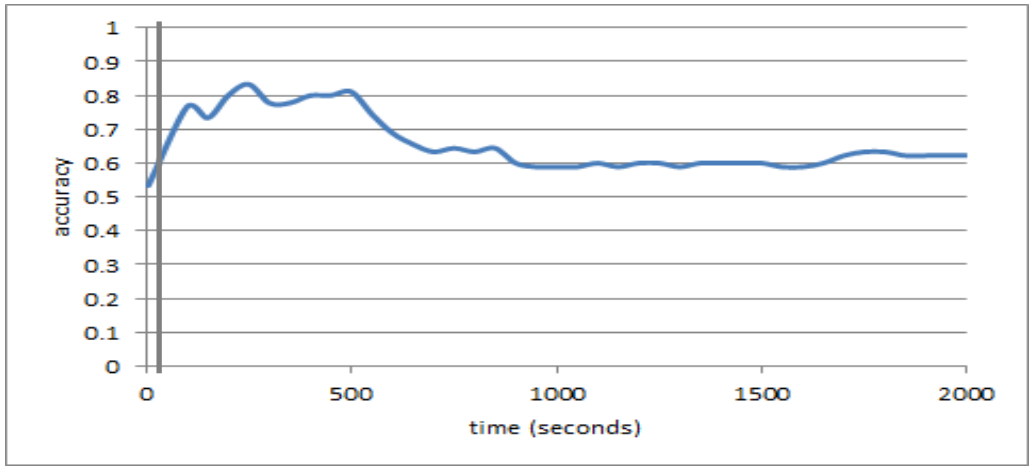


Mouse Events

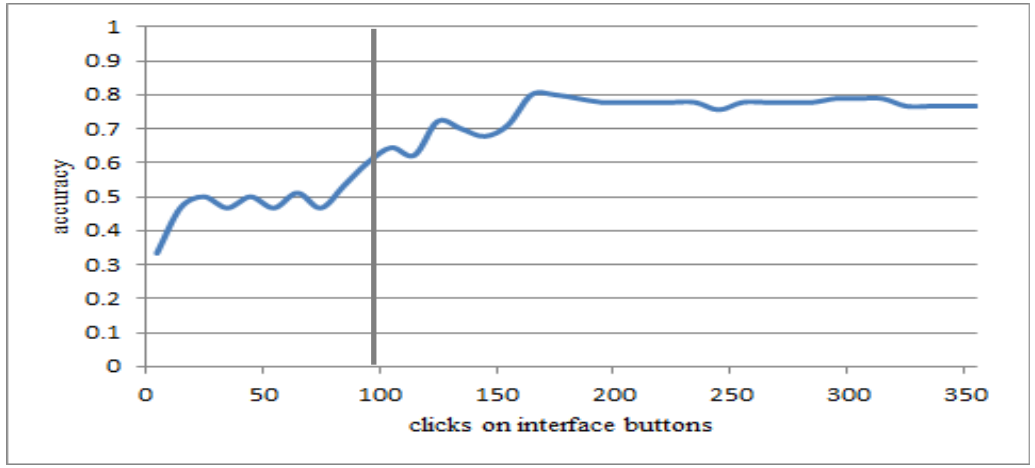
Can we do this real time?



State Space

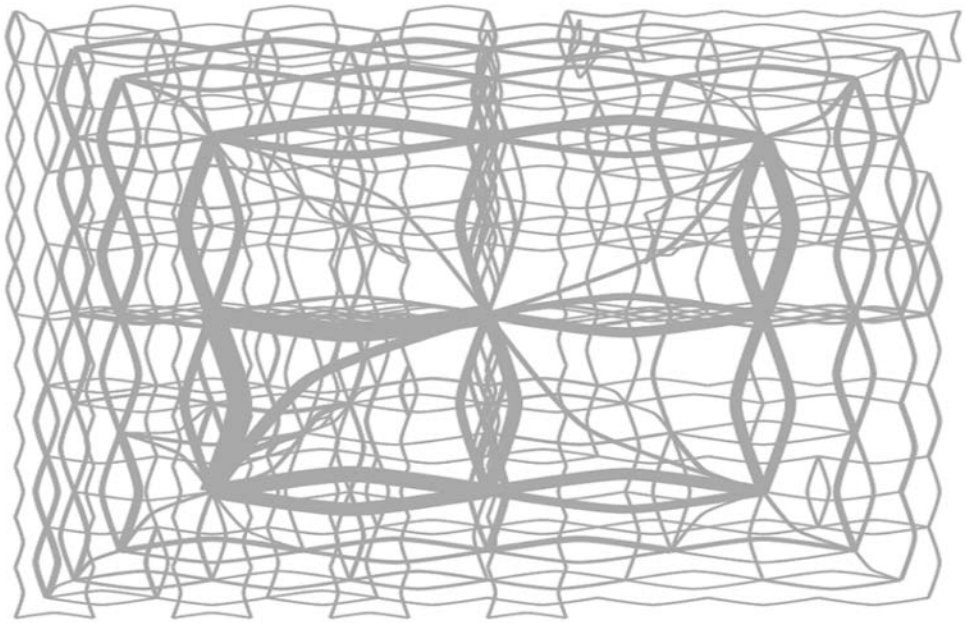


Mouse Events

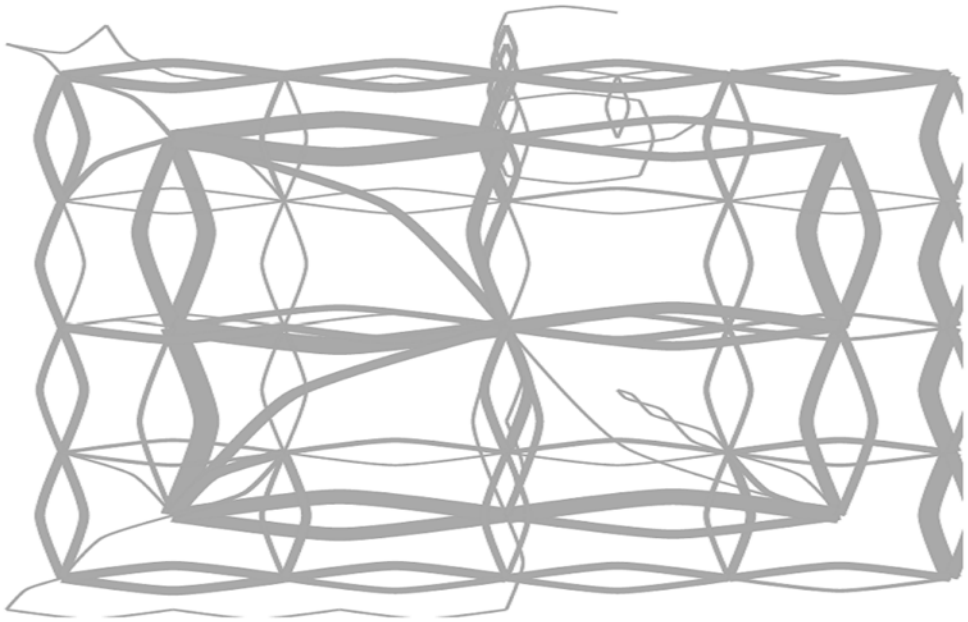


Sequences

What about individual differences?



External Locus of Control

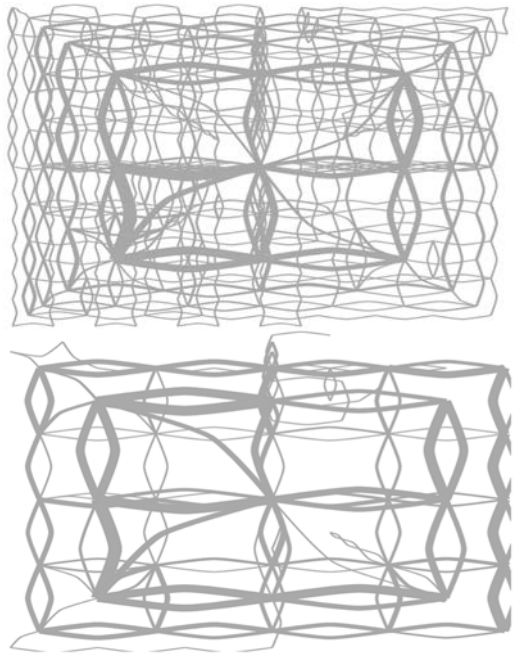


Internal Locus of Control

We were able to detect Extraversion, Neuroticism, and Locus of Control each at ~60% accuracy.

Summary

- It is possible to automatically extract information about users from their interacts
  - Performance
  - Individual traits

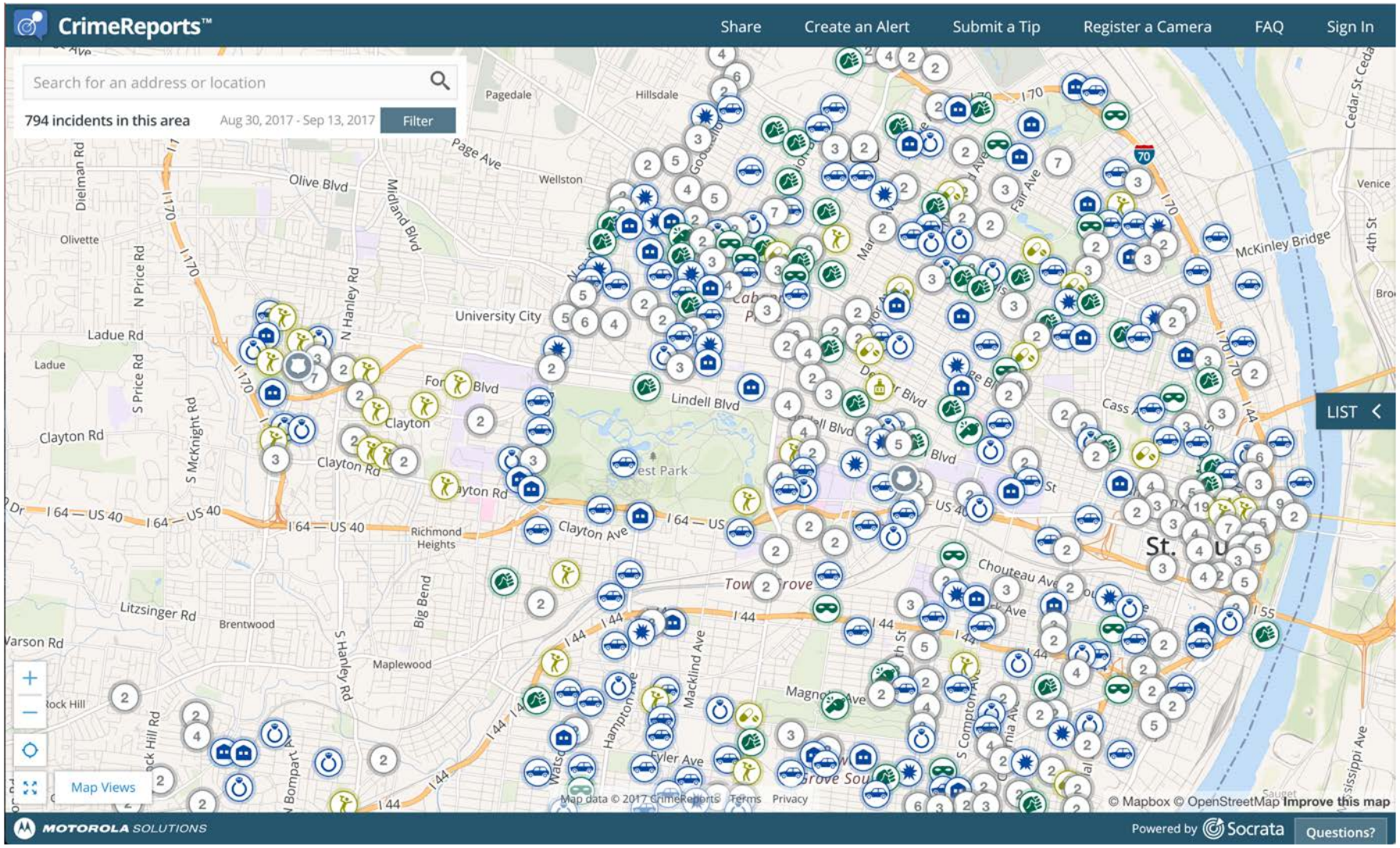


Questions or Comments?

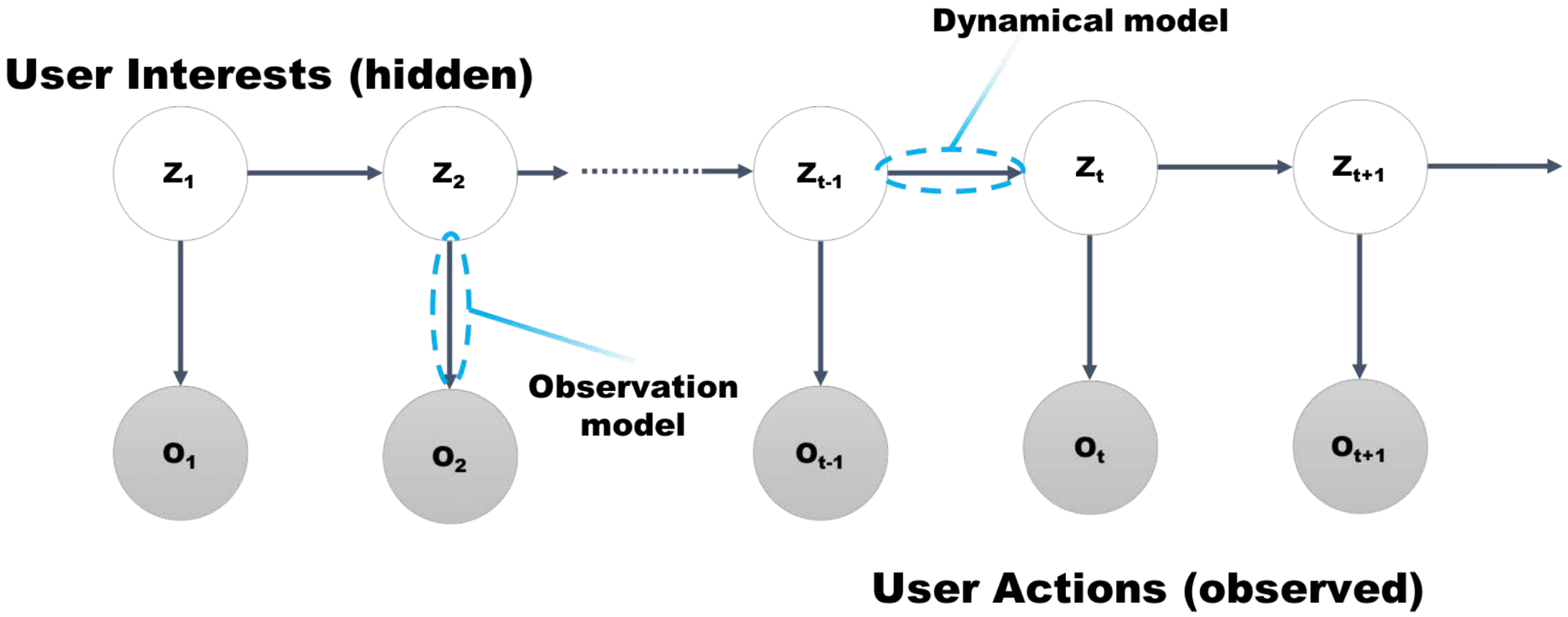
Research questions:

WHAT: Can we learn what the analyst is doing?

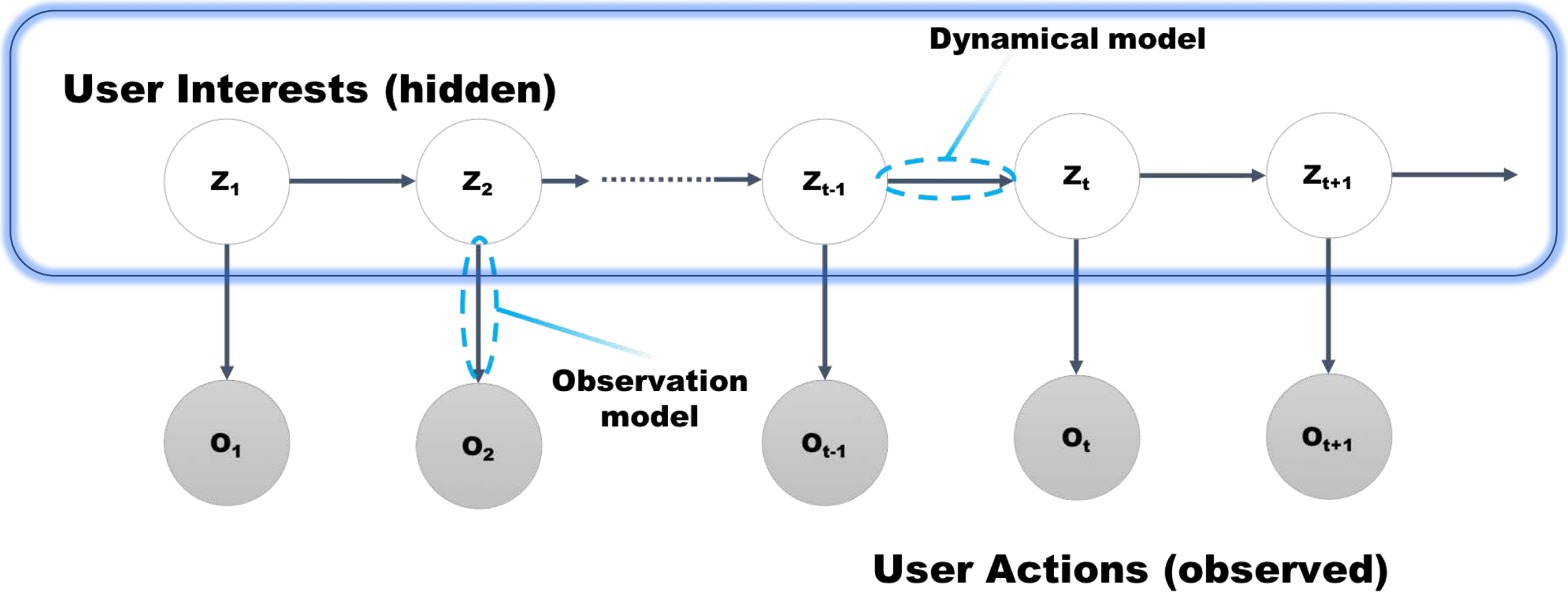
Can we predict moves on a crime map?



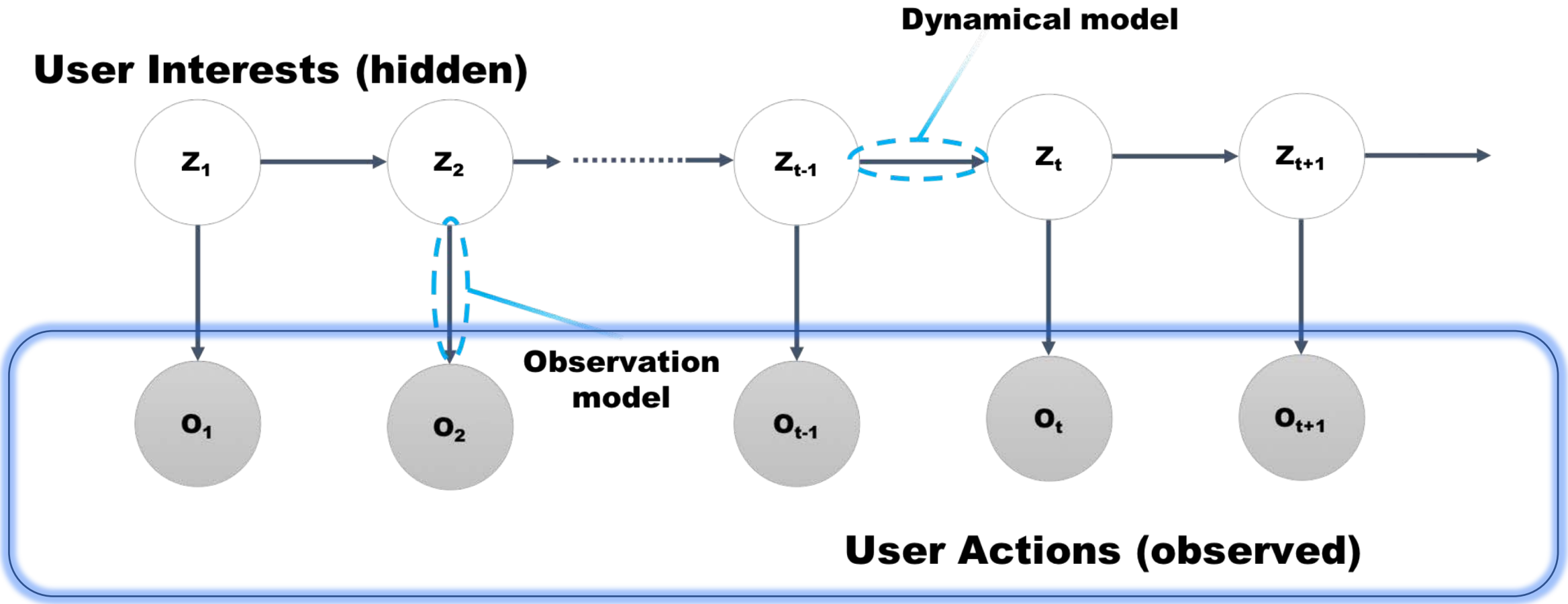
First: Hidden Markov Models



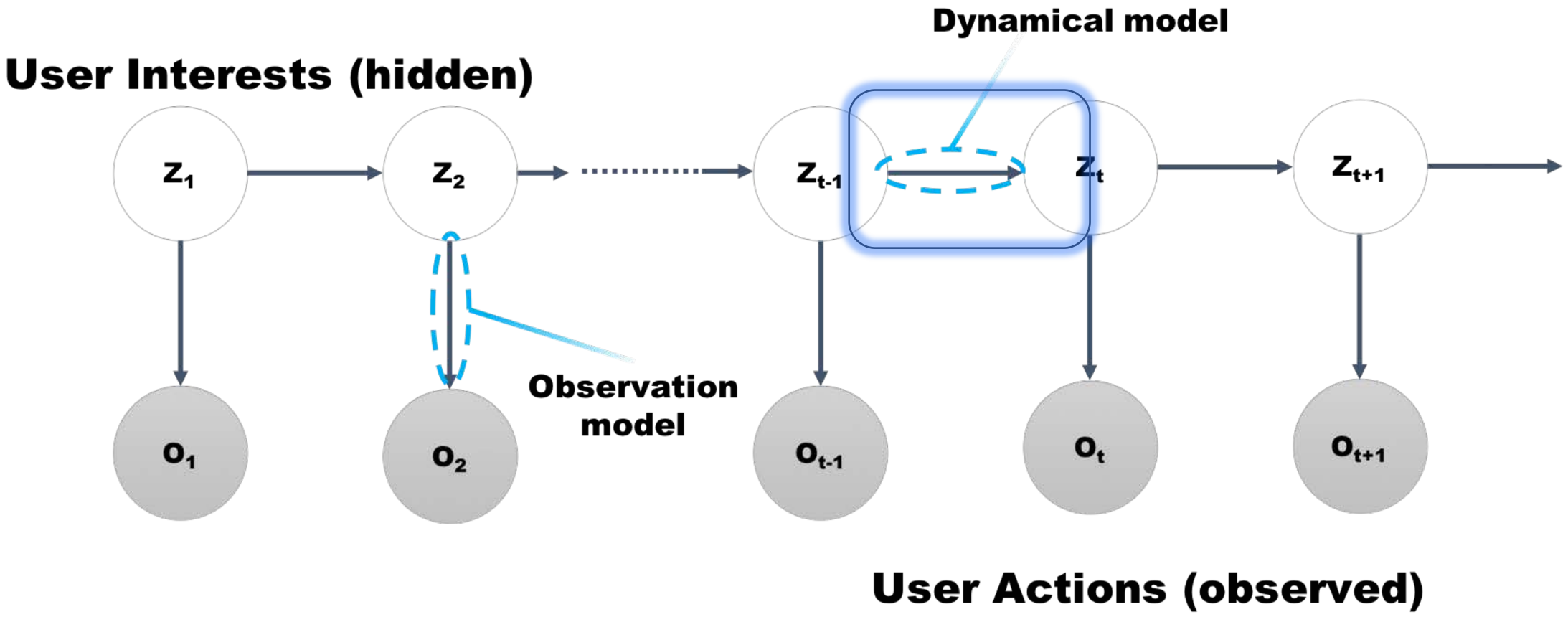
First: Hidden Markov Models



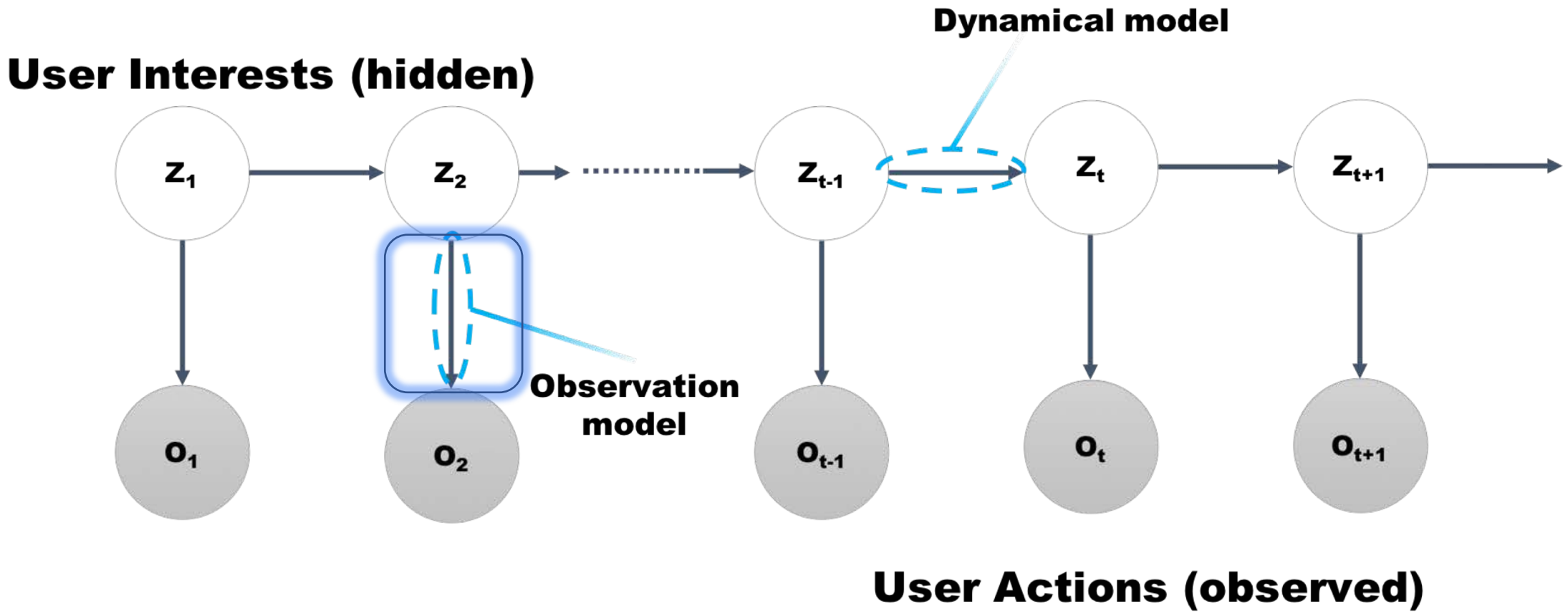
First: Hidden Markov Models



First: Hidden Markov Models



First: Hidden Markov Models



Example

```
states = ('Rainy', 'Sunny')
```

Example

```
states = ('Rainy', 'Sunny')

observations = ('walk', 'shop', 'clean')
```

Example

```
states = ('Rainy', 'Sunny')

observations = ('walk', 'shop', 'clean')

start_probability = {'Rainy': 0.6, 'Sunny': 0.4}
```

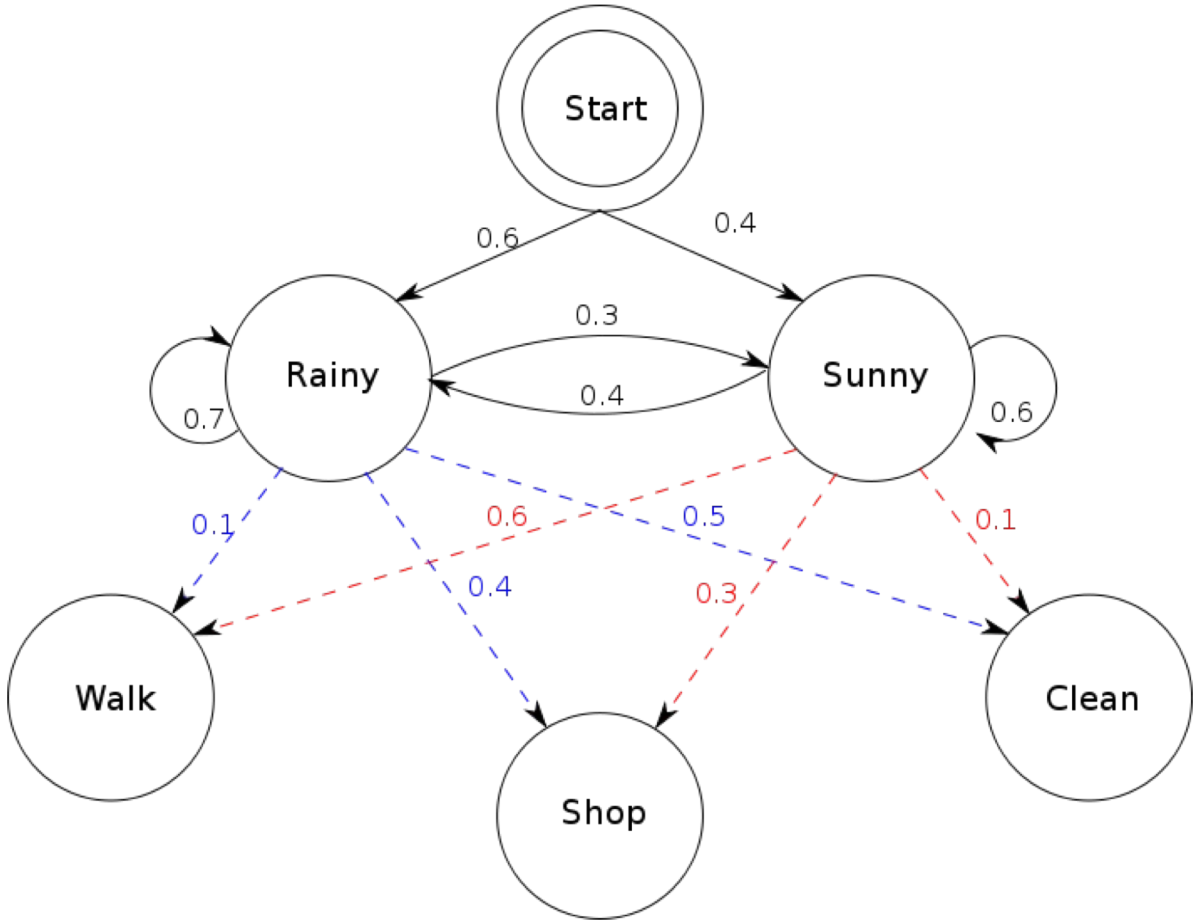
Example

```
states = ('Rainy', 'Sunny')

observations = ('walk', 'shop', 'clean')

start_probability = {'Rainy': 0.6, 'Sunny': 0.4}

transition_probability = {
   'Rainy' : {'Rainy': 0.7, 'Sunny': 0.3},
   'Sunny' : {'Rainy': 0.4, 'Sunny': 0.6},
   }
```

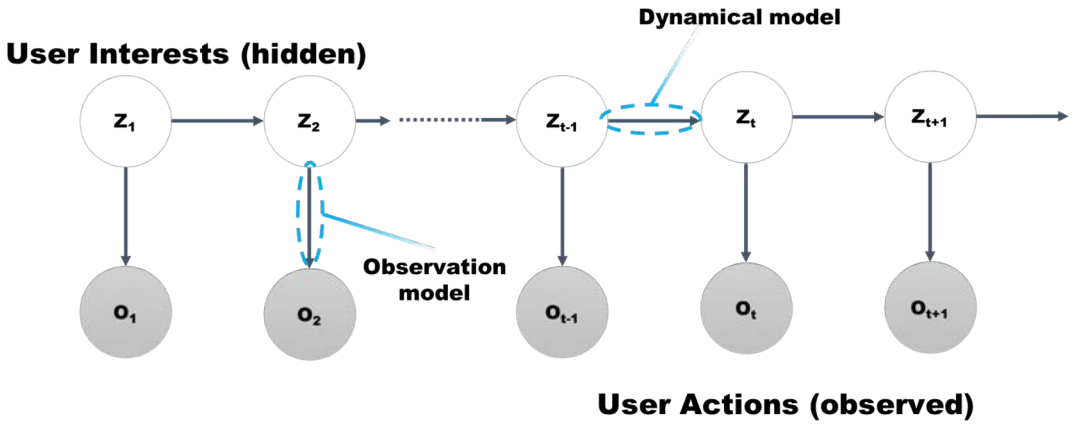
Example

```
states = ('Rainy', 'Sunny')

observations = ('walk', 'shop', 'clean')

start_probability = {'Rainy': 0.6, 'Sunny': 0.4}

transition_probability = {
   'Rainy' : {'Rainy': 0.7, 'Sunny': 0.3},
   'Sunny' : {'Rainy': 0.4, 'Sunny': 0.6},
   }

emission_probability = {
   'Rainy' : {'walk': 0.1, 'shop': 0.4, 'clean': 0.5},
   'Sunny' : {'walk': 0.6, 'shop': 0.3, 'clean': 0.1},
   }
```

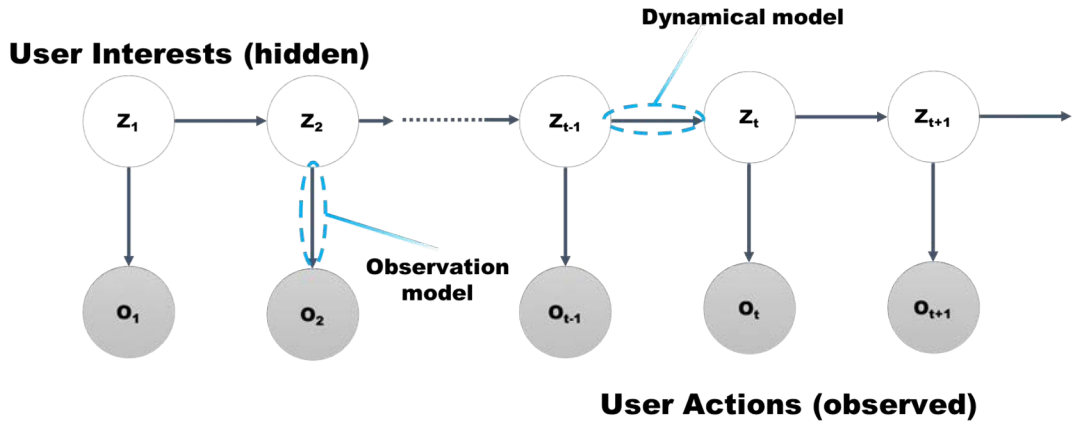
Example


Start
0.6
0.4
Rainy
Sunny
0.3
0.4
0.7
0.6
0.1
0.6
0.5
0.1
0.4
0.3
Walk
Shop
Clean

Apply hidden Markov model to visualizations



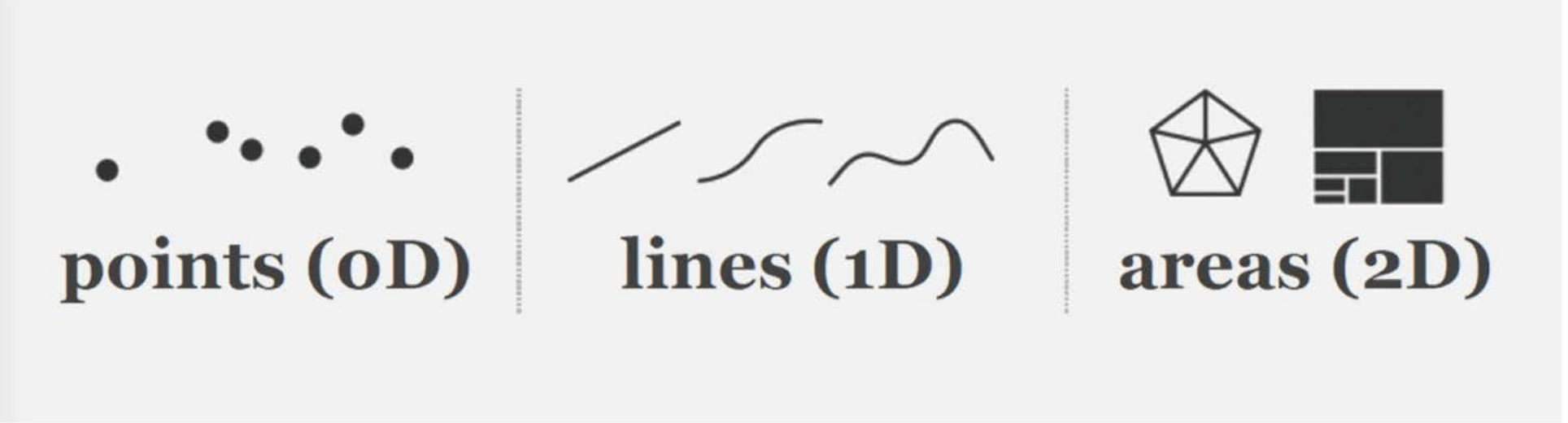
Apply hidden Markov model to visualizations



1. Define hidden states
2. Define observations
3. Define Dynamical model
4. Define Observational model

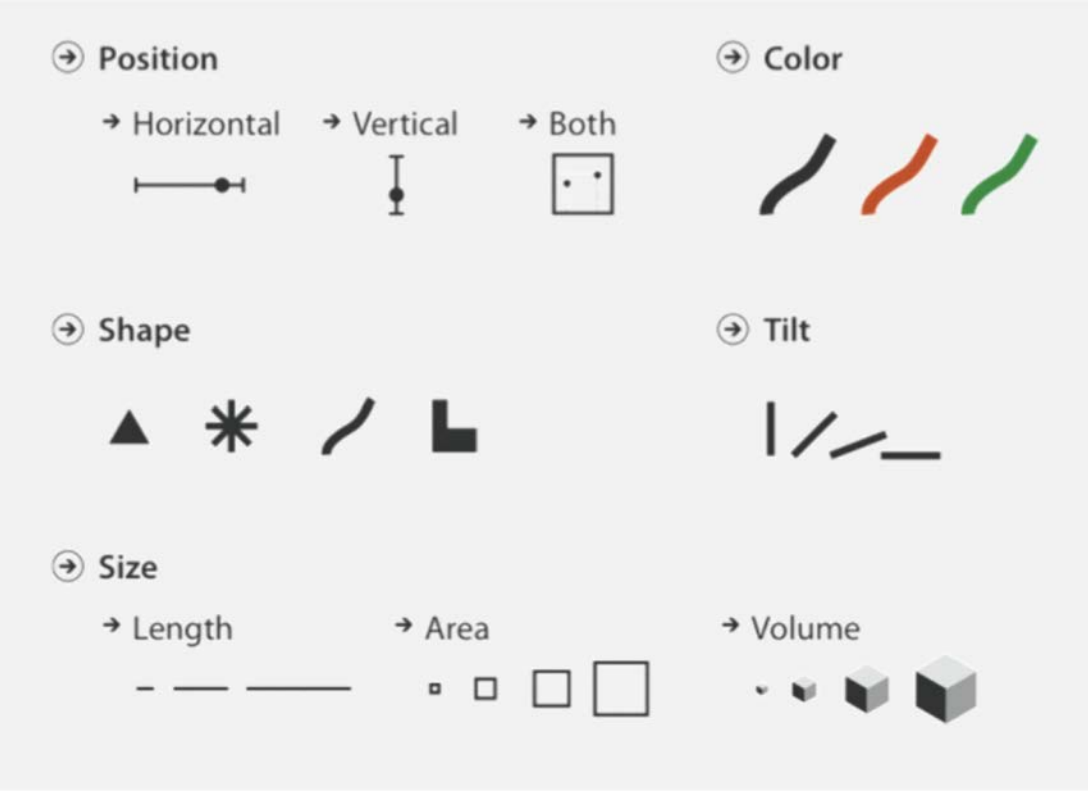
Visual Mark

- Graphical element in an image

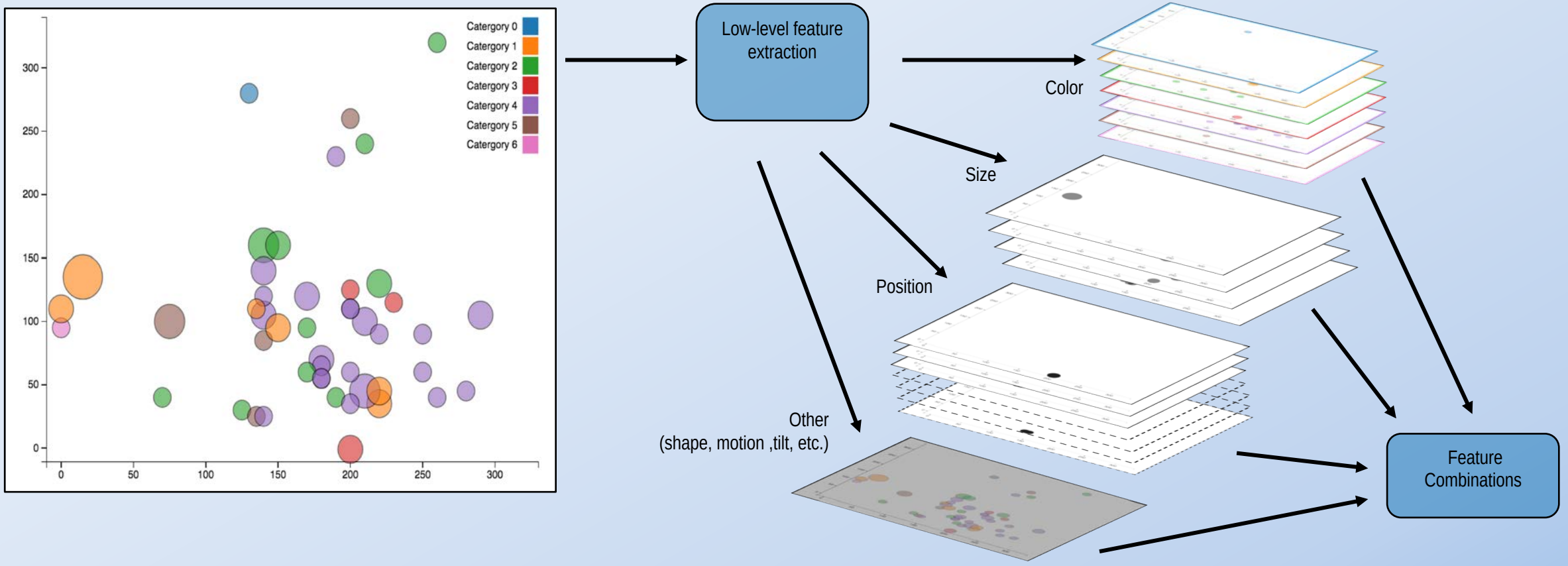


Visual Channel

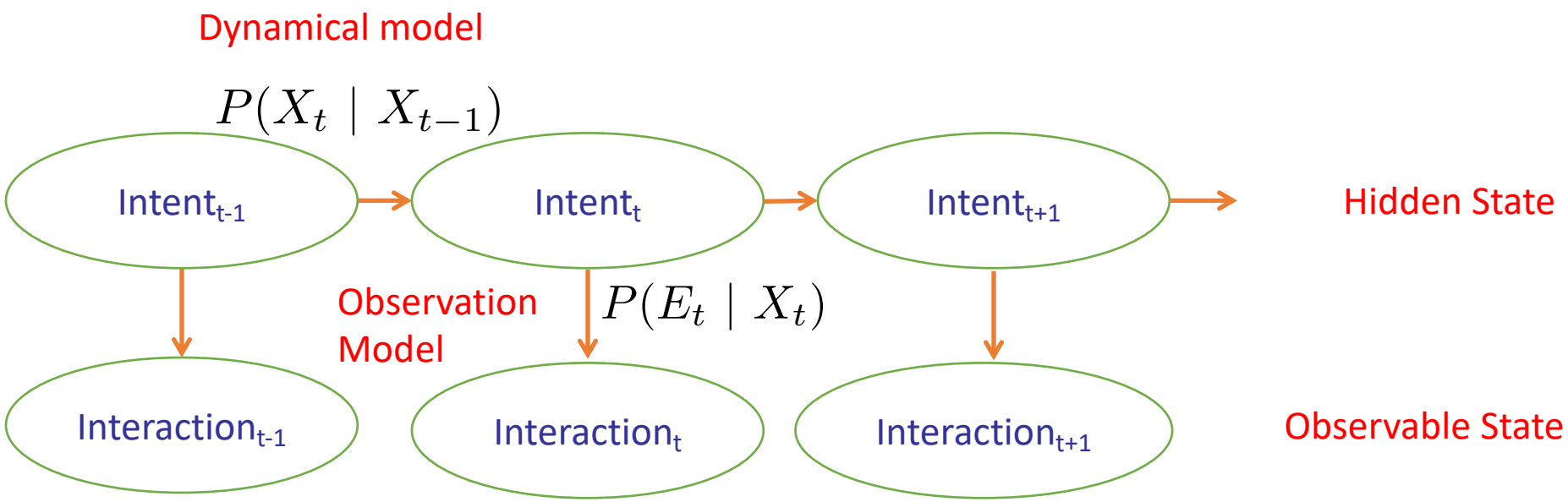
- Parameters that control the appearance of marks



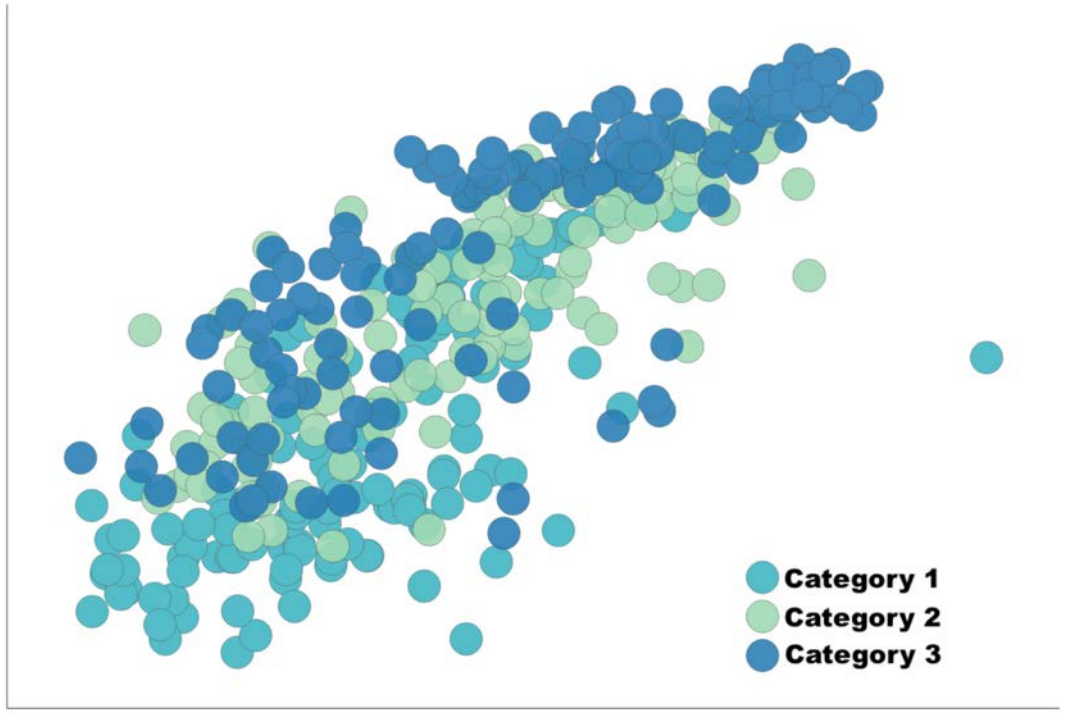
Decomposing the visualization



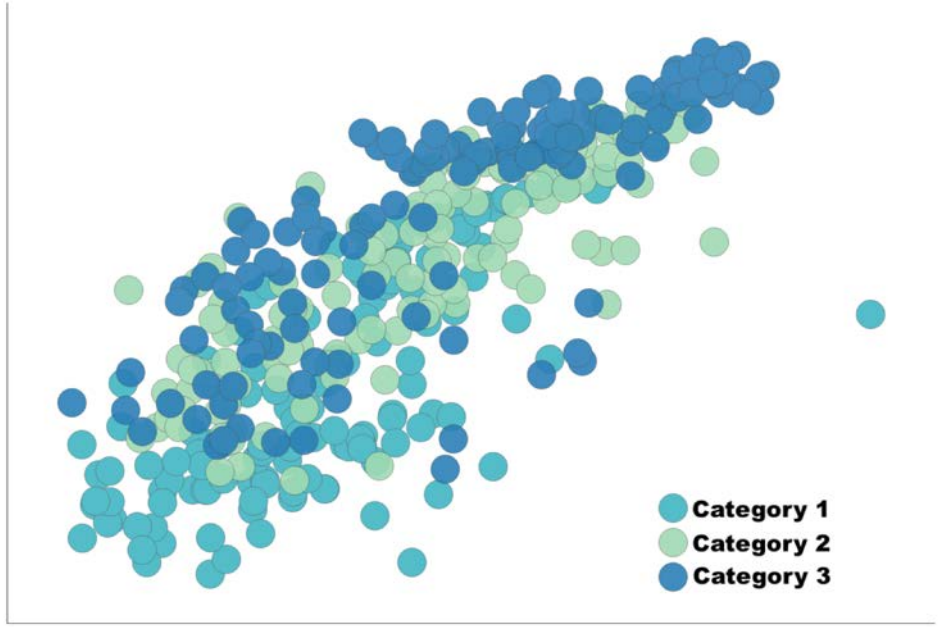
Constructing Probabilistic model



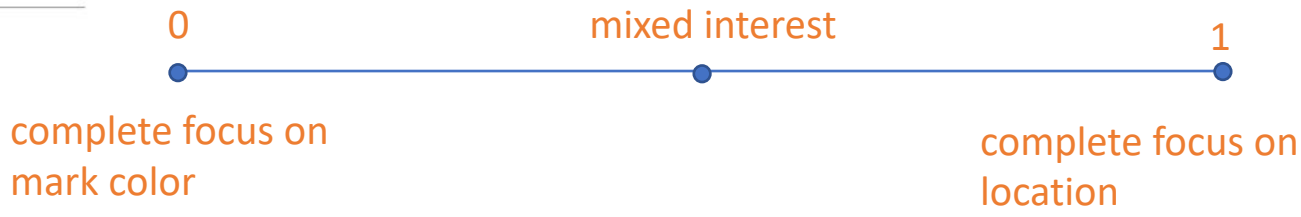
Example



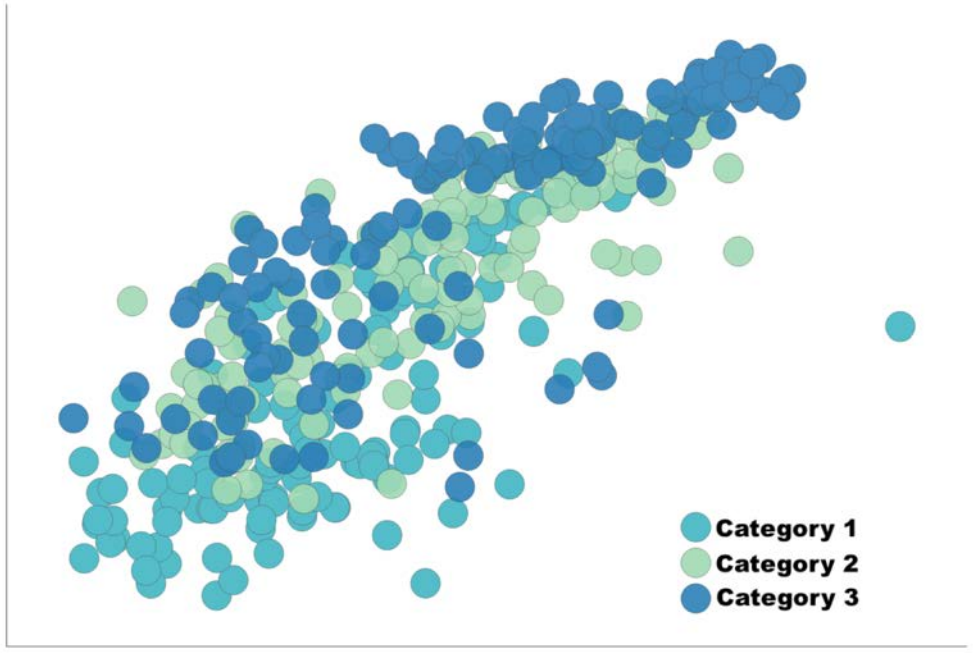
Defining Hidden States



- *M*: the mark space
- *I*: hidden intent space
  - Subspace of *M*
- *In this case, hidden space is a vector*
  - *1) location(x,y)*
  - *2) mark color k*
  - *3) trade of parameter π*
  - *4) z = (x,y,z,π)*



Defining Observable States



- O_t: an observation at time t
  - Mouse click, mouse move, eye gaze
  - represent each observable state as a vector specifying the visual attributes of o_t
- *In this case,*
  - *Mouse click*
  - *A vector of c = (x,y,k)*

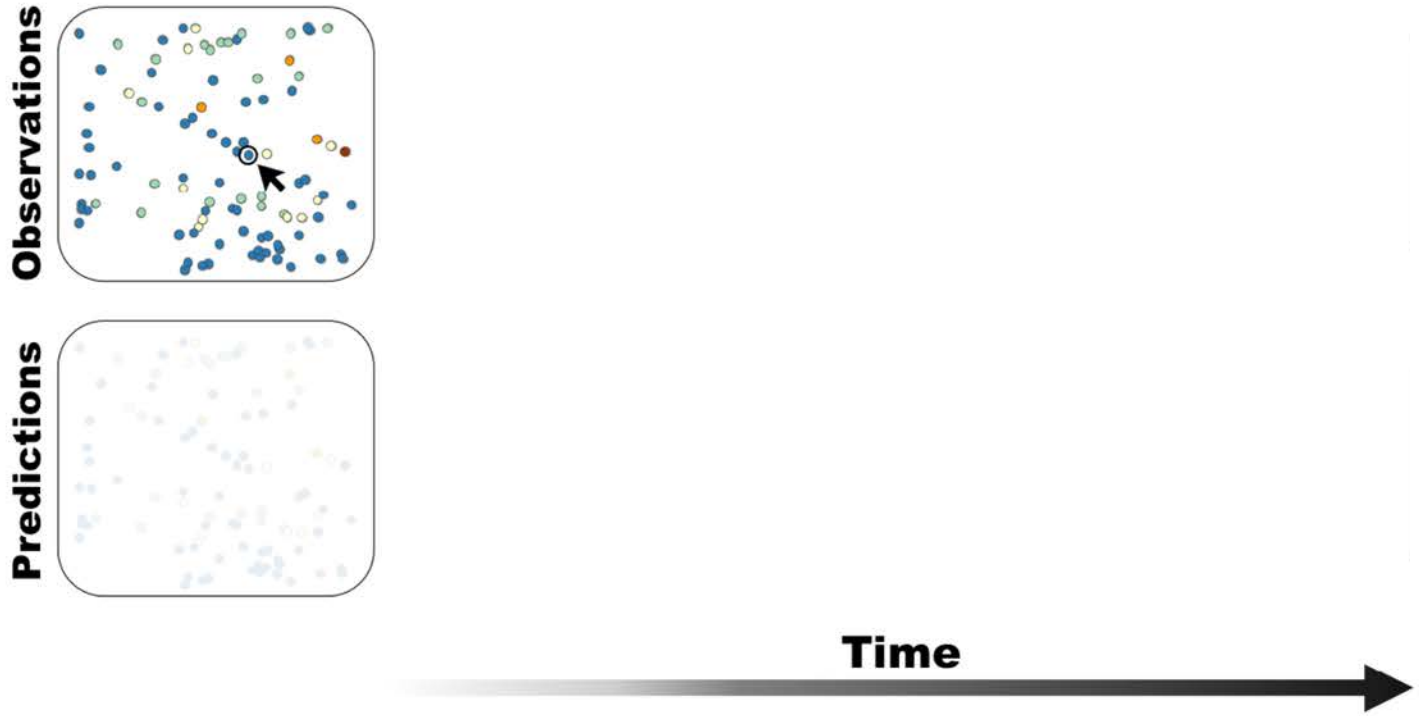
Particle Filtering

- Robot localization example
- https://www.youtube.com/watch?v=5nswUcByXLo

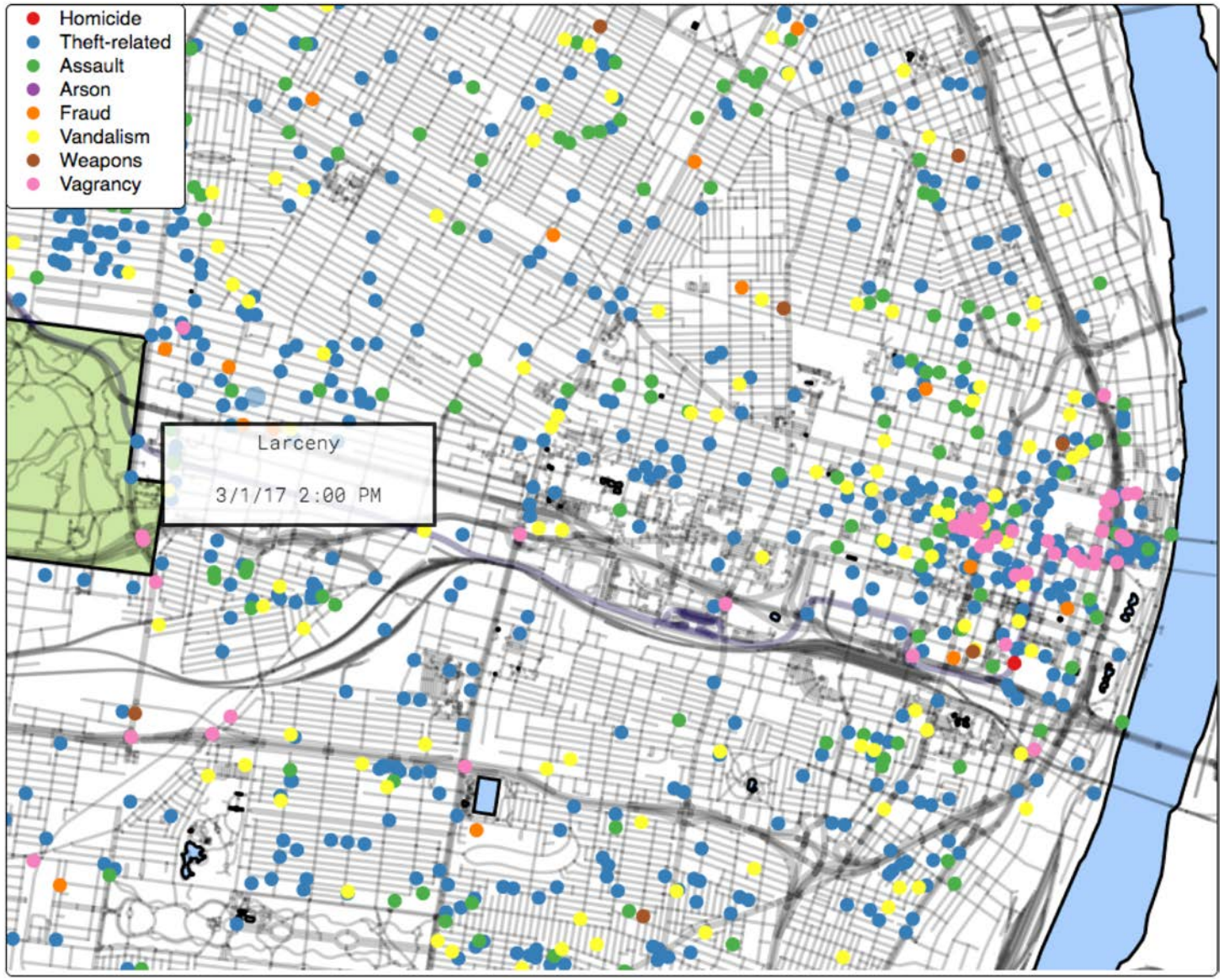
Predicting movement

$$d* = \arg\max_{d \in D} \sum_{p \in P} \mathbb{P}(d|p)$$

Simulation

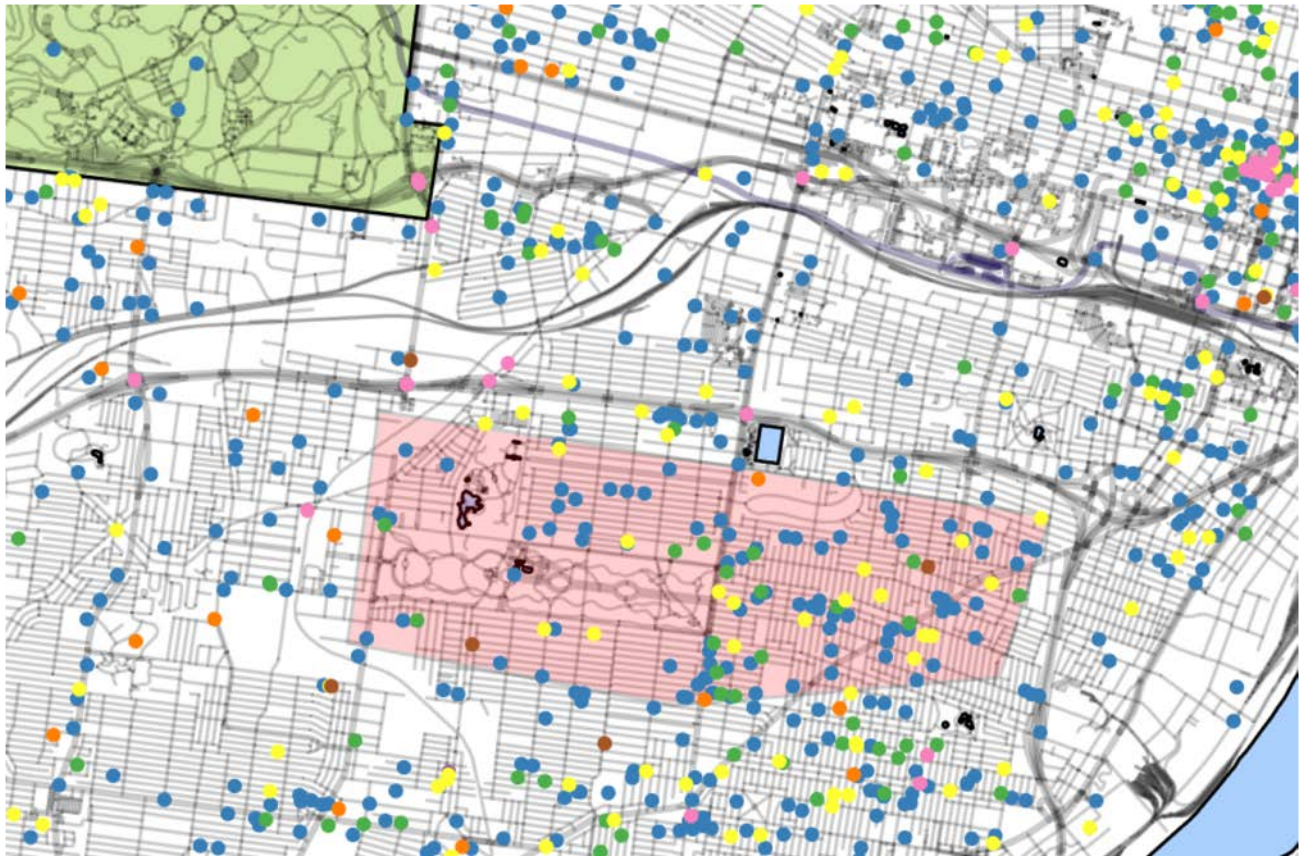


Case Study

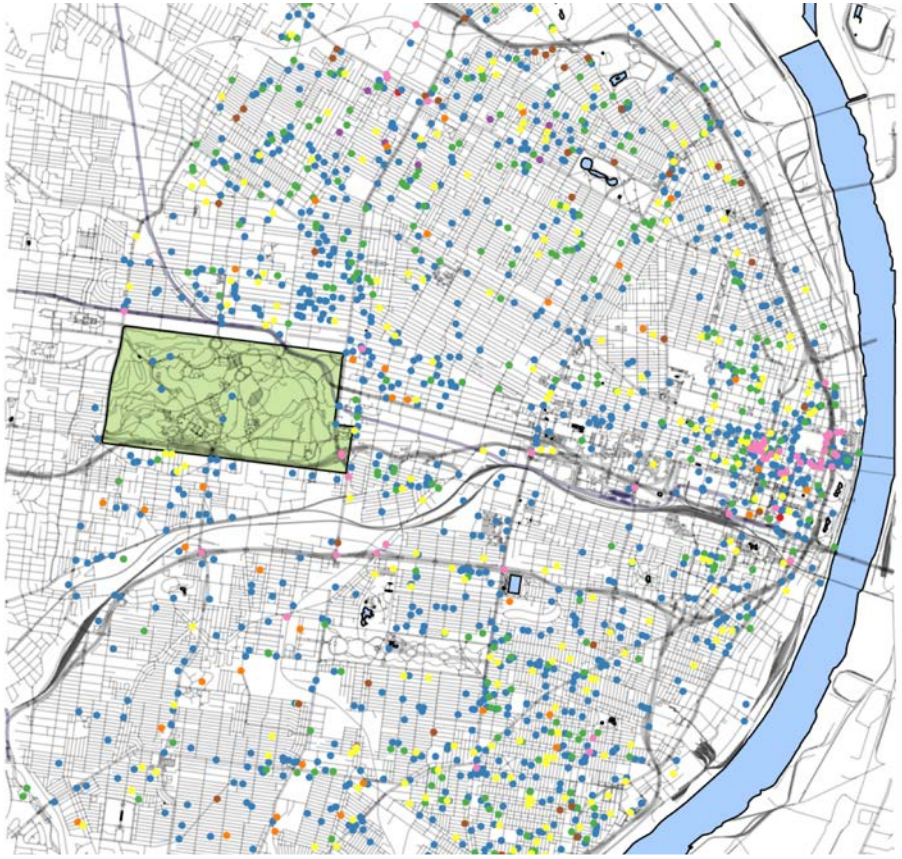

Homicide
Theft-related
Assault
Arson
Fraud
Vandalism
Weapons
Vagrancy
Larceny
3/1/17 2:00 PM

Geo-Based Task

Mixed Type



Type-Based Task



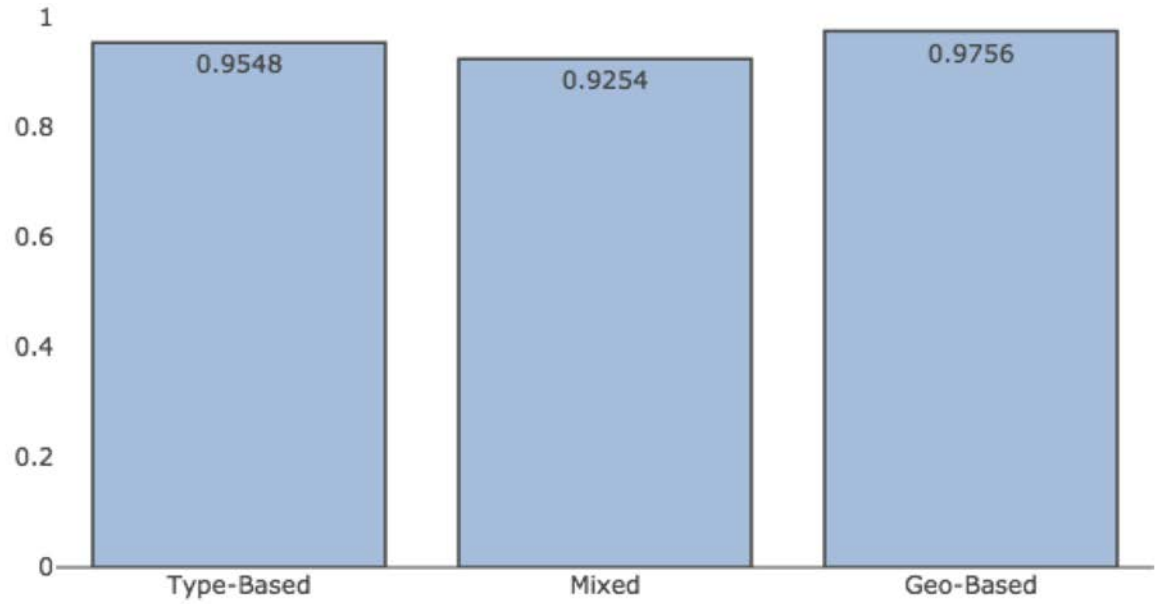
Data Collection & cleaning

- Mechanical Turk
- 30 participants:180 trials in total
- 78 trials remained (23, 27, and 28 trials for Type-Based, Mixed and Geo-Based tasks respectively

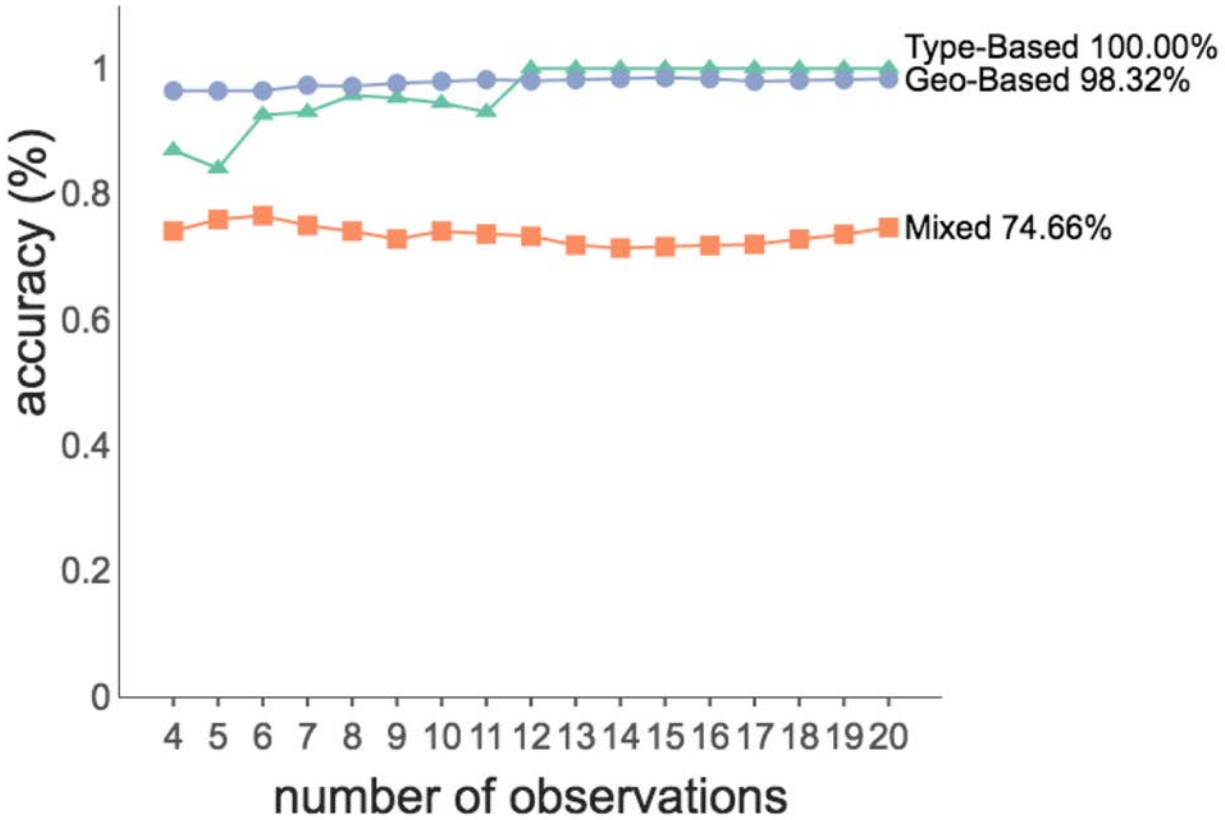
Apply Model

- 1. Initialize particles
- 2. Dynamical model
- 3. Make prediction
- 4. Observe a click
- 5. Calculate weight and resample
- 6. Back to 2

Prediction Accuracy



Accuracy Over Time



Design Implications

How would you design adaptive interface?

Design implication

- Prompt potential options
- Eliminate irrelevant information
- Prefetching data point
- Accessibility: Bubble Cursor
  - https://youtu.be/JUBXkD_8ZeQ?t=1m20s

Current Work: Predicting “high-level goals”

Research questions:

WHO: What learn personal characteristics?

- I love to talk
- Socializing helps me feel energized and inspired
- I like to solve problems by discussing them
- People often describe me as friendly and approachable

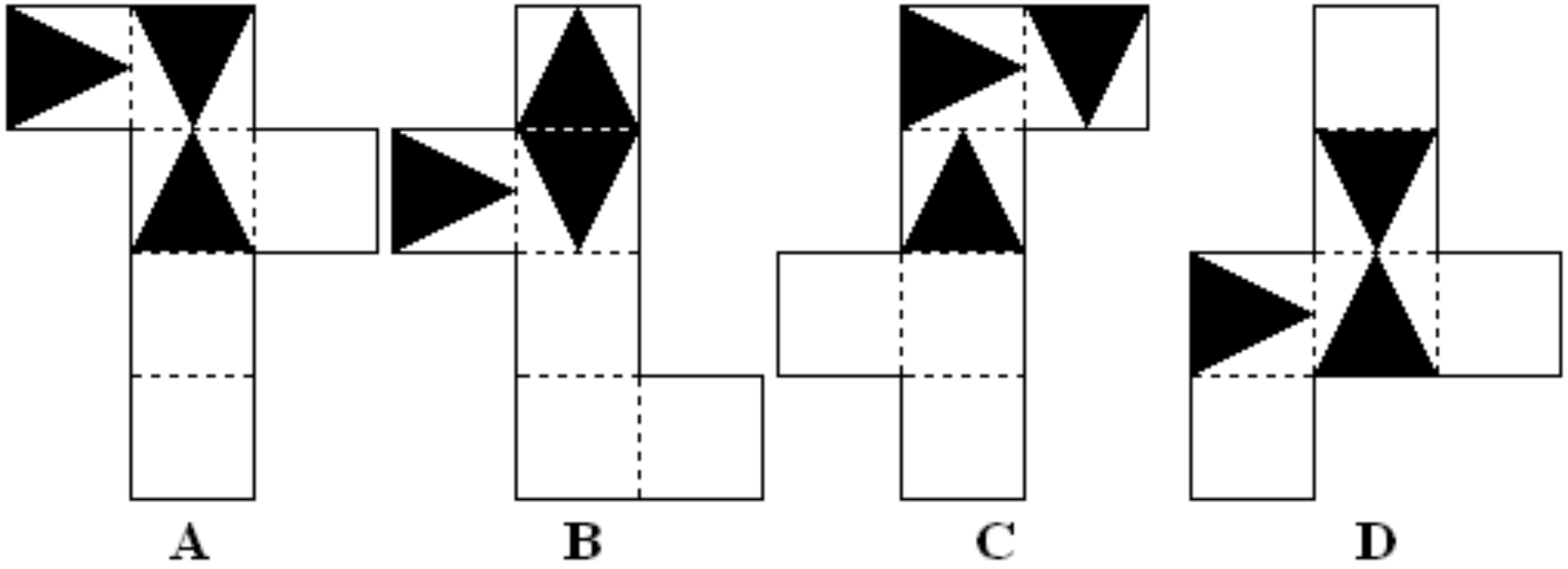
If these describe you then you are an **Extravert**

- What is going to happen will happen.
- No matter how hard you try, some people just don’t like you
- Getting a good job depends mainly on being in the right place at the right time

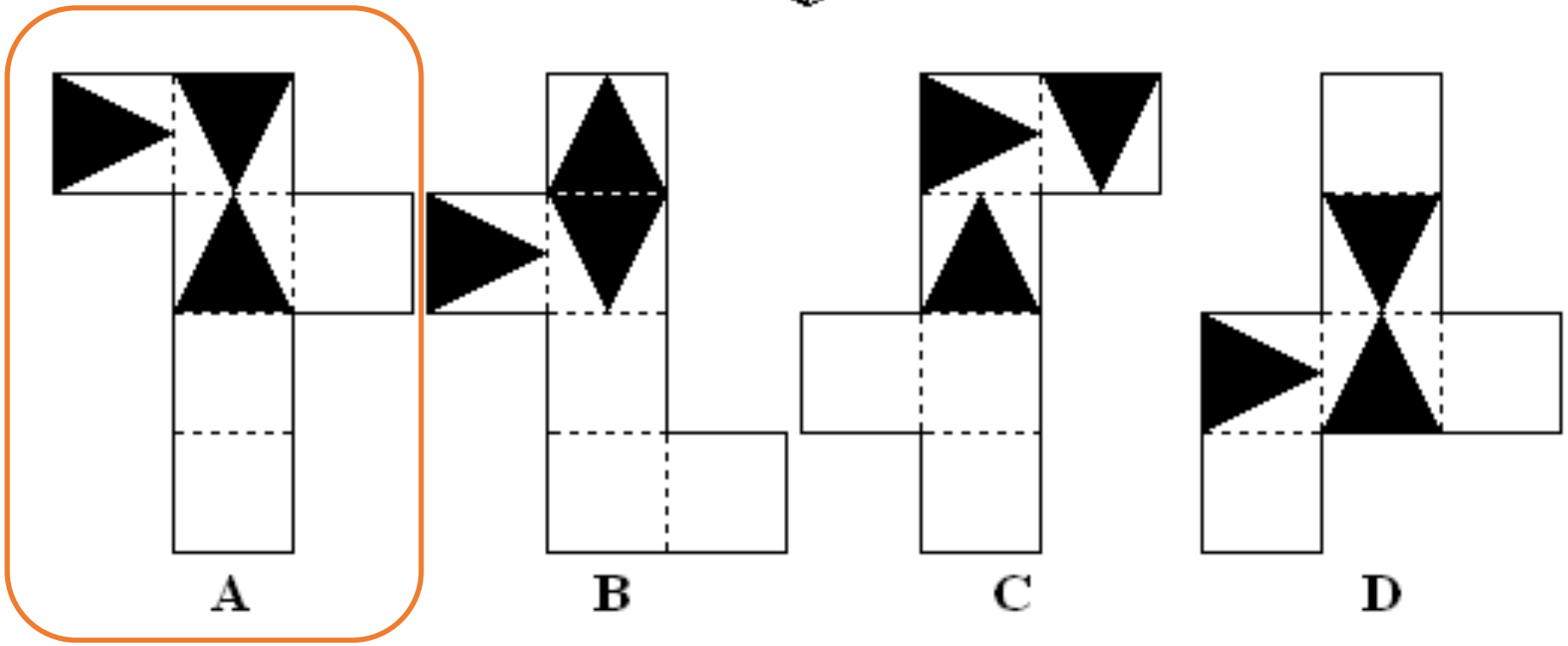
If you believe these to be true then you have an

External Locus of Control

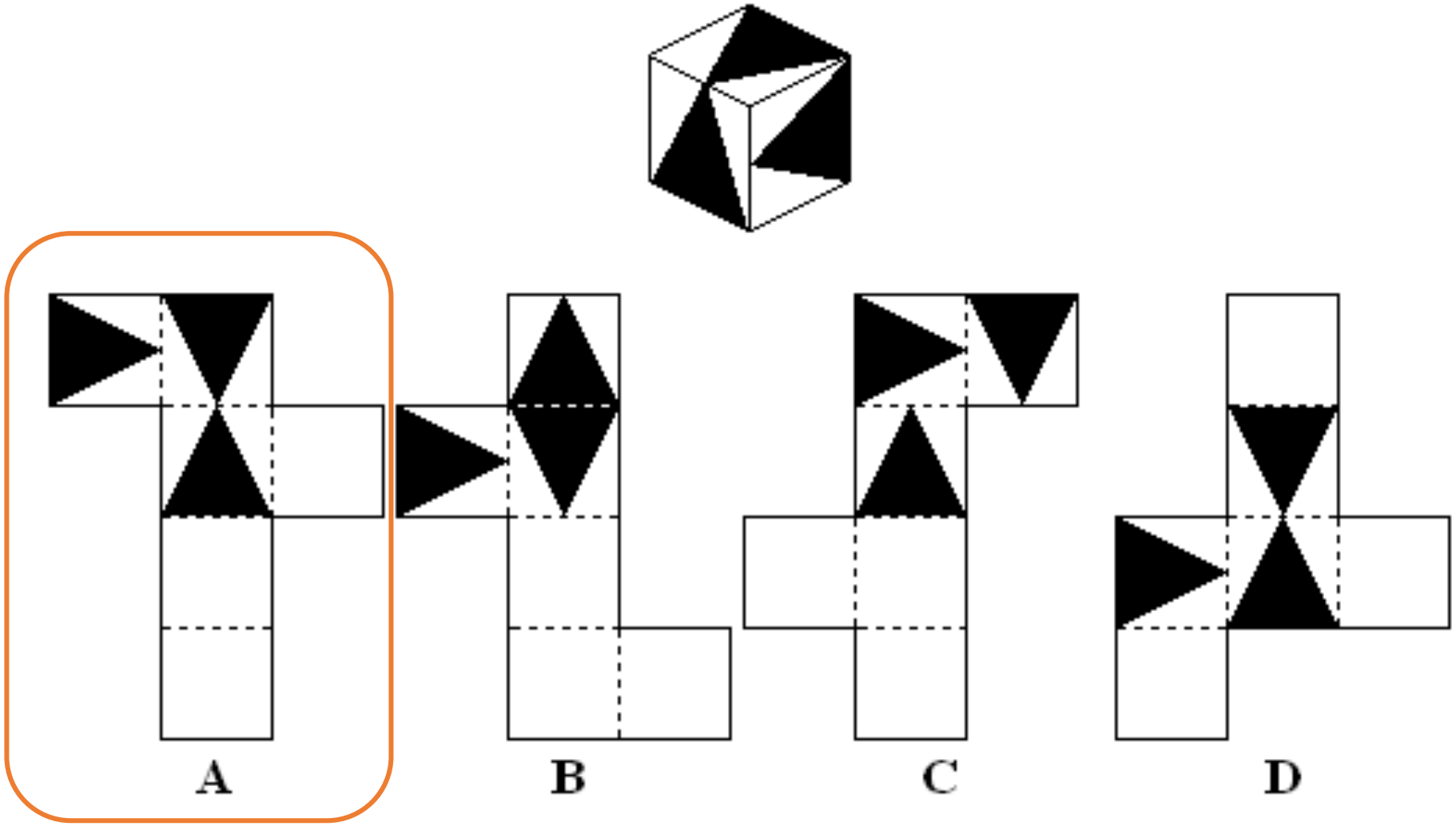
Which pattern can be folded to make the cube shown below?



Which pattern can be folded to make the cube shown below?



Spatial Ability measure how well we can mentally manipulate objects.



Individual Differences

- Personality
- Views
- Cognitive Abilities

Why should we care?

Communicating Bayesian reasoning is important for medical decision-making.

Consider the following problem:

The probability of breast cancer is 1% for women at age forty who participate in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also get a positive mammography.

If a woman at age 40 is tested positive, what are her chances of actually having breast cancer?

The chance of **actually** having breast cancer given a **positive** mammogram:

7.8%

95 out of 100 **doctors**[1] estimate this probability to be:

80%

[1]Eddy, David M. "Probabilistic reasoning in clinical medicine: Problems and opportunities." (1982).

The New York Times

HEALTH

359 COMMENTS

American Cancer Society, in a Shift, Recommends Fewer Mammograms

By DENISE GRADY OCT. 20, 2015

Email
Share
Tweet
Save
More

One of the most respected and influential groups in the continuing breast-cancer screening debate said on Tuesday that women should begin mammograms later and have them less frequently than it had long advocated.

The American Cancer Society, which has for years taken the most aggressive approach to screening, issued new guidelines on Tuesday, recommending that women with an average risk of breast cancer start having mammograms at 45 and continue once a year until 54, then every other year for as long as they are healthy and likely to live another 10 years.

The organization also said it no longer recommended clinical breast exams, in which doctors or nurses feel for lumps, for women of

A doctor reading mammograms at the University of Texas M.D. Anderson Cancer Center in Houston. The American Cancer Society said women should begin mammograms later and

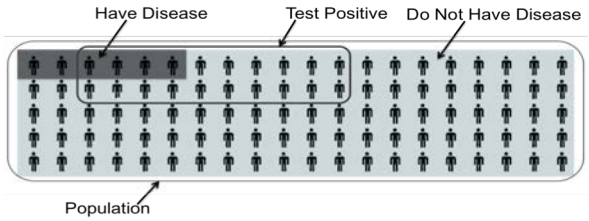
Can visualizations help?

There is a newly discovered disease, Disease X, which is transmitted by a bacterial infection found in the population. There is a test to detect whether or not a person has the disease, but it is not perfect.

There is a total of 1000 people in the population. Out of the 1000 people in the population, 10

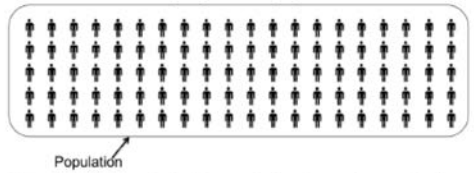
Imagine 1000 people are tested for the disease.
(a) How many people will test positive? ____
(b) Of those who test positive, how many will actually have the disease? ___

There is a newly discovered disease, Disease X, which is transmitted by a bacterial infection found in the population. There is a test to detect whether or not a person has the disease, but it is not perfect.

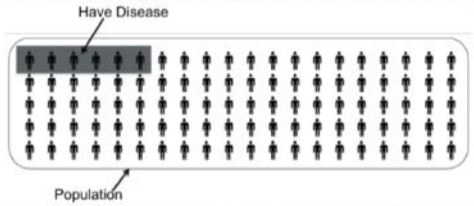


Imagine 1000 people are tested for the disease.
(a) How many people will test positive? ____
(b) Of those who test positive, how many will actually have the disease? ___

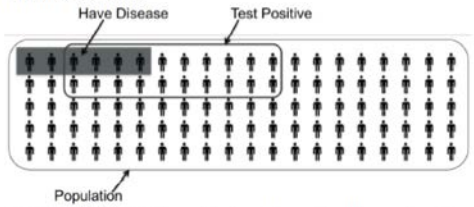
There is a total of 100 people in the population.



Out of the 100 people in the population, 6 people actually have the disease.

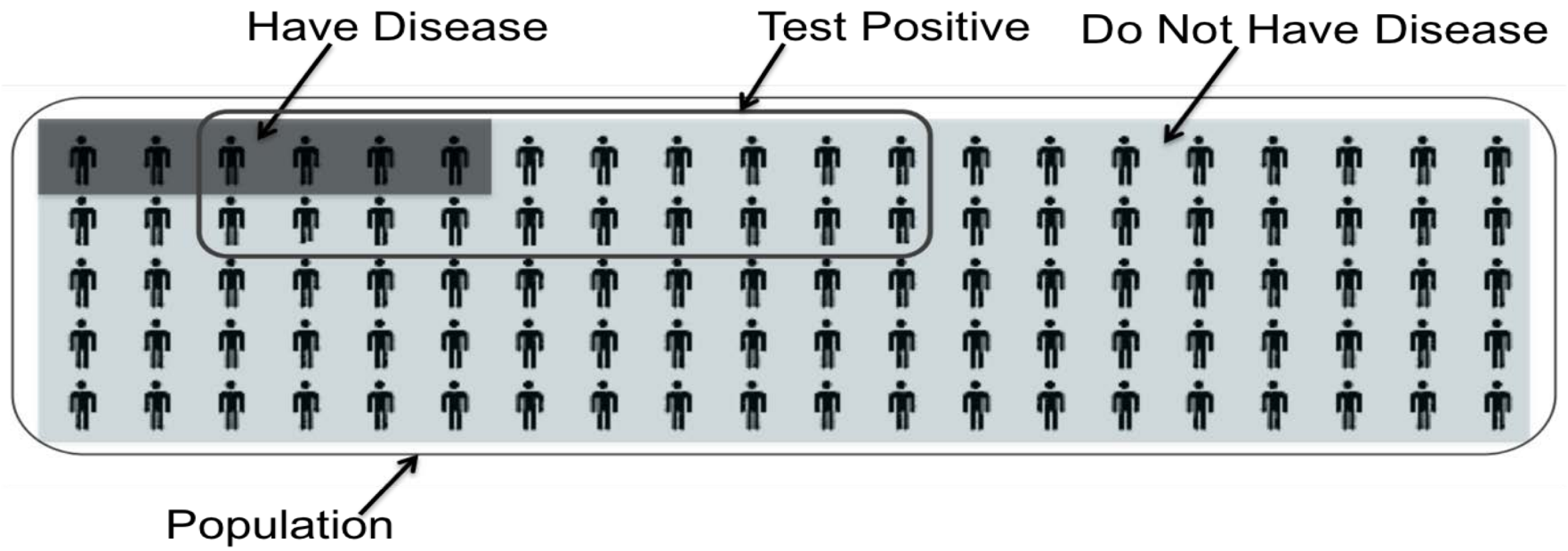


Out of these 6 people, 4 will receive a positive test result and 2 will receive a negative test result.



On the other hand, 94 people do not have the disease (i.e., they are perfectly healthy).

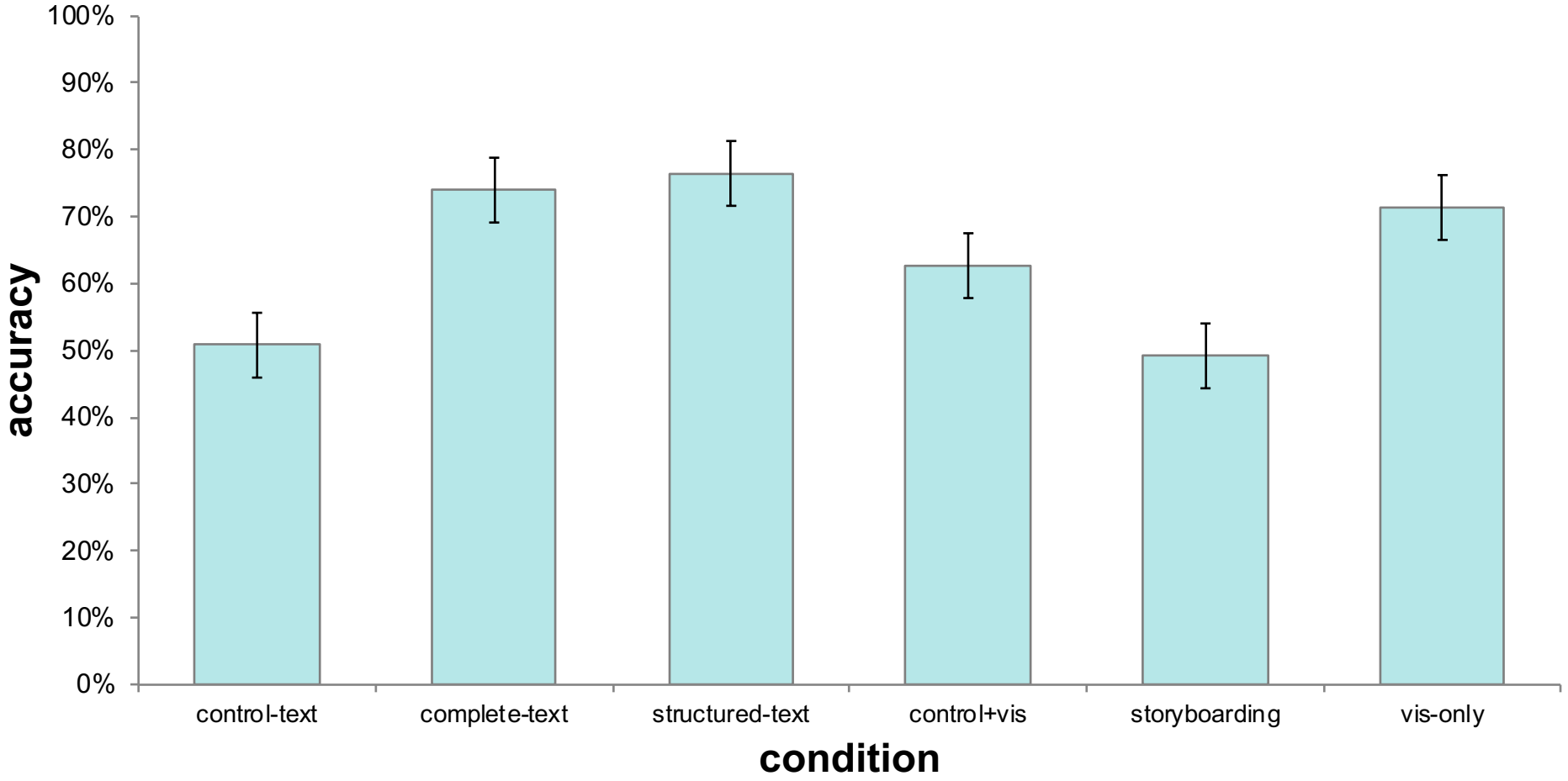
Why this visual representation?



Experiment Design

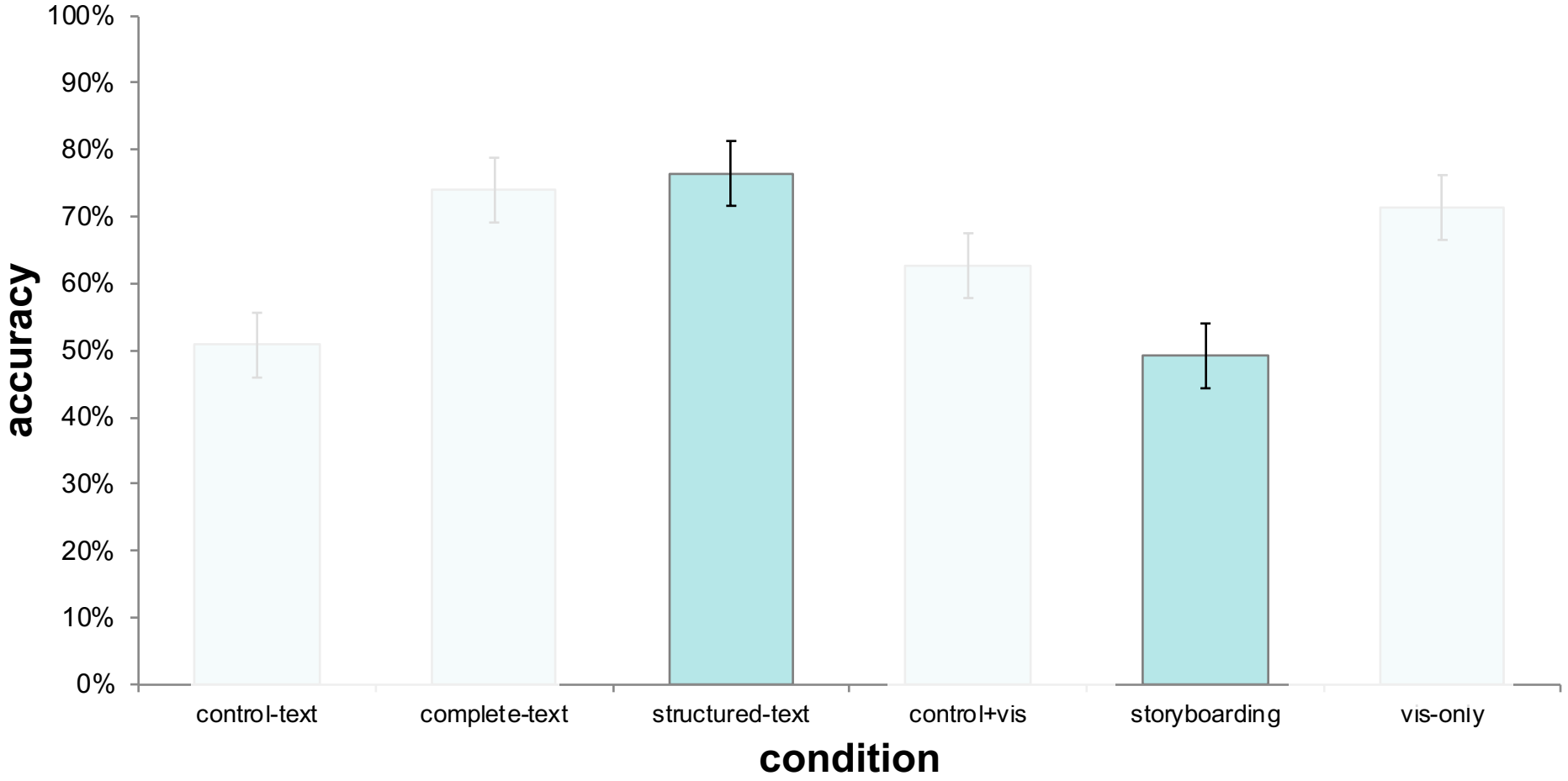
- 6 conditions
- 377 participants
- Also measured spatial ability

Findings



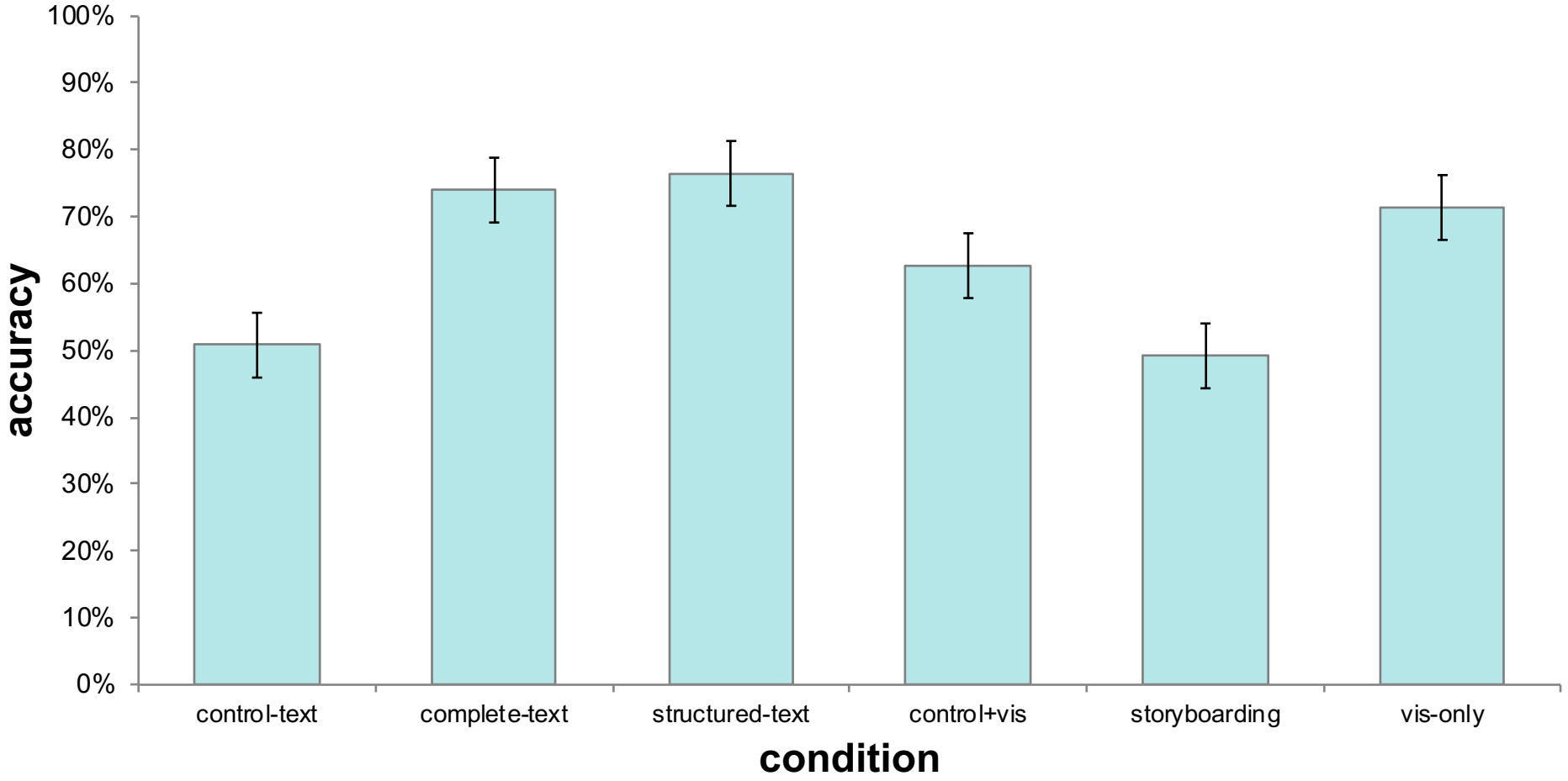
p = 0.004

Findings

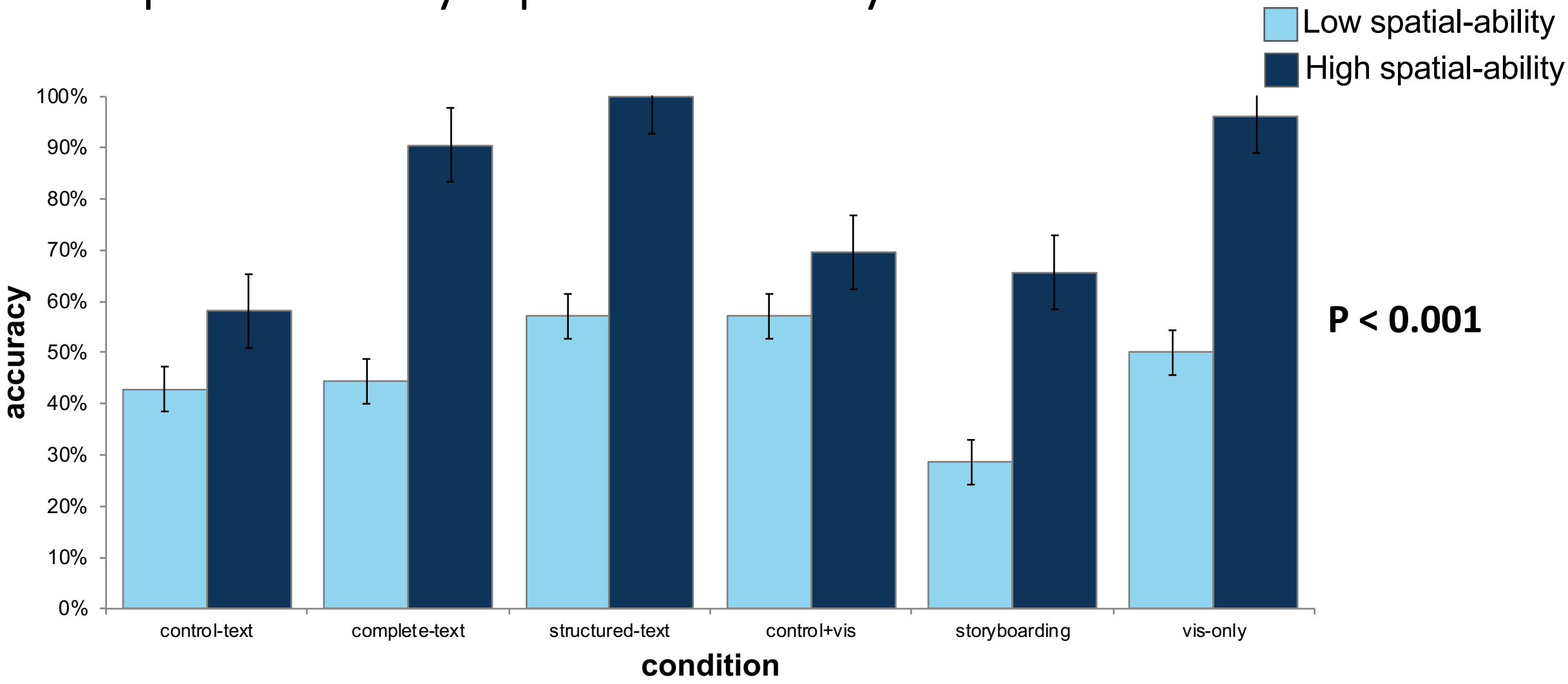


p = 0.003

Findings

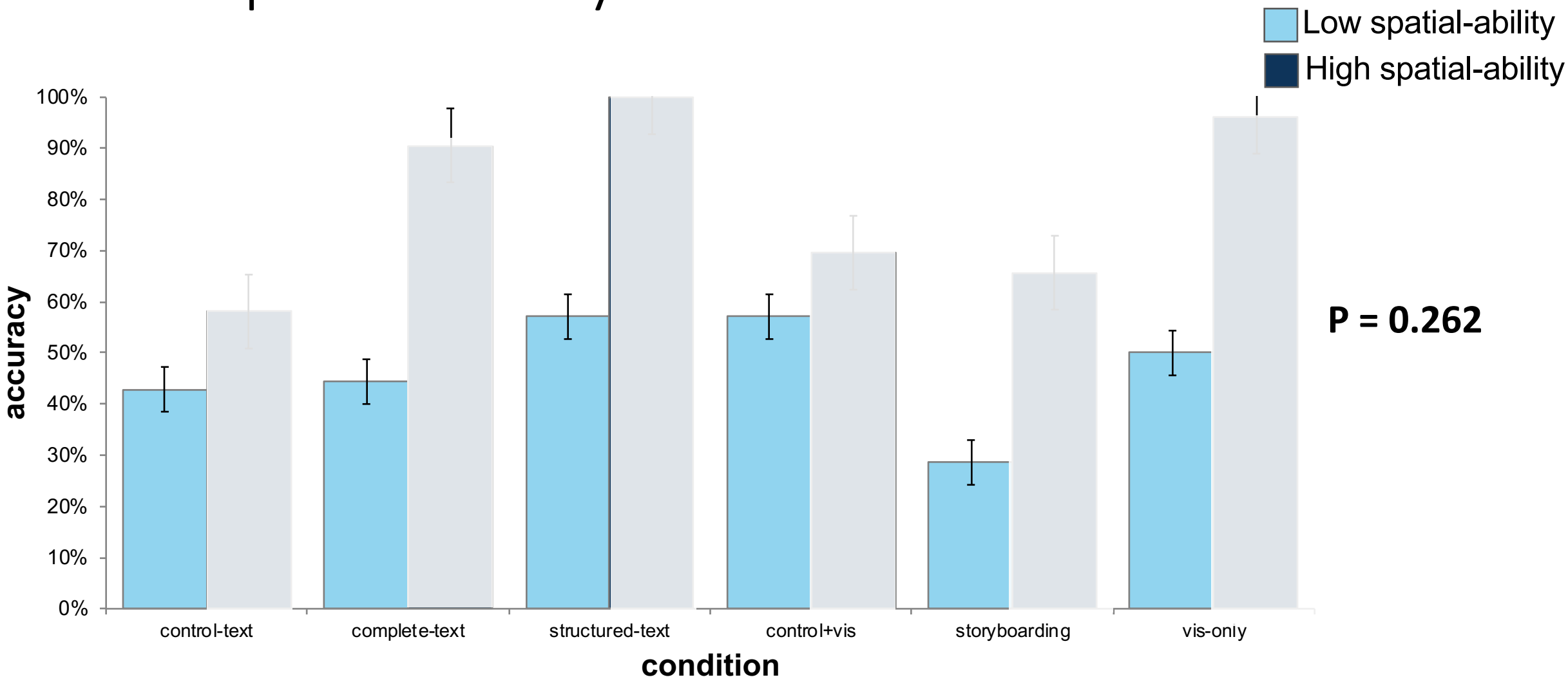

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
accuracy
control-text
complete-text
structured-text
control+vis
storyboarding
vis-only
condition

Separated by Spatial Ability



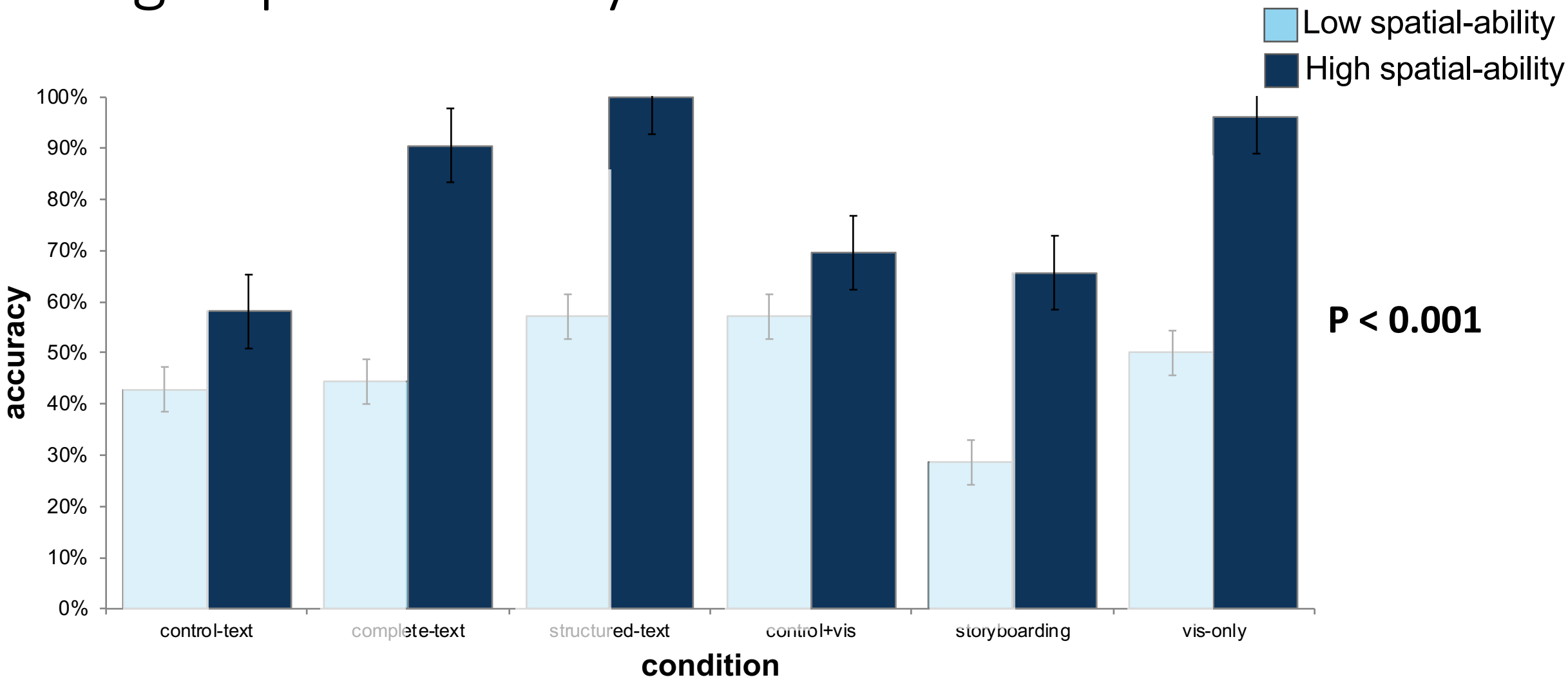
P < 0.001

Low Spatial Ability Users



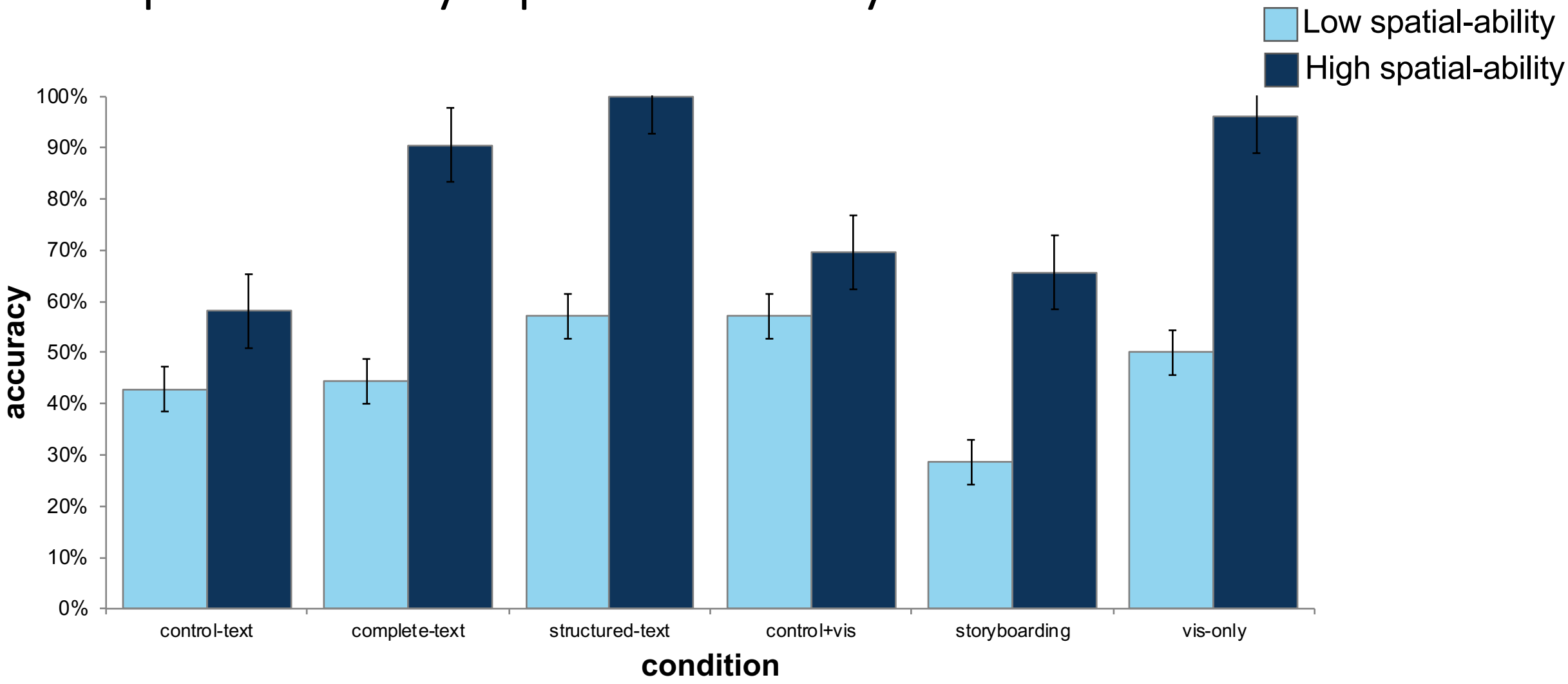
P = 0.262

High Spatial Ability Users



P < 0.001

Separated by Spatial Ability



Take away

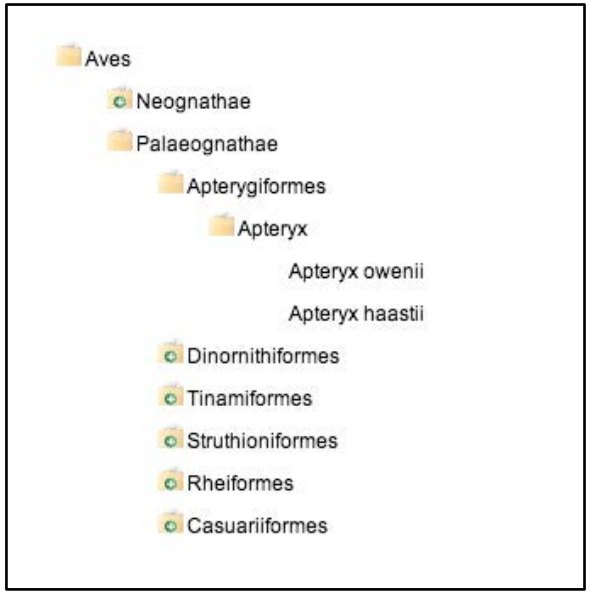
Who we are matters

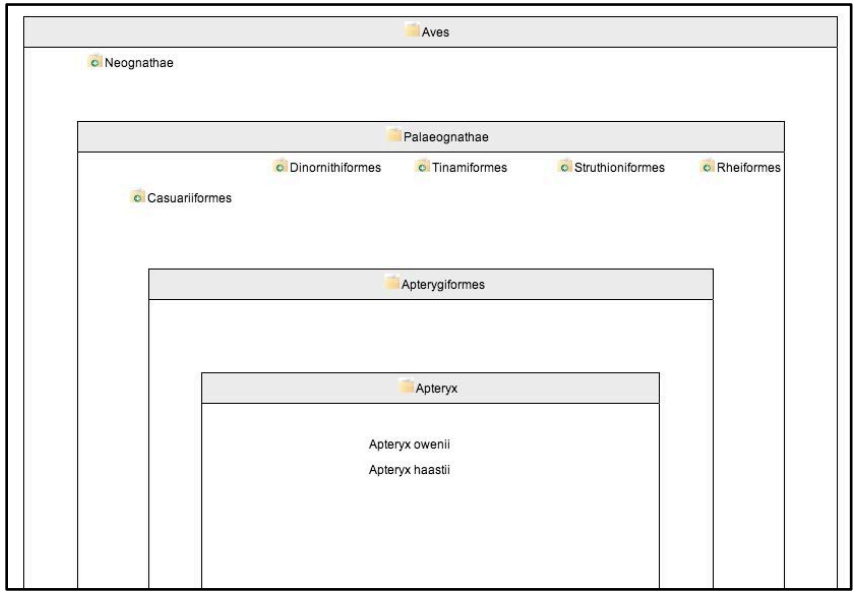
Improving Bayesian Reasoning: The effects of phrasing, visualization, and spatial ability

A. Ottley, E. M. Peck, D. Afergan, L. Harrison, H. Taylor, P. Han and R. Chang | InfoVis 2015.

What about other individual differences and tasks?

What about other individual differences and tasks?





Extraversion, Neuroticism, and Locus of Control can predict speed and accuracy

How visualization layout relates to locus of control and other personality factors

C. Ziemkiewicz, **A. Ottley**, R. J. Crouser, A. Yauilla, S. Su, W. Ribarsky, R. Chang | TVCG 2012.

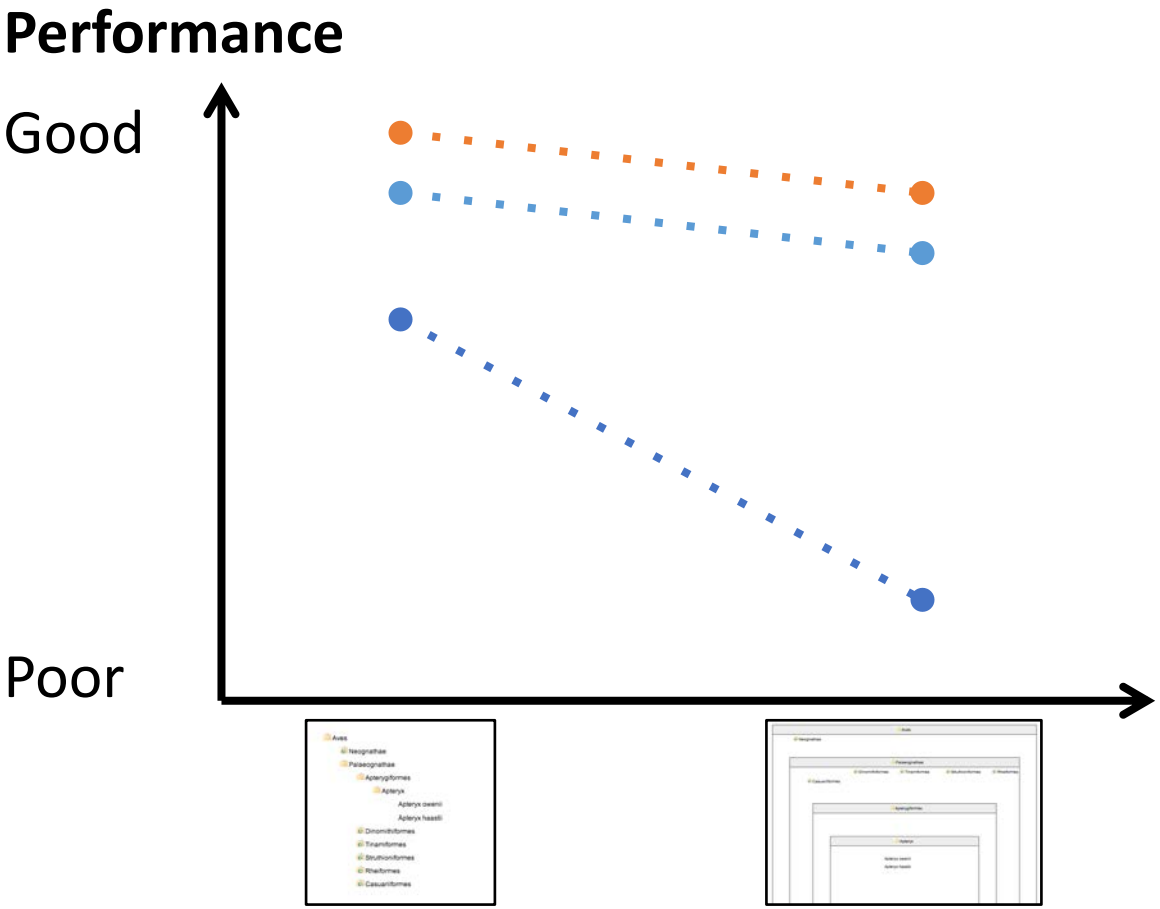
Locus of Control

Recall: Locus of Control describes the extent to which someone feels in control of external events.

External LOC
Believe that events are guided by fate, luck or powerful beings

Internal LOC
Believe that events are guided by their own actions

Locus of Control Results



External LOC
Average LOC
Internal LOC

P < 0.01

But how robust is this effect?

What is Priming?

Priming Locus of Control

- **Internals** can be primed by asking them to think about times when they **did not feel in control**

- **Externals** can be primed by asking them to think about times when they **felt in control**

Experiment Design

- 300 participants
- Primed locus of control
- Participants performed search tasks with the two hierarchical visualizations

Priming text for reducing locus of control

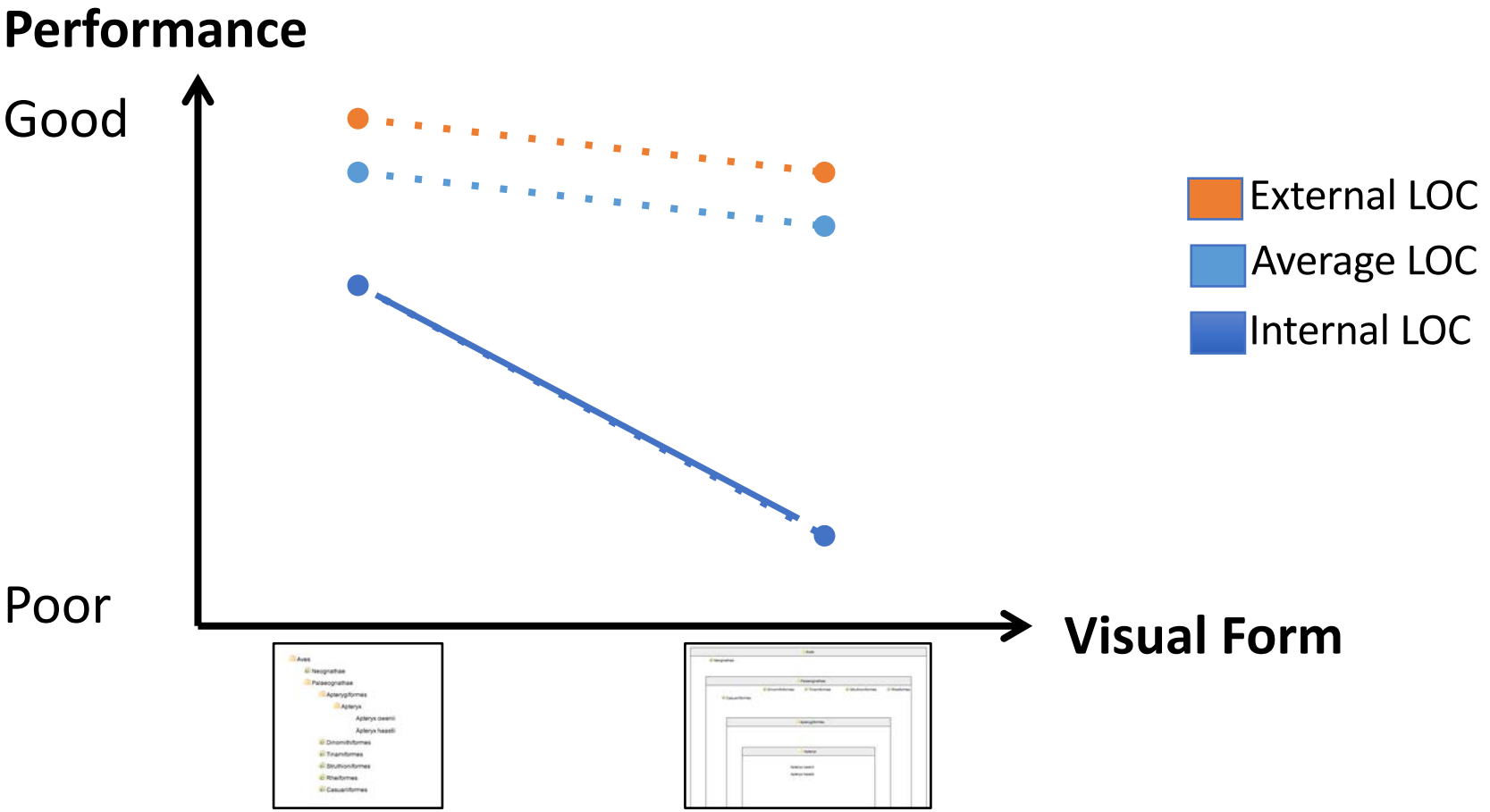
“We know that one of the things that influence how well you can do everyday tasks is the number of obstacles you face on a daily basis. If you are having a particularly bad day today, you may not do as well as you might on a day when everything goes as planned. Variability is a normal part of life and you might think you can’t do much about that aspect.

In the spaces provided below, give 3 examples of times when you have felt out of control and unable to achieve something you set out to do. Each example must be at least 100 words long.”

Hypothesis/ Expectations

Hypothesis/ Expectations

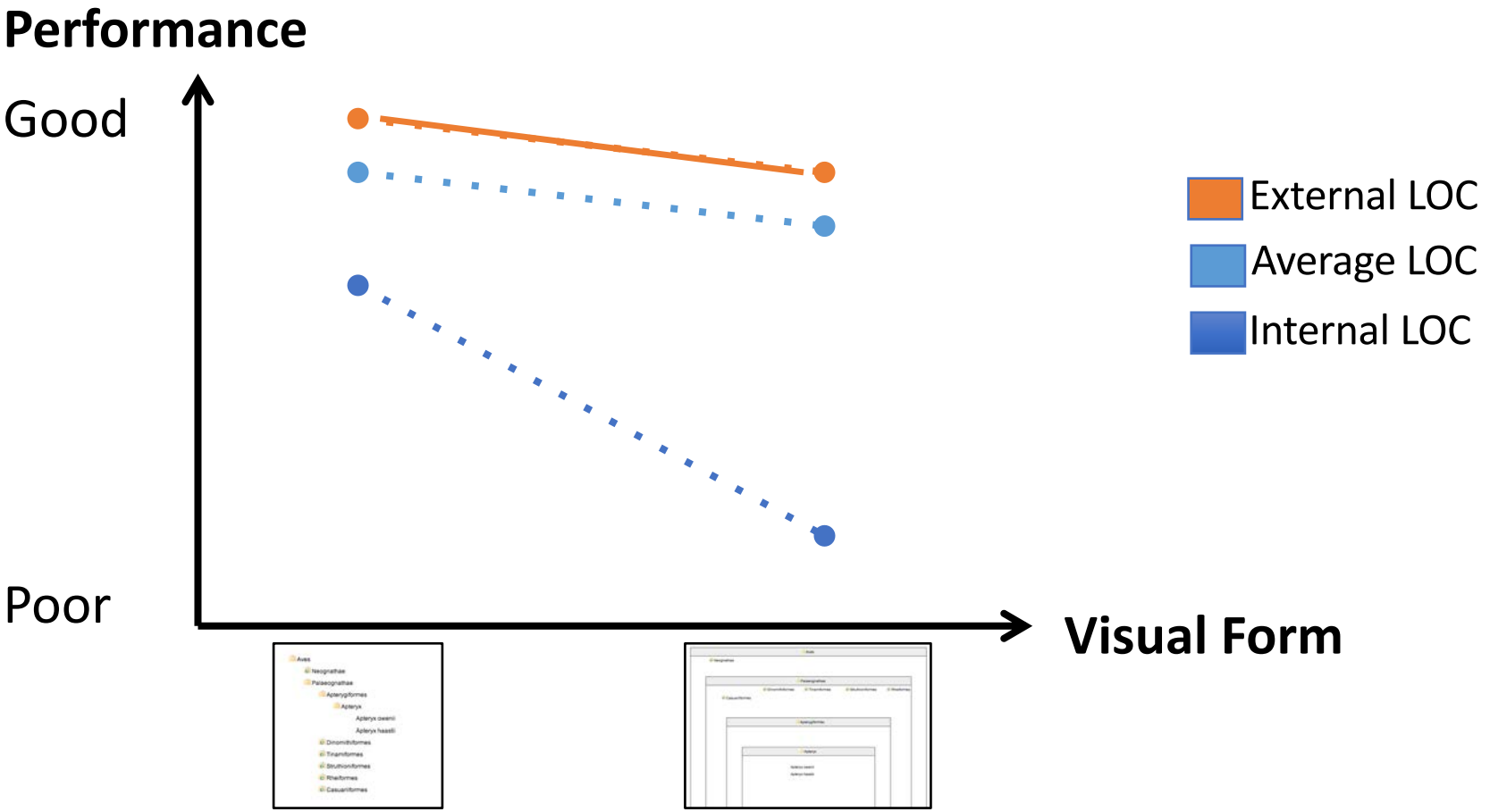
Summary of Results



Manipulating and Controlling for Personality Effects on Visualization Tasks

A. Ottley, C. Ziemkiewicz, R. J. Crouser, R. Chang | IVI 2013

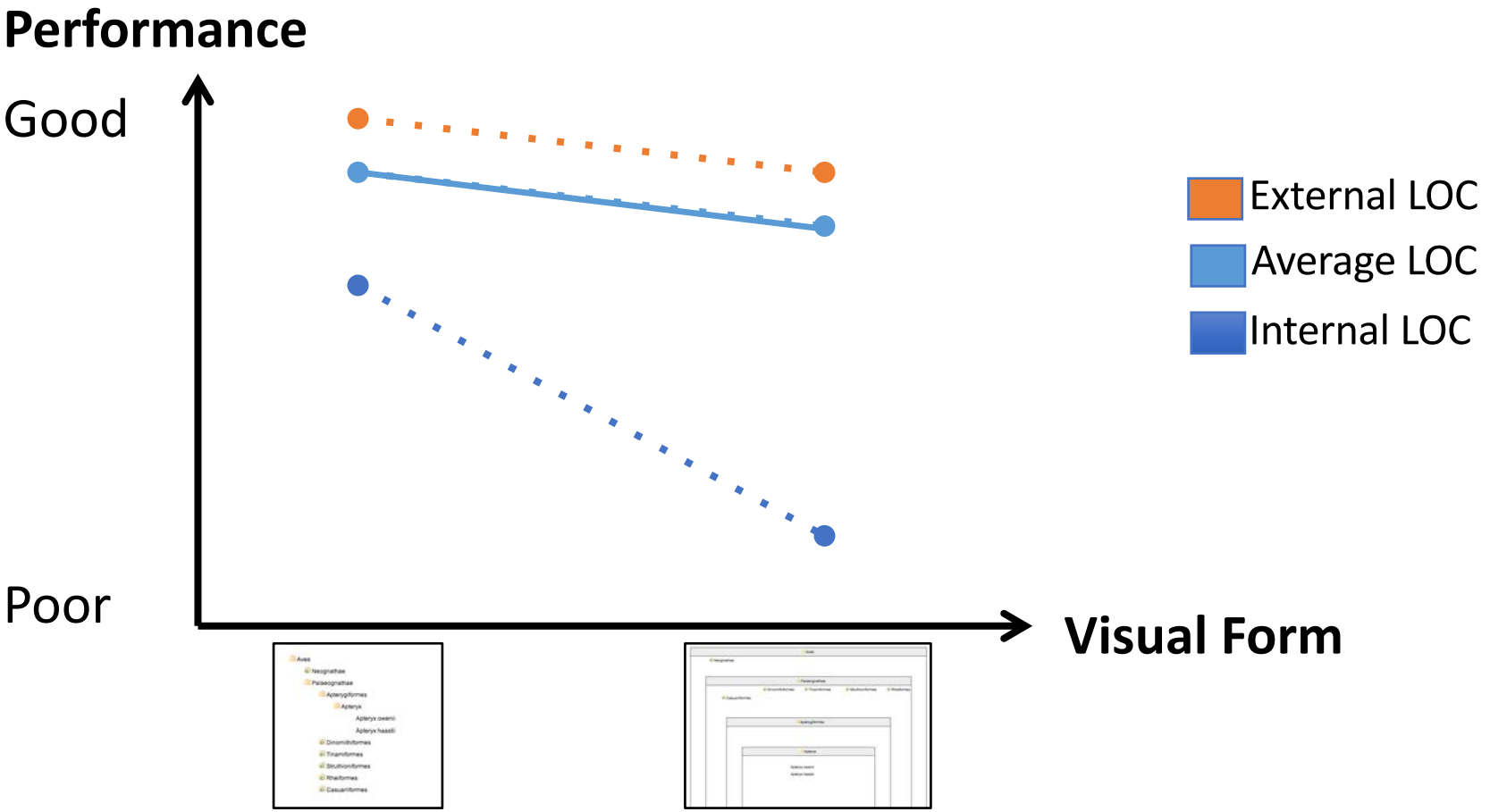
Summary of Results



Manipulating and Controlling for Personality Effects on Visualization Tasks

A. Ottley, C. Ziemkiewicz, R. J. Crouser, R. Chang | IVI 2013

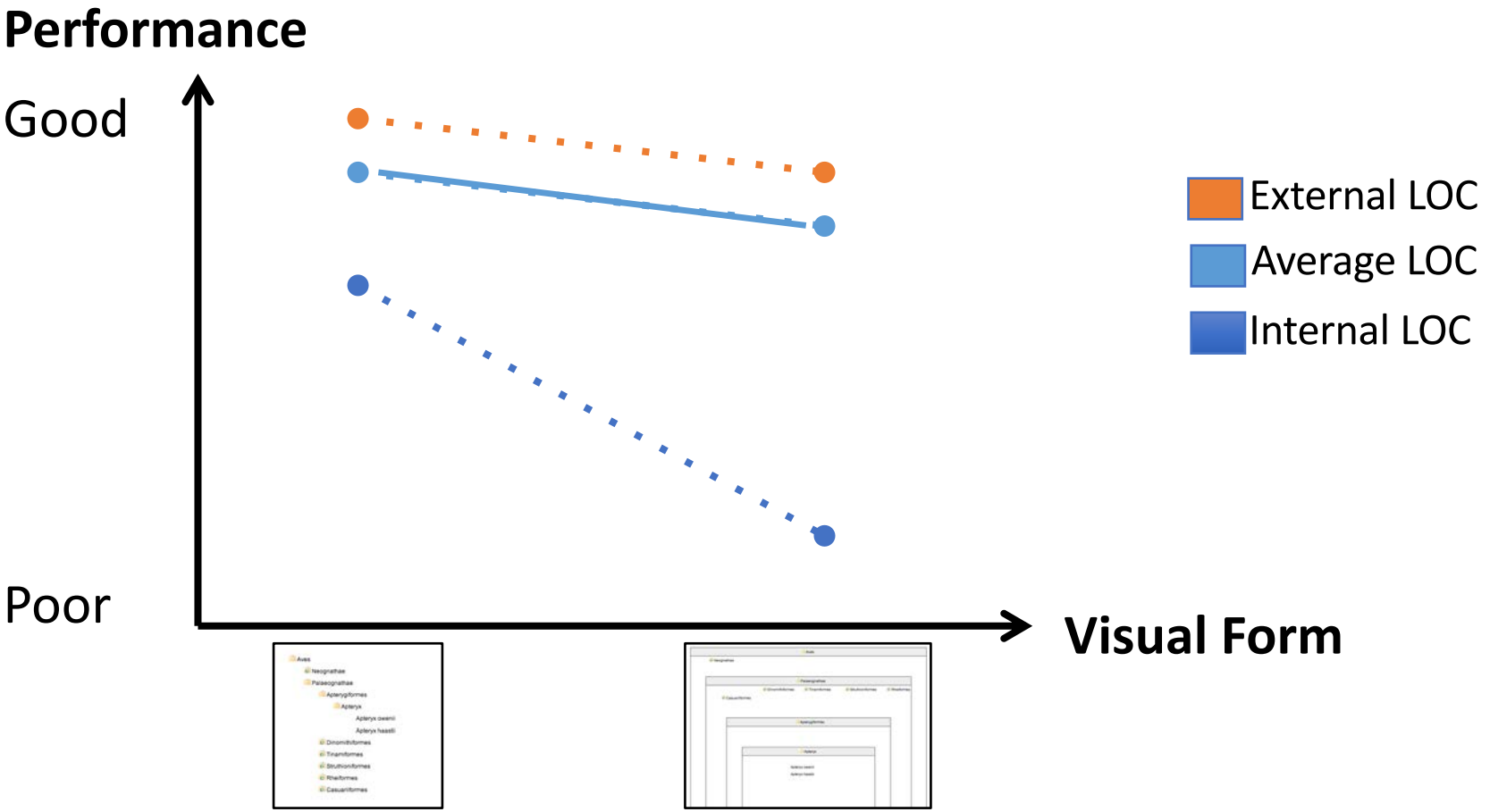
Summary of Results



Manipulating and Controlling for Personality Effects on Visualization Tasks

A. Ottley, C. Ziemkiewicz, R. J. Crouser, R. Chang | IVI 2013

Summary of Results



Manipulating and Controlling for Personality Effects on Visualization Tasks

A. Ottley, C. Ziemkiewicz, R. J. Crouser, R. Chang | IVI 2013

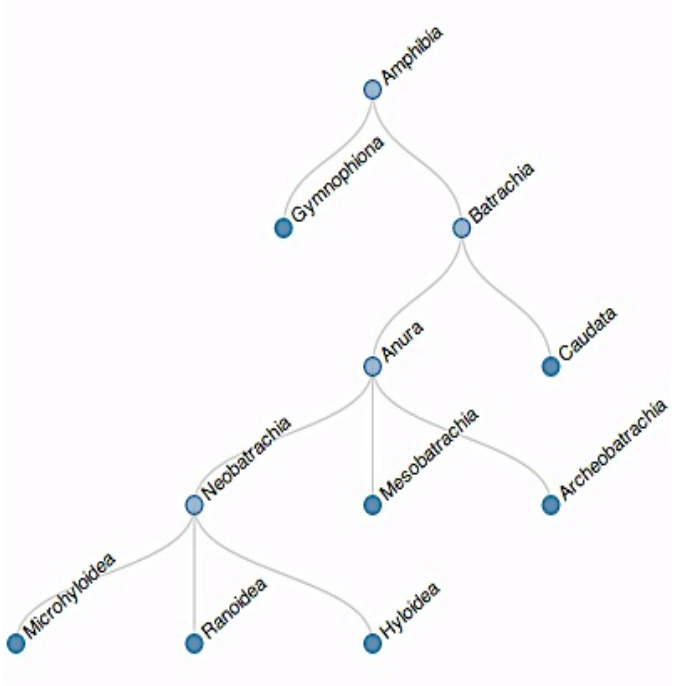
Take away

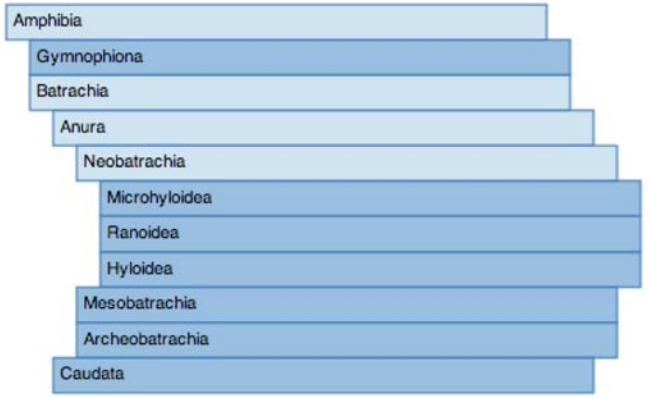
- Individual differences matter
  - Spatial Ability
  - Extraversion
  - Neuroticism
  - Locus of Control

- There is a strong connection between locus of control and performance

Do individual differences impact strategies?

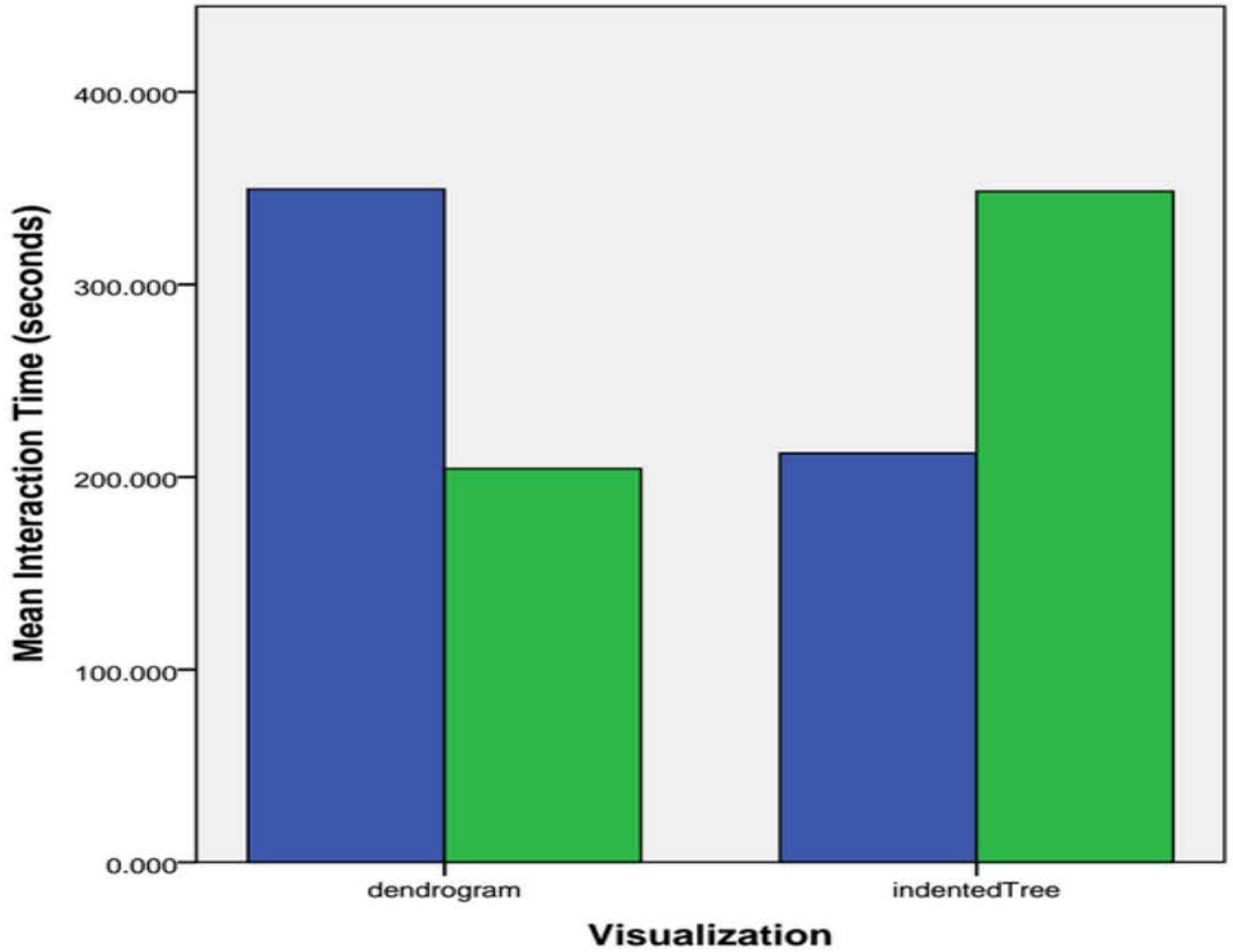
Does Locus of Control impact user strategies?





Experiment Design

- 54 participants
- Participants performed search tasks with the two hierarchical visualizations
- Captured mouse interaction

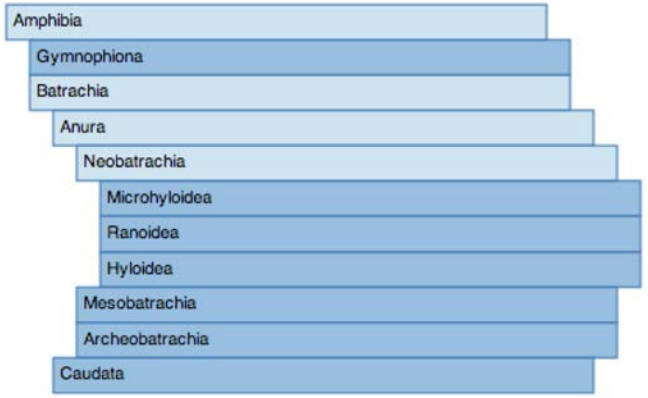

400.000
300.000
200.000
100.000
0.000
Mean Interaction Time (seconds)
dendrogram
indentedTree
Visualization
External
Internal

p = 0.03

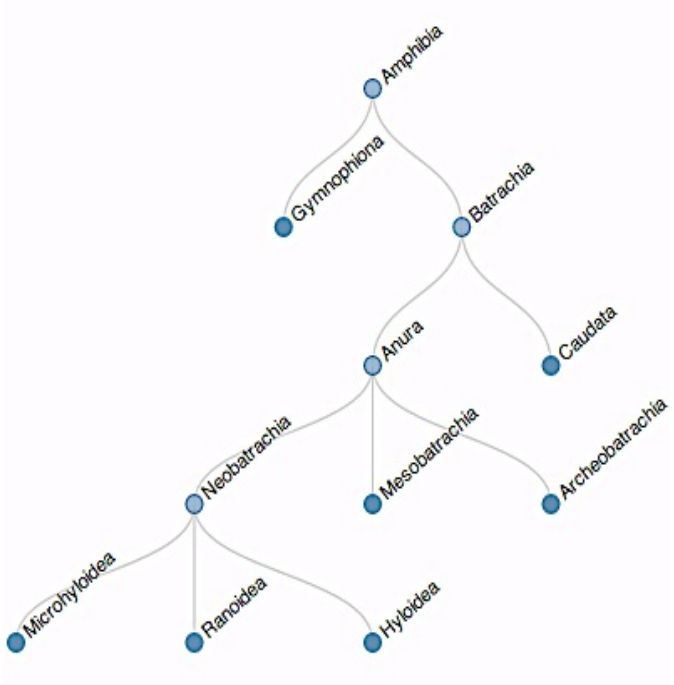
Results Highlights

Internals were more likely to perform a breadth-first search

Externals were more likely to perform depth-first search



Results Highlights

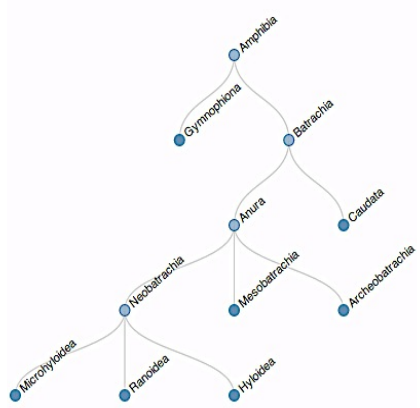


Internals were more likely to perform a combination of breadth-first and depth-first search

Externals were more likely to perform random search

Locus of control influences strategies

- Locus of Control affects not only speed and accuracy, but also search strategies
- Interaction logs contain a great deal of information!
- Manual analysis is time consuming and not scalable.



Other Interests

ADVICE: Active Decision and Visualization Interface for Collaborative Environments

Decision support for smart homes

supporting
medical decision-
making

Visualization of Tomorrow

Questions?

Visual Data Analysis Group