



# Research Methods

## The Cognitive Psychology of HCI

Stephen Gilbert, with material from Jonathan Kelly

# Diversity of Methods

Will focus on quantitative methods,  
mostly experimental

HCI relies on many methodologies,  
depending on the goal.

- See [Dix et al. 2003 Ch 9](#) for  
overview



<https://youtu.be/yGsHq-mZI8U>

# There's a new drug being studied.

200 participants who have heart issues.

100 get placebo, 100 get drug  
They come weekly to the clinic for 4 weeks.

They wear a heart monitor 24/7.

What are ways I  
can hack the  
study so that  
the results  
come out the  
way I want?

# Scientific Method

Based on empiricism

- Testable predictions
- Observation

Hypothesize

Operationalize

Measure

Evaluate

Revise or Replicate



# Hypothesize

Hypothesis: testable prediction about the conditions under which an outcome will occur

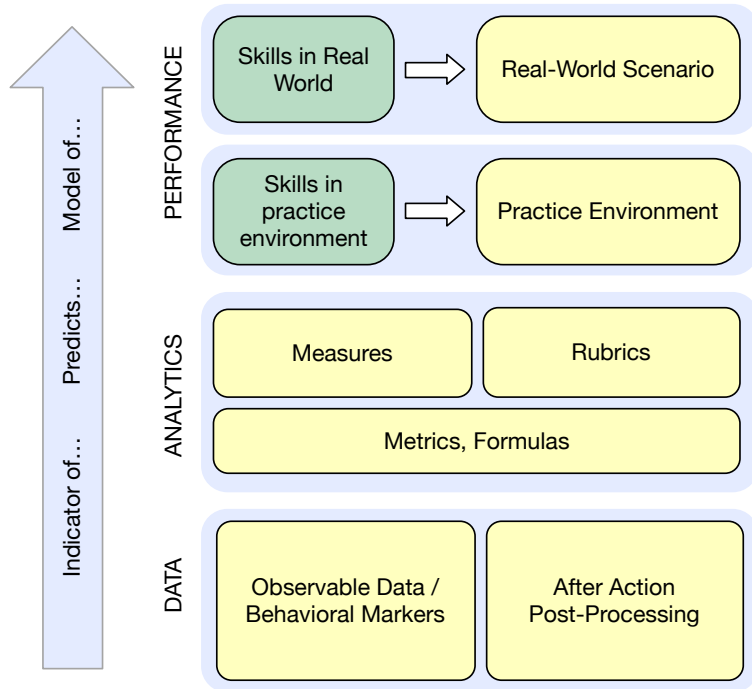
Examples:

- Workers perform worse under stress
- Student attention is highest in mid-day classes

At least 2 concepts and their relationship  
Can be operationalized and tested



# Operationalize



Conceptual

Performance  
Stress  
Attention  
Trust

Operational

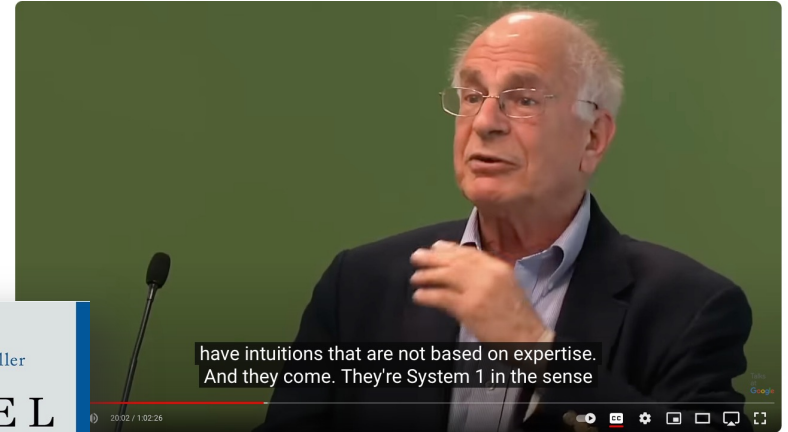
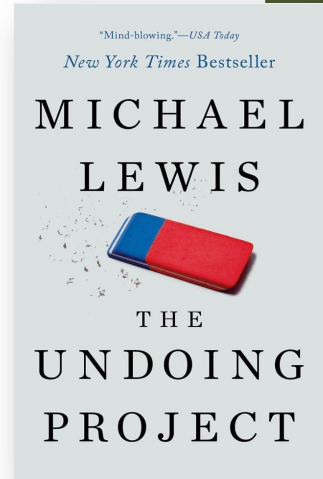
Errors  
Response time  
Heart rate  
Survey response

# Thinking, Fast and Slow

## Daniel Kahneman, Nobel Prize Winner

System 1 – Fast thinking, unconscious

System 2 – Slow thinking, deliberation





<https://youtu.be/CjVQJdIrDJ0>



# Decision making study

## Which food to eat?

- 7 digits vs 2 digits

Affective   
Cognitive 



Affective   
Cognitive 



Conceptual?

Operational?



# New Topic



# Measure

Observational

Correlational

Experimental

Quasi-experimental



# Measure

**Observational**

Correlational

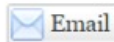
Experimental

Quasi-experimental



# Women wash hands more than men do

September 26, 2005 | From Times wire reports



## BRIEFLY

Men are dirtier than women. So scientists have confirmed by spying in public restrooms.

Wednesday's results mark the American Society of Microbiology's latest look at how many people take what is considered the single easiest step to staying healthy: spending 20 seconds rubbing with soap under the faucet.

Back in 1996, the society first studied how often people wash up after using the toilet. Researchers lingered in public restrooms, putting on makeup or combing their hair, while surreptitiously counting. They concluded about one-third of people did not wash. Every few years, researchers repeat the spying.

This time, 83% of people washed, reported Harris Interactive, a research company that last month monitored more than 6,300 public restroom users for the society.

\* The worst hygiene was at Atlanta's Turner Field baseball stadium, where 37% of men left the restroom without washing, compared with 16% of women who didn't wash.

\* New York's Penn Station had the biggest gender disparity, where 64% of men washed their hands compared with 92% of women. Grand Central Station was almost as bad.

\* The best hygiene was at San Francisco's Ferry Terminal Farmers Market and Chicago's Museum of Science and Industry and Shedd Aquarium, where only about 12% of people left without washing.

Not  
observational  
because they  
observe, but  
because that's  
ALL they do.

# Observational Methods

Describe the nature of a phenomenon by observing behavior

Two types:


- Unobtrusive observation (hidden cameras, etc.)
  - usually anonymous
- Participant observation

# Observational example

Why does phishing work?

Displayed  
fraudulent and  
authentic websites

[Dhamija, Tygar & Hearst \(2006\)](#)



The screenshot shows a web browser window with the address bar displaying a URL that mimics Bank of America's website. The page content includes the Bank of America logo and the text "Higher Standards" and "Online Banking". The main heading is "Confirm your Bank of America credit/debit card details". Below this, there is a paragraph of text explaining the confirmation procedure. The form contains several fields: "Type of banking" with radio buttons for "personal", "small business", and "corporate & institutional"; "Select your state" with a dropdown menu; "Your ATM or Credit Card Number" with a text input field; "Exiration date MM/YYYY" with two dropdown menus showing "01" and "2007"; and "Your ATM or Credit Card PIN" with a text input field. A large, semi-transparent blue watermark with the word "Phishing" is overlaid diagonally across the form.

# Observational example

23% ignored all browser-based cues

- Page content only

36% used page content + domain name

- Noticed change from domain name to IP address

9% used page content + domain + https

9% used page content + domain + https + ssl

- Some thought “Padlock more important on the page than browser” 

Could I do this study in an unobtrusive way? Not really.

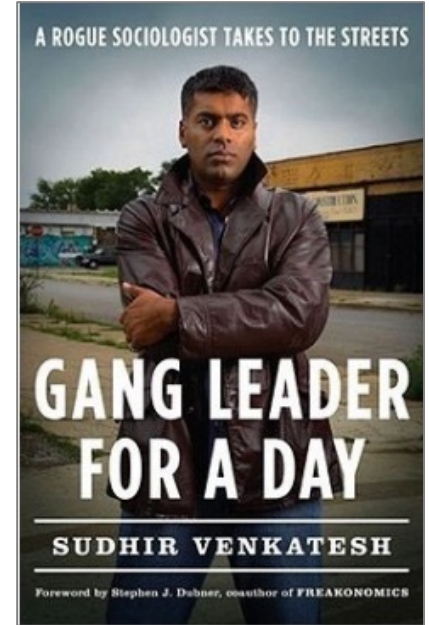
# Observational Methods

## Participant observation

- Observer interacts but tries not to influence
- Ethnography



Jane Goodall



Sudhir Venkatesh



# Strengths and limitations of observation

Strengths? Limitations?

# Strengths and limitations of observation

Strengths? Limitations?

Strengths: easy, natural environment

Limitations:

- Behavior can be difficult to observe
- Few predictions
- No causation
- Difficult to generalize
- Hawthorne effect



# Measure

Observational

**Correlational**

Experimental

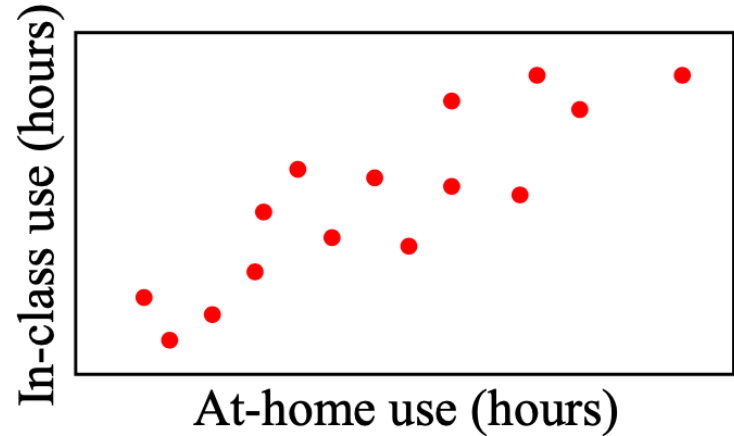
Quasi-experimental

# Correlational Research

Correlation: Relationship between variables

- How does Y change when X changes?

Teachers' home  
and in-class  
computer usage



# Strengths and limitations of correlational research

Strengths? Limitations?

# Strengths and limitations of correlational research

Strengths? Limitations?

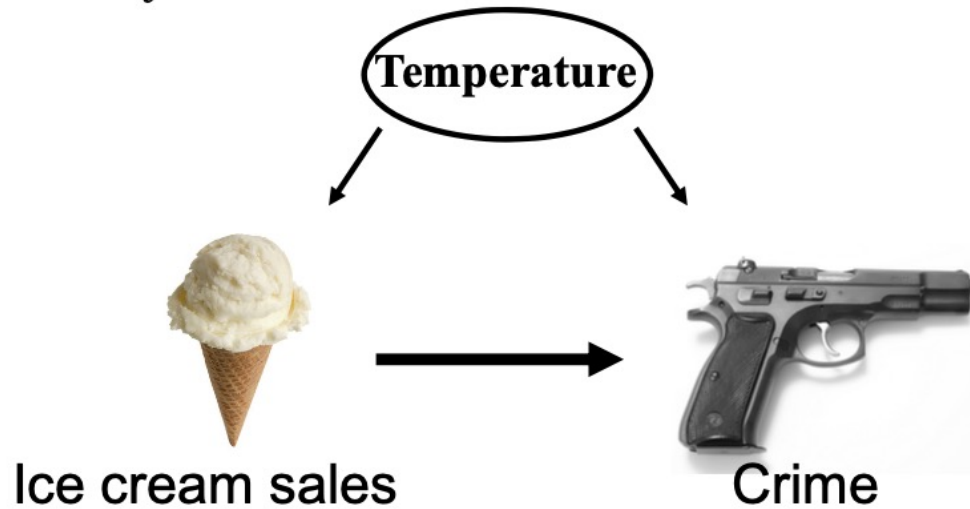
Strengths:

- Variables might be difficult or unethical to manipulate
- Surveys allow quick data collection

Limitations:

- Lack of causality
- 3<sup>rd</sup> variables

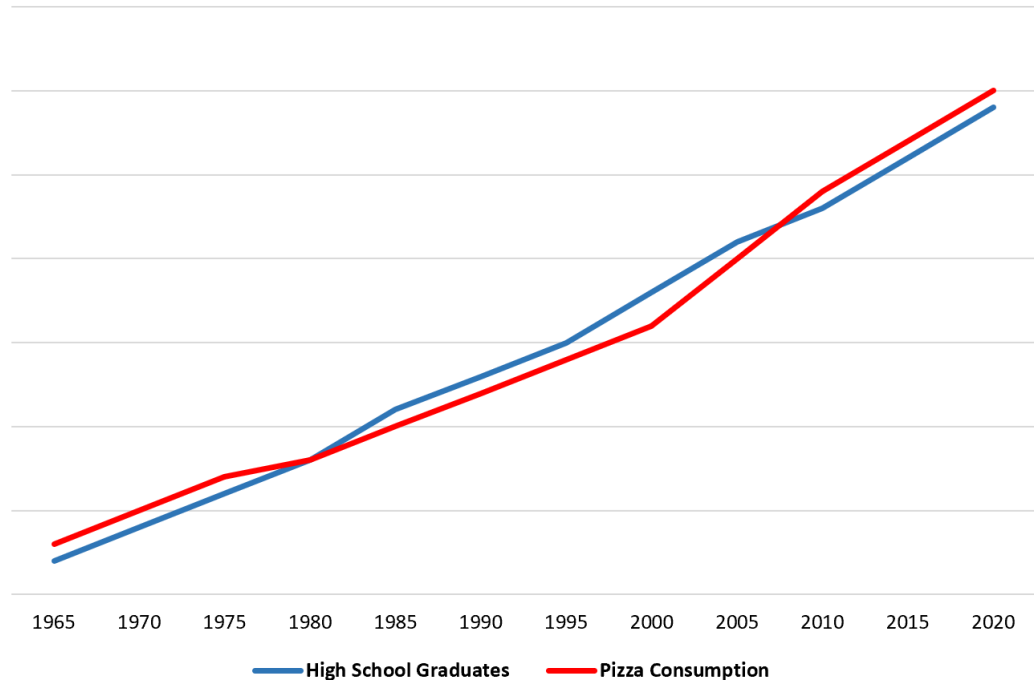
# 3<sup>rd</sup> variable problem



Never clear why  
variables are correlated

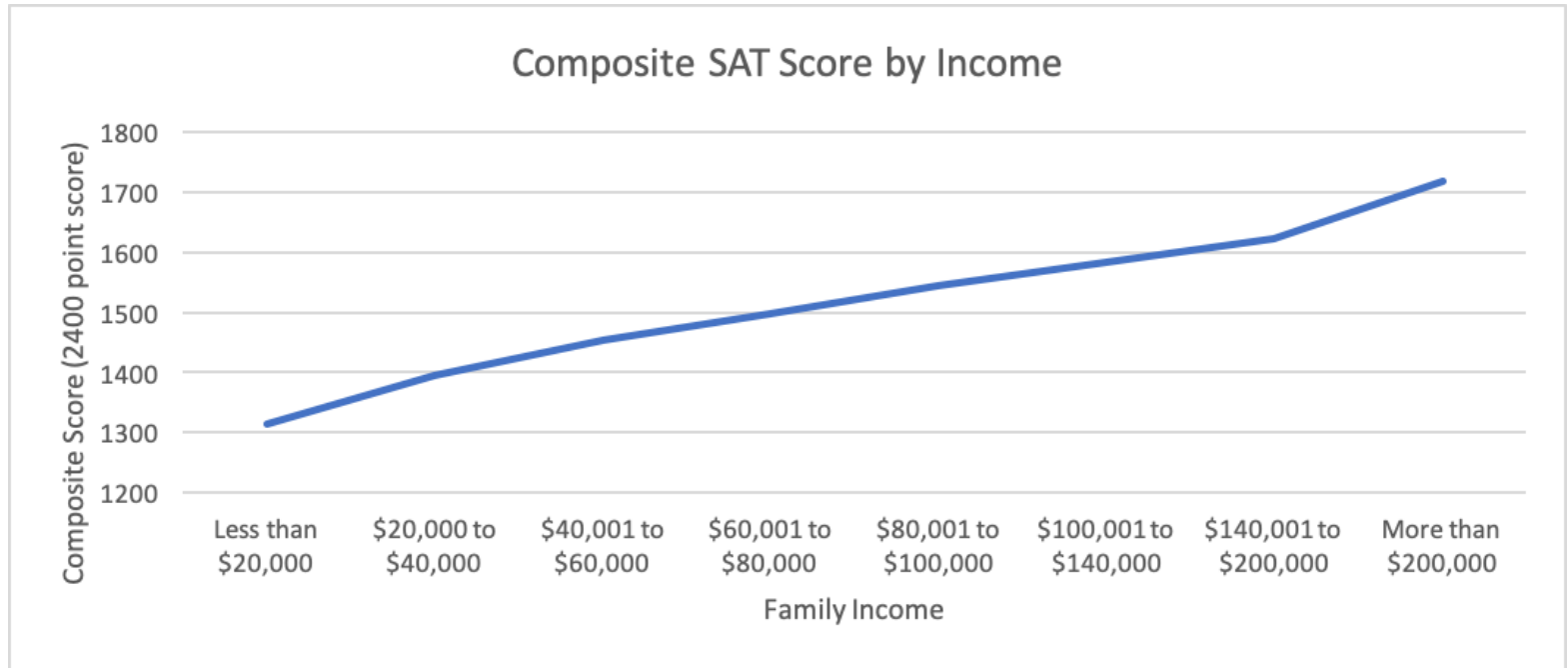
# 3<sup>rd</sup> variable problem

High School Graduates vs. Pizza Consumption





# 3<sup>rd</sup> variable problem



<https://www.futurescienceleaders.com/yvr1b/2019/01/08/standardized-testing-can-it-predict-college-success/>

# Measure

Observational

Correlational

**Experimental**

Quasi-experimental

# Measure

Observational

Correlational

**Experimental**

Quasi-experimental

Did change in X **cause** change in Y?

Manipulate one variable and observe effects on a second variable

# Independent vs Dependent Variables

## Independent variables (IVs)

- You **manipulate** to observe effects on another variable
  - 2-digit vs 7-digit number
  - Mouse vs touch interface
  - VR training vs. real-world training

## Dependent variables (DVs)

- You **measure** to determine the influence of the IV
  - Choice of food (fruit or chocolate cake)
  - Reaction time, pointing accuracy, test performance, etc.

## Causation

- Changes in DV can be only explained by changes in IV

# Manipulating IVs

## Within-participants design

- **Same** people in multiple conditions
  - 1<sup>st</sup> day: memorize 2-digit number and choose food
  - 2<sup>nd</sup> day: memorize 7-digit number and choose food



## Between-participants design

- **Different** people in each condition
  - Half memorize 2-digit, half memorize 7-digit



Strengths? Weaknesses?

# Within-Participants Design

## Strengths

- More sensitive
- Differences between people reflected in all conditions
- Fewer participants needed

## Limitations

- Carryover effects
- Practice or fatigue effects

## Counterbalancing

- Condition A → Condition B
- Condition B → Condition A

# Between-Participants Design

## Strengths

- No carryover, practice, or fatigue effects

## Limitations

- Requires more participants so that groups are similar

# Summary

Three unique requirements of a true experiment

1. A manipulated IV
2. Random assignment of participants
  - Ensures similar characteristics
3. Experimental control



# Measure

Observational

Correlational

Experimental

**Quasi-experimental**

# Quasi-experimental design

No random assignment to different levels

- Men vs. women vs. non-binary
- Android vs iOS users
- Gamer vs. non-gamer

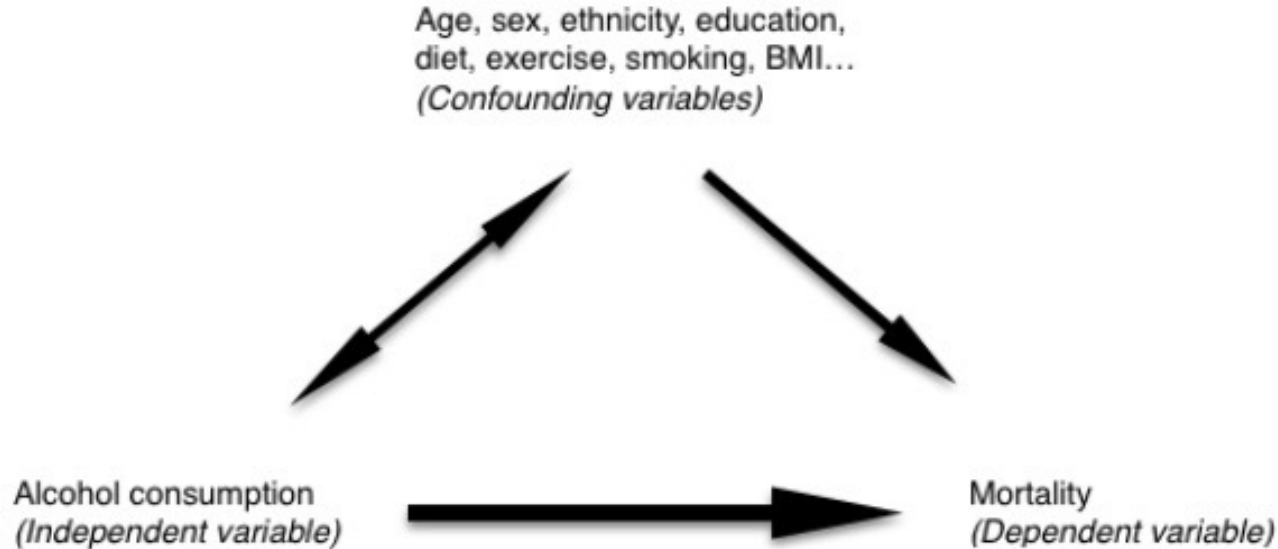
Looks like an experiment

- Multiple levels of an IV

Can't make causal conclusions

**Confounding variables** are a problem, but you can sometimes account for them with multiple regression.

# Quasi-experimental design: confounding variables



# New Topic



# Multitasking Study

Multitaskers perform many tasks at once

Do multitaskers have greater cognitive control?

[Ophir, Nass & Wagner \(2009\)](#)

# Multitasking Study

Multitaskers perform many tasks at once

Do multitaskers have greater cognitive control?

Questionnaire

- YouTube, music, games, web browsing...
- How much do you use X per week?
- When using X, how often do you concurrently use the other media?

[Ophir, Nass & Wagner \(2009\)](#)

# Multitasking Study

Multitaskers perform many tasks at once

Do multitaskers have greater cognitive control?

Questionnaire

- YouTube, music, games, web browsing...
- How much do you use X per week?
- When using X, how often do you concurrently use the other media?

Heavy vs. Light Media Multitaskers

- HMM vs. LMM

[Ophir, Nass & Wagner \(2009\)](#)

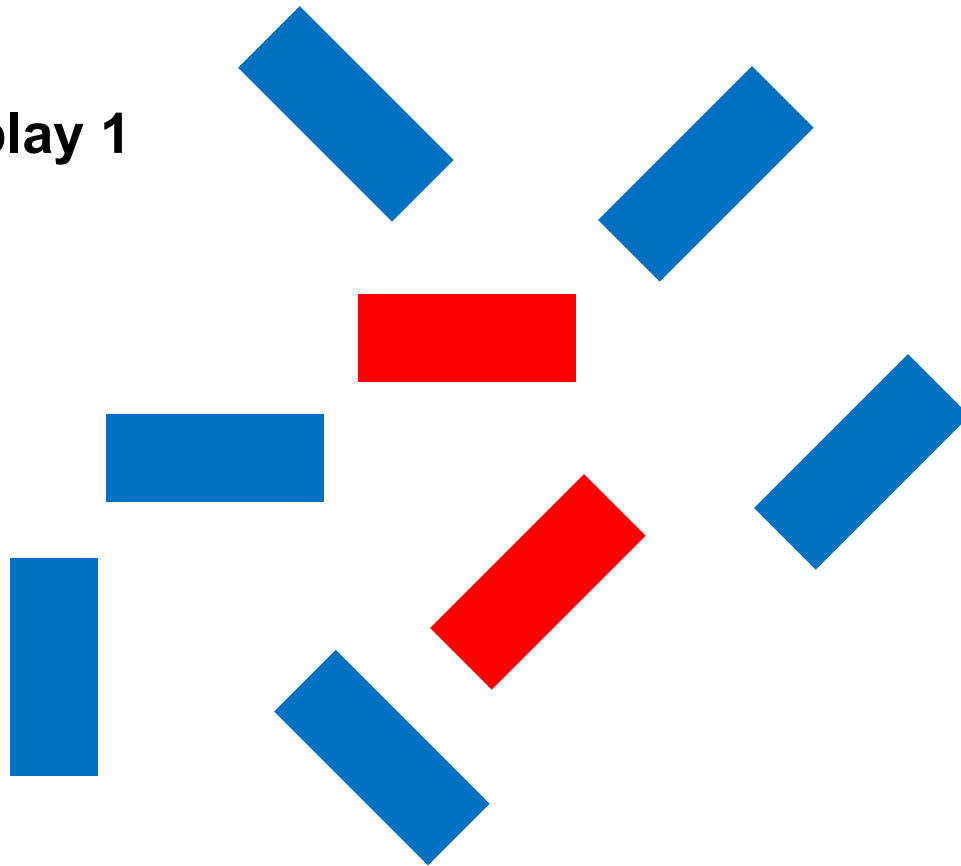
# Multitasking Study

## Cognitive control task

- Filtering environmental distractions



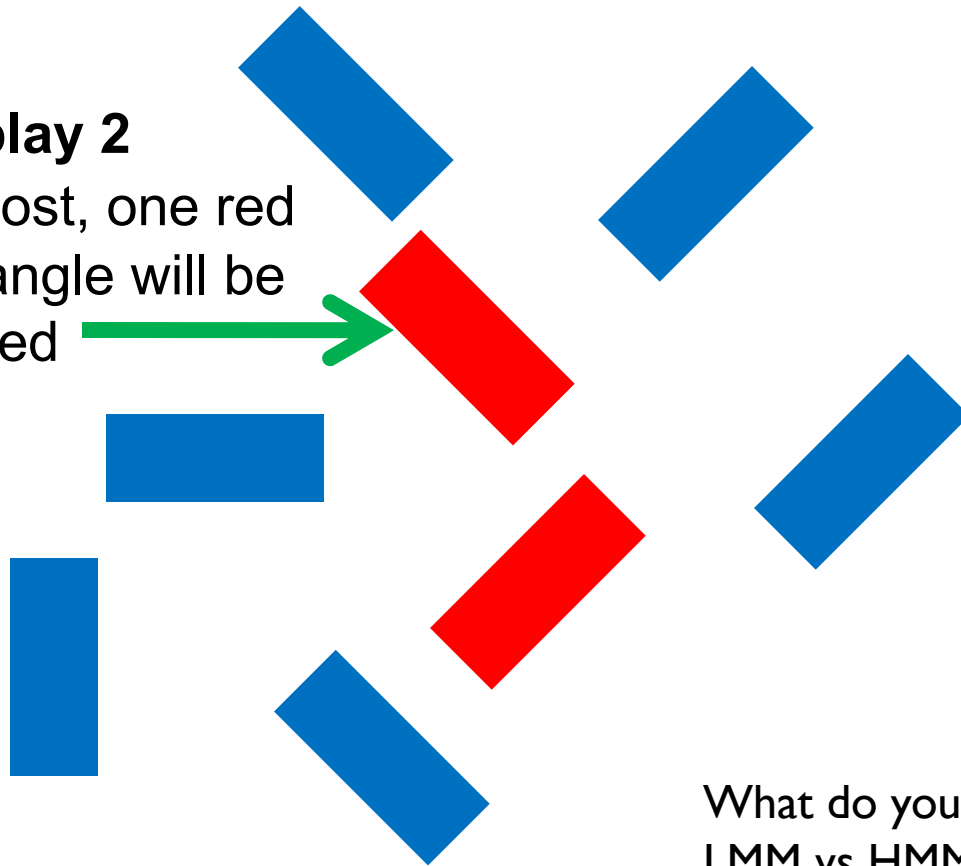
# Display 1





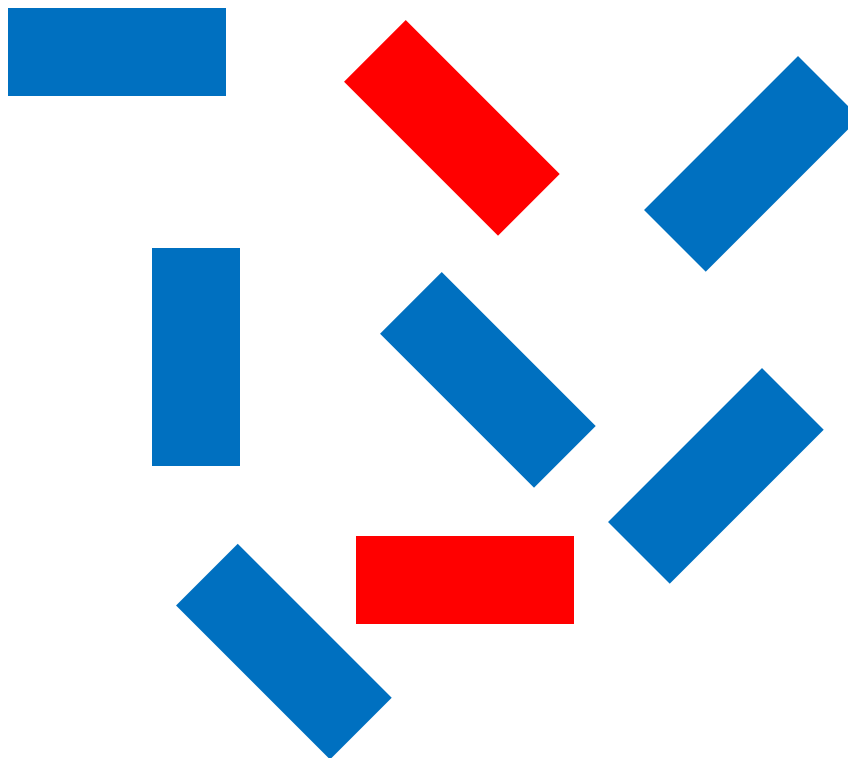
## Display 2

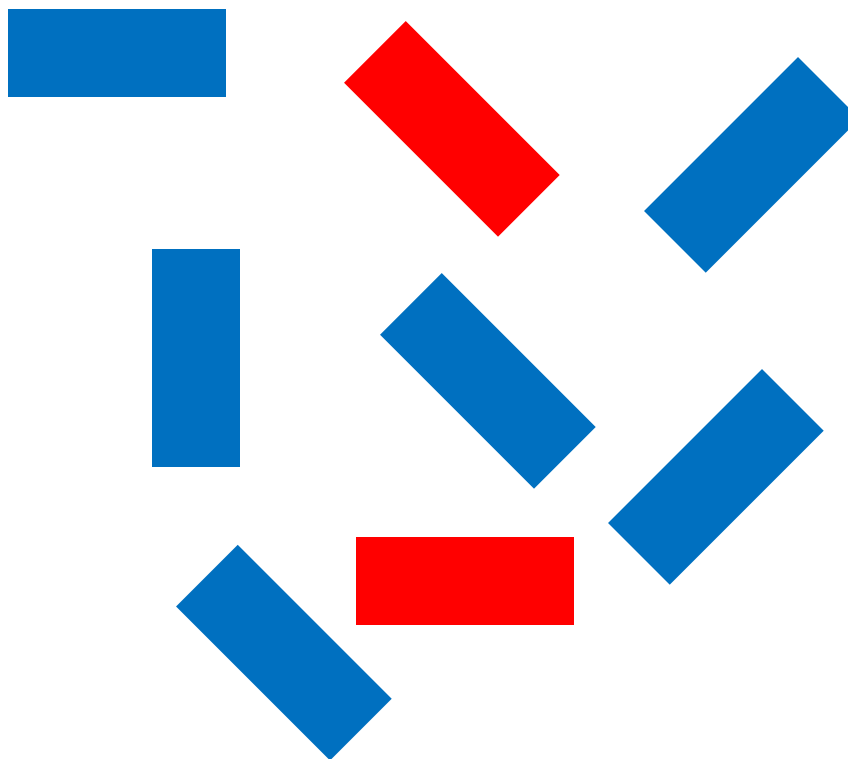
At most, one red rectangle will be rotated



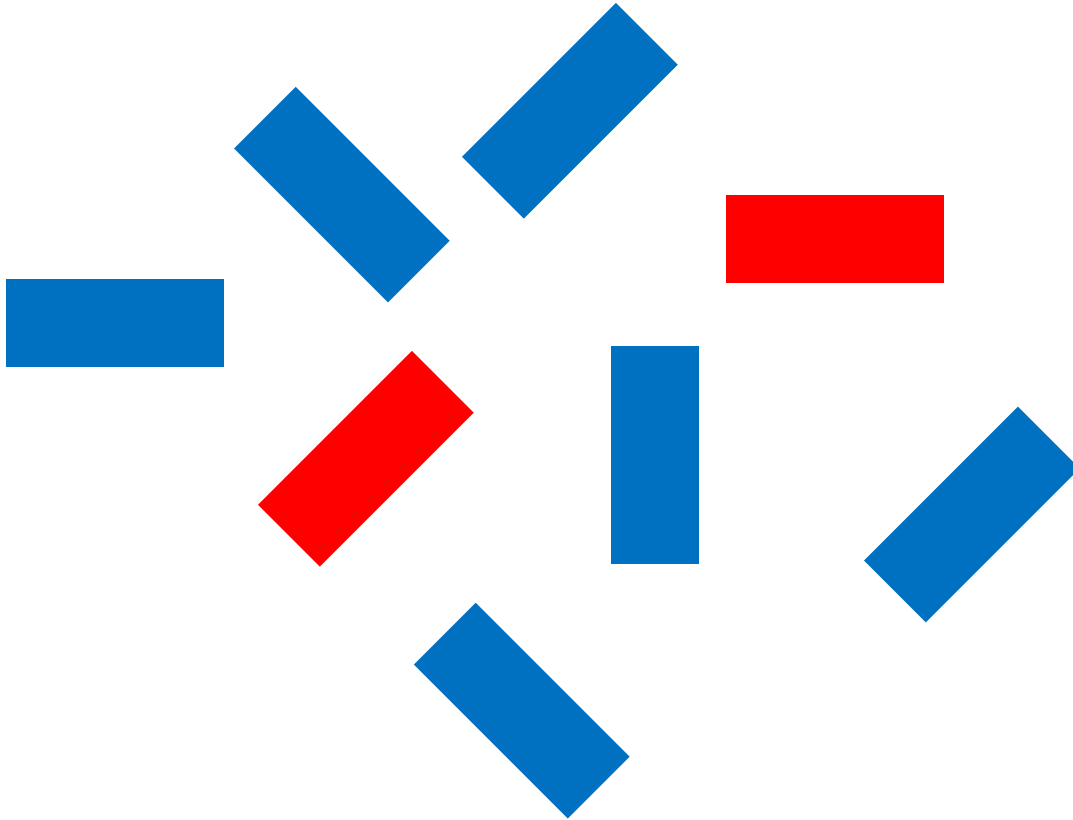
What do you predict about how LMM vs HMM will perform?

# Ready?

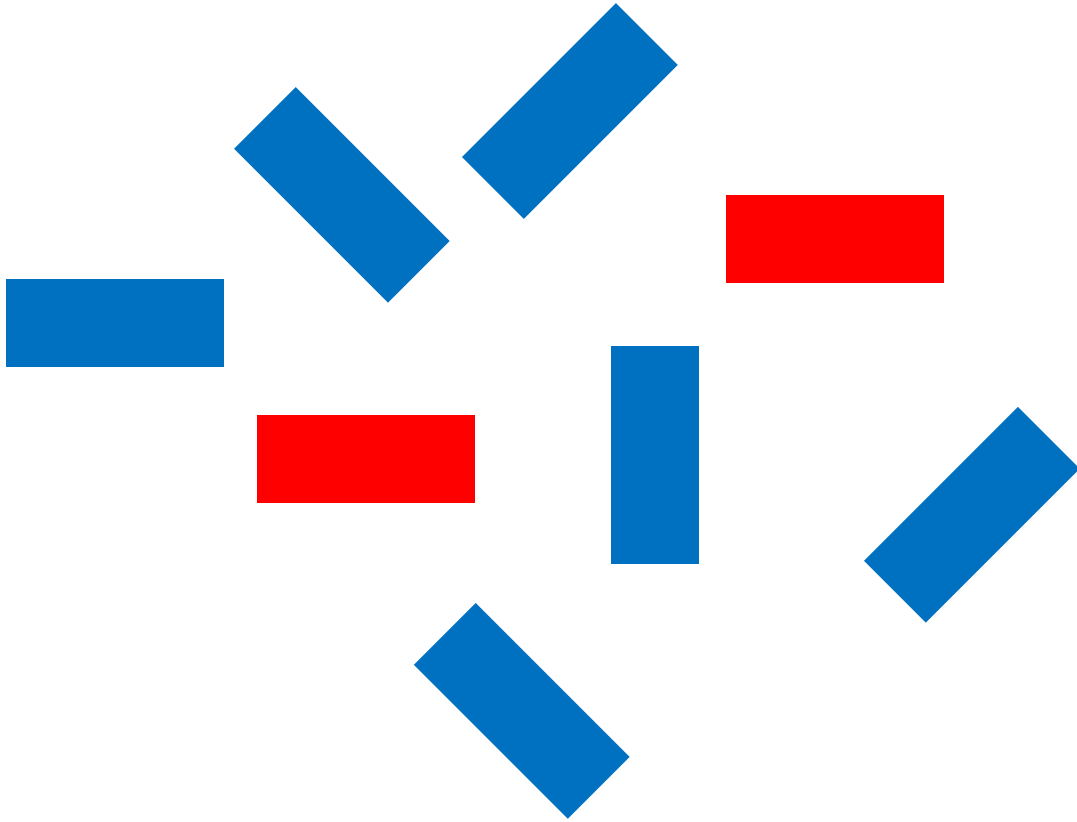






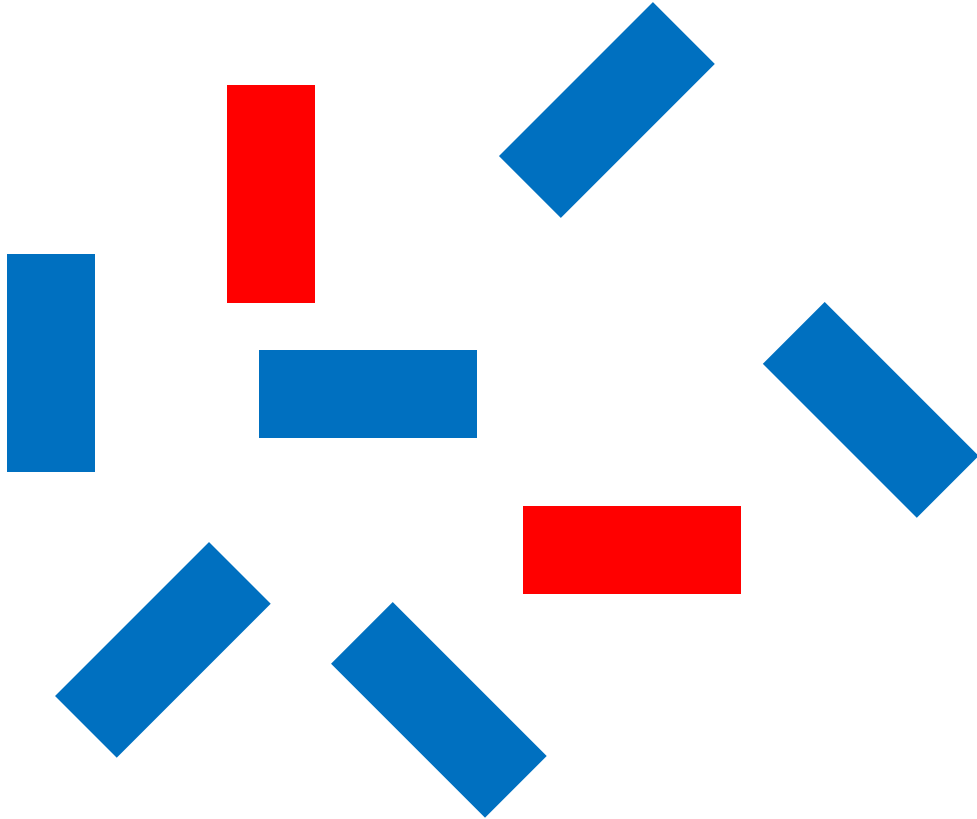


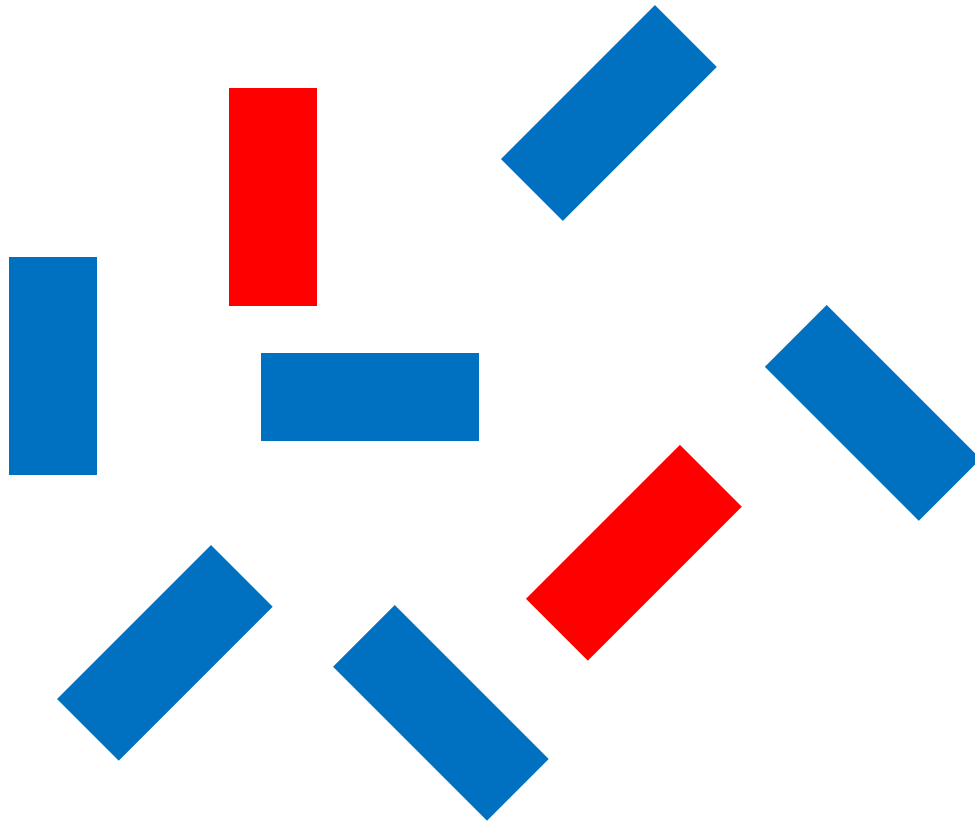




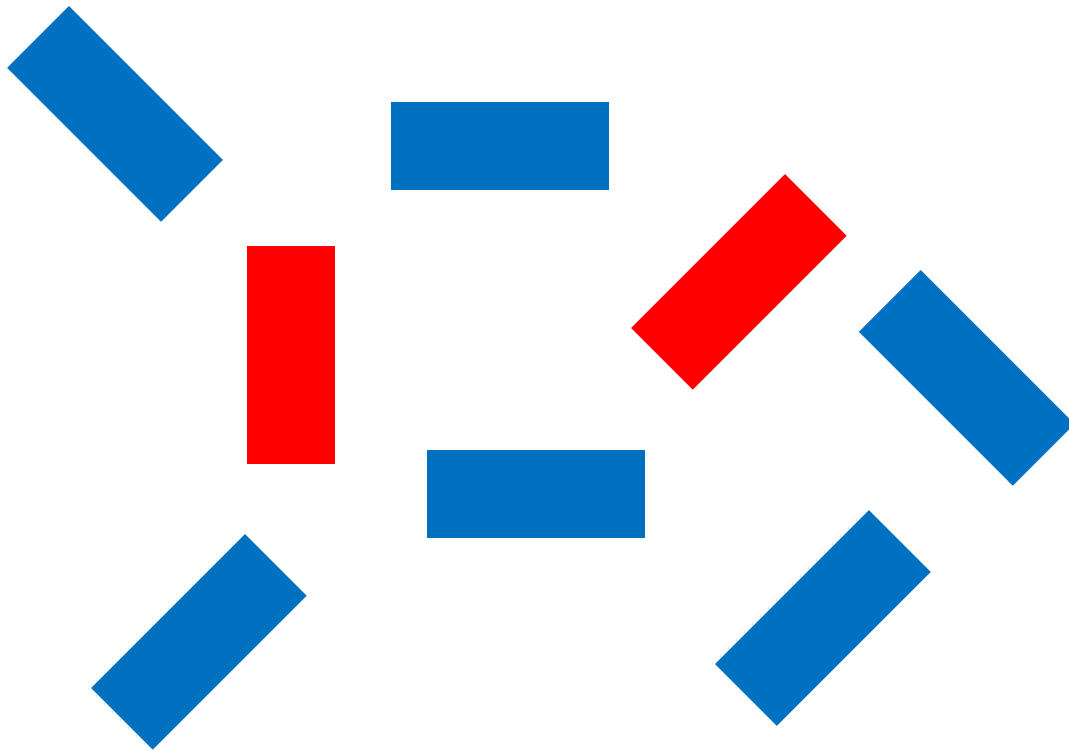
# Change: Yes or no?

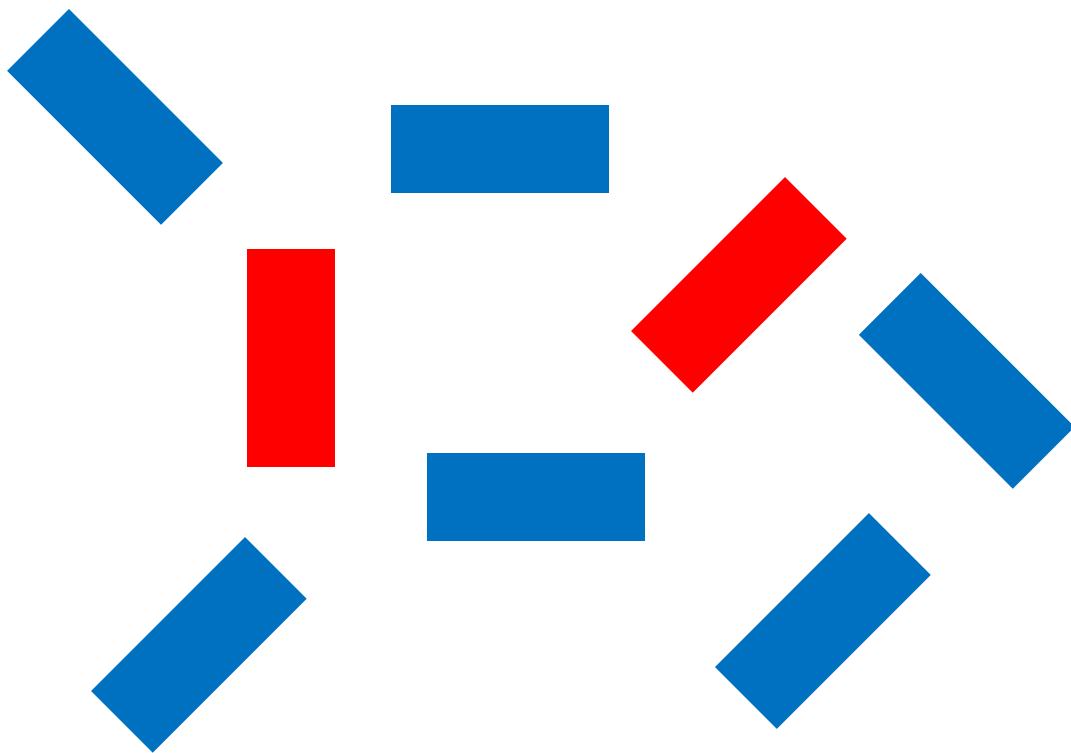




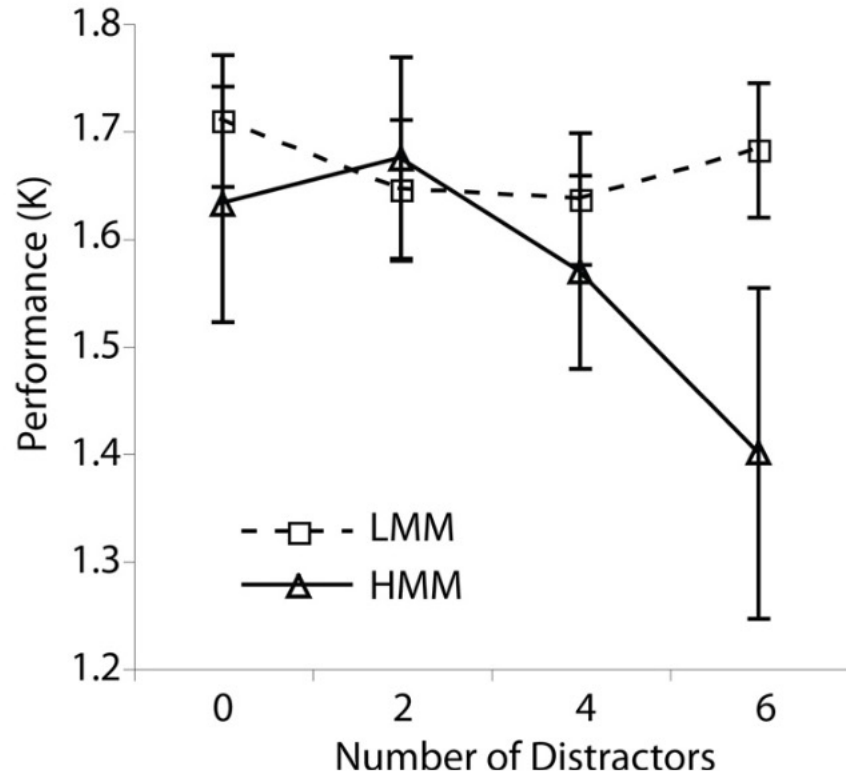








# Multitasking Study





# Multitasking Study

What method did they use?

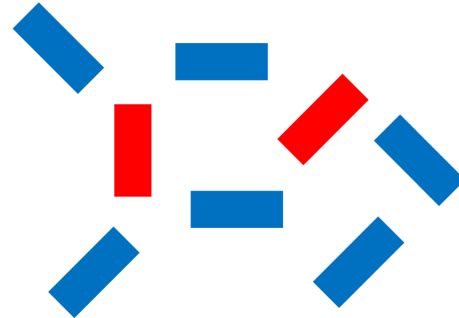
- Observational
- Correlational
- Experimental
- Quasi-experimental

Independent variables(s)?

- Multitasking: HMM or LMM
- # of distractors (blue rectangles)
  - 0, 2, 4, 6

Dependent variable(s)?

- Detecting changes to red rectangles



[Ophir, Nass & Wagner \(2009\)](#)

# Multitasking Study

What can we conclude from this quasi-experimental study?

# Multitasking Study

What can we conclude from this quasi-experimental study?

How could we make it an experimental study?

What's the difference between correlational and quasi-experimental?

# New Topic



# Scientific Method

Based on empiricism

- Testable predictions
- Observation

Hypothesize

Operationalize

Measure

**Evaluate**

Revise or Replicate



# Evaluating Research

How valid is this study?

- Construct validity
- Ecological validity
- Internal validity
- External validity

# Construct Validity

Are you testing and measuring what you intent to?

- How good are your operational definitions?
- Does blue-red rectangle task test “filtering of distractions”?
- Does the questionnaire accurately identify HMM vs. LMM?

You have full control over this.

# Ecological Validity

Does the research setting resemble everyday situations?

Effect of cell phones on driving performance

Mouse &  
keyboard  
(low  
ecological  
validity)



Wheel &  
pedals  
(higher  
ecological  
validity)





# Internal Validity

Could **only** the IV have caused changes in the DV?

- True experiment → high internal validity

Did multitasking cause changes in filtering distractions?

# External Validity

How well do results generalize to **other situations** and **other people**?

Does the red-blue rectangle task generalize to driving while texting?

Do Stanford students behave similarly to other university students? Older adults?

# New Topic



# Scientific Method

Based on empiricism

- Testable predictions
- Observation

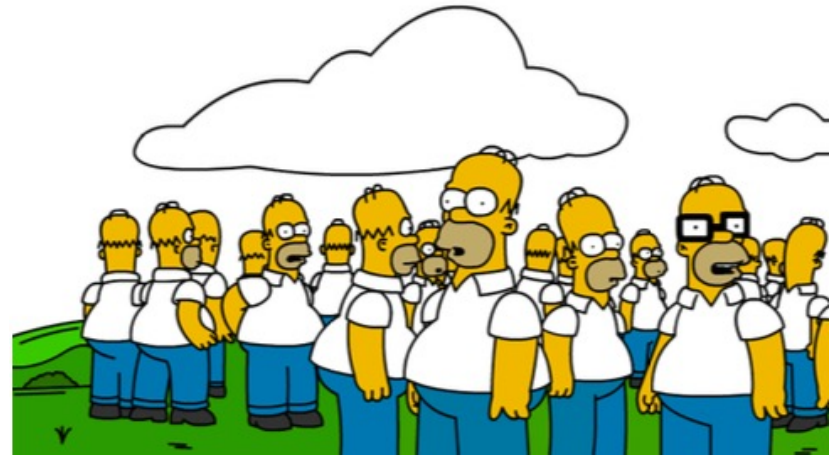
Hypothesize

Operationalize

Measure

Evaluate

Revise or Replicate



# Revise or Replicate

If findings do not support our hypotheses:

- Revise operational definitions

If findings do support hypotheses:

- Replicate findings, possibly including different conditions
- Ex: Other types of multitasking situations

# New Topic



# Research Design Activity

Design a study on smart phones and driving safety.

Hypothesize, Operationalize, Measure

Study can be: ~~Observational~~, correlational, experimental, or quasi-exp

Assume access to research tools

- Instrumented car, driving simulator, eye tracker, accident records...whatever you might need

Expectations, conclusions, limitations?

Write this up in a blog post. Include your research question, the variables, your design, how you'll measure them, your hypothesis, and what you'll do.

