Lecture 20: Inferential Statistics

Wednesday, November 8, 2023

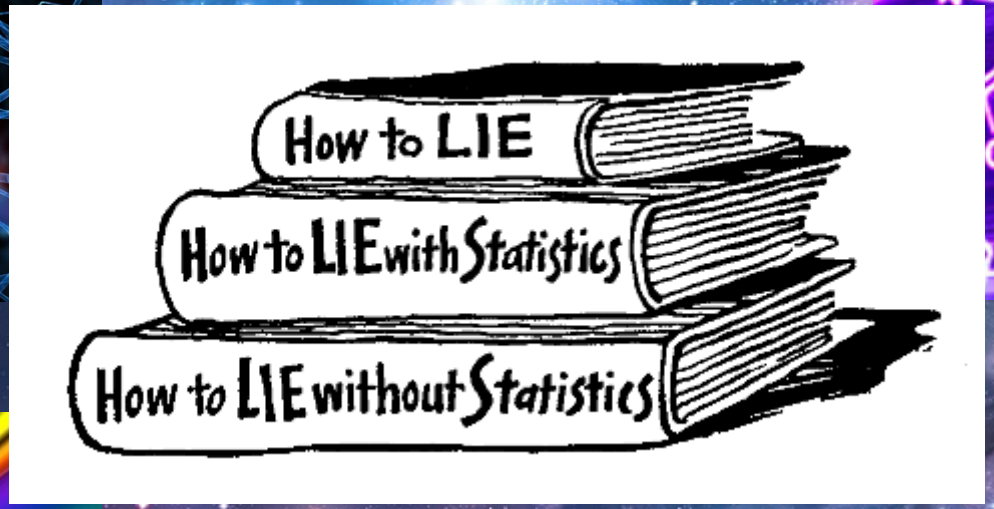
Your Teaching Fellows:

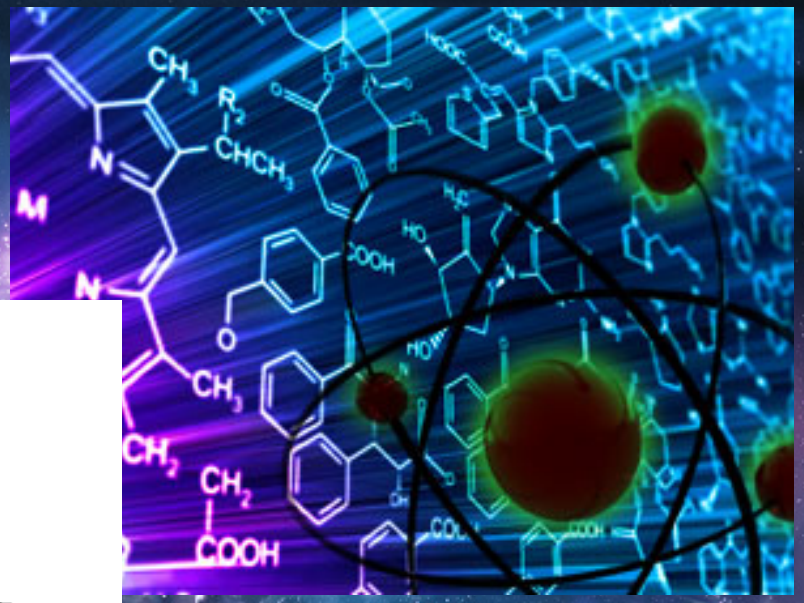
003/004:	Zahra Abolghasem	Bronwen Grocott
	Vasileia Karasavva	Ni An
010:	Thalia Lang	Malina Lemmons
	Ruoning Li	Irene Wen

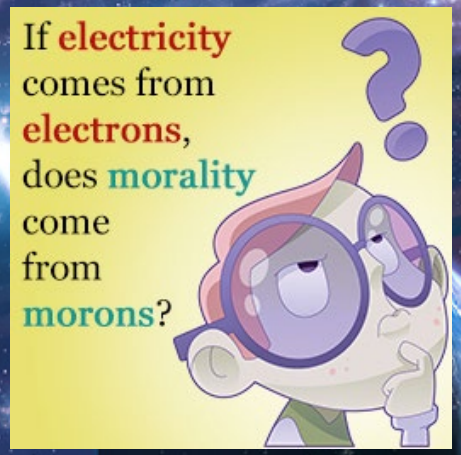
Lectures: MWF 12:00 PM – 1:00 PM (003); 1:00 PM – 2:00 PM (004); 2:00 PM – 3:00 PM (010)
Office hours: Tuesdays 2:00 PM – 4:00 PM

Learning objectives

- By the end of this class, you'll be able to
  - Differentiate between population, probabilistic trend, and sample
  - Understand the relation between sample size and accuracy in estimation population values
  - Explain the common errors in judgment that people make


How to LIE
How to LIE with Statistics
How to LIE without Statistics




If electricity comes from electrons, does morality come from morons?

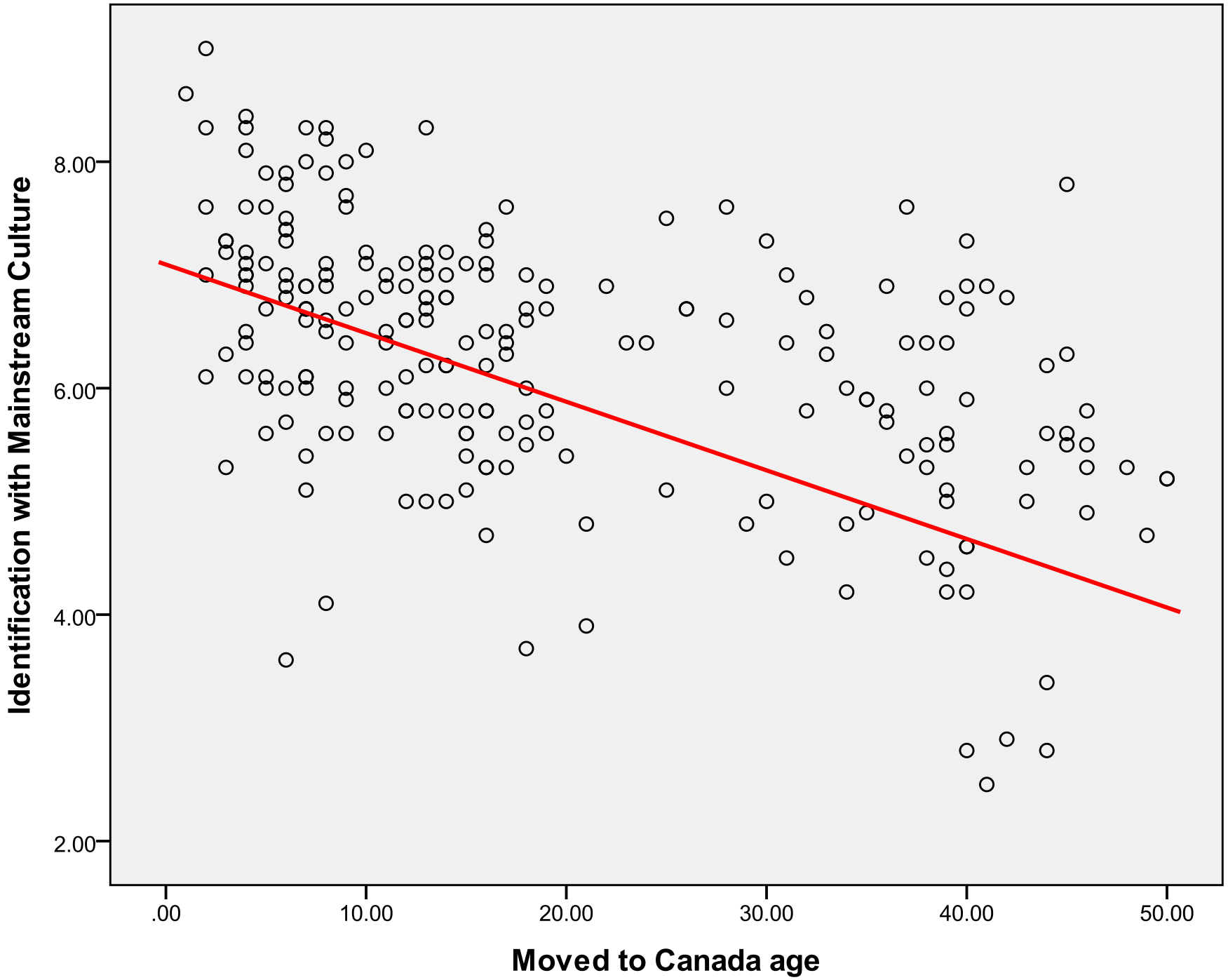
Sample vs. Population

- Population Level
  - As a general state of affairs in this universe, how effective is the Pfizer vaccine for COVID?
- Probabilistic trend
  - The expected effectiveness at the population level – often applied to a study
- Random Sample
  - In any random group of people, how effective is the Pfizer vaccine for COVID?

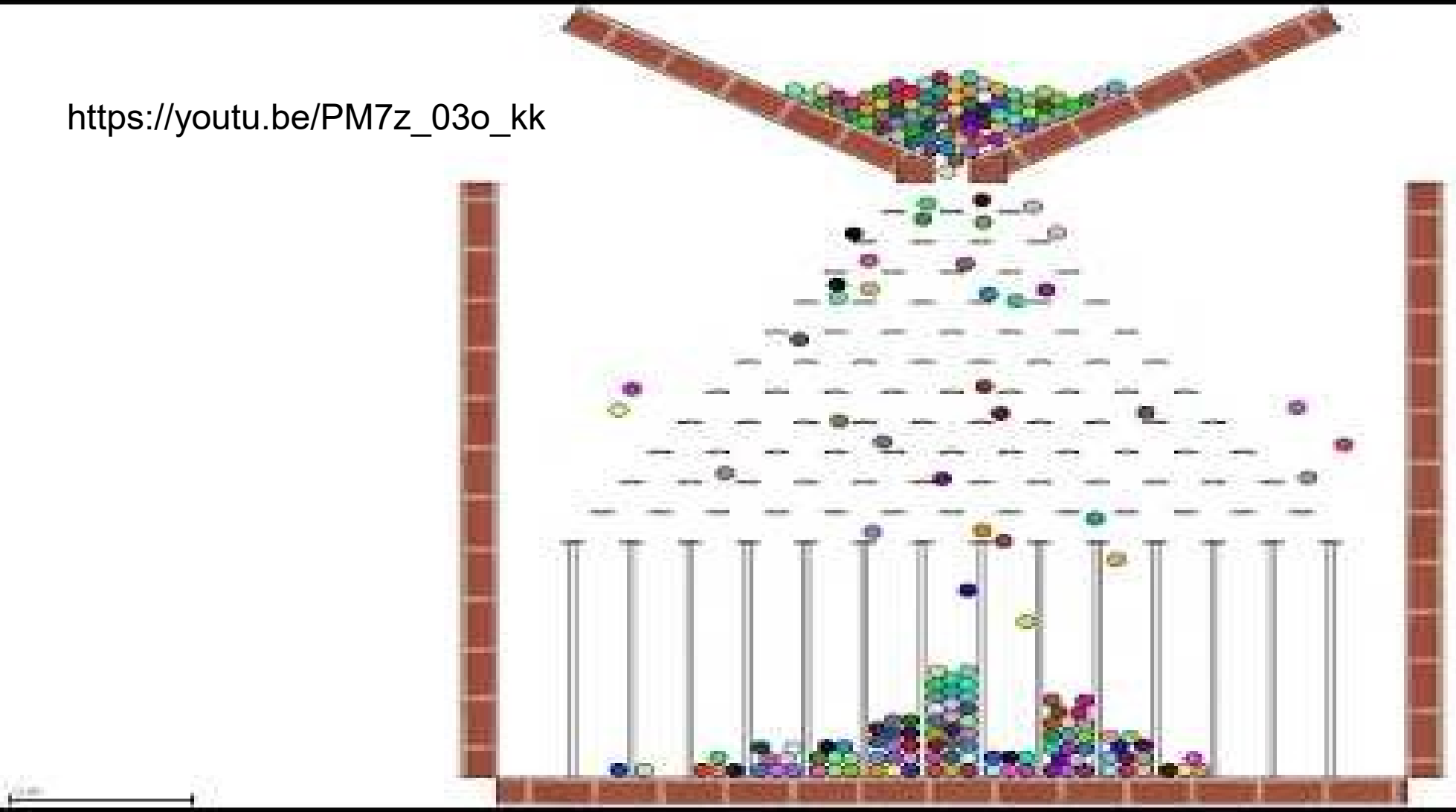
- As our sample size increases, the pattern in the sample will better represent what we should see at the population level
  - Trying to obtain estimate of what happens in the population

Sample vs. Population

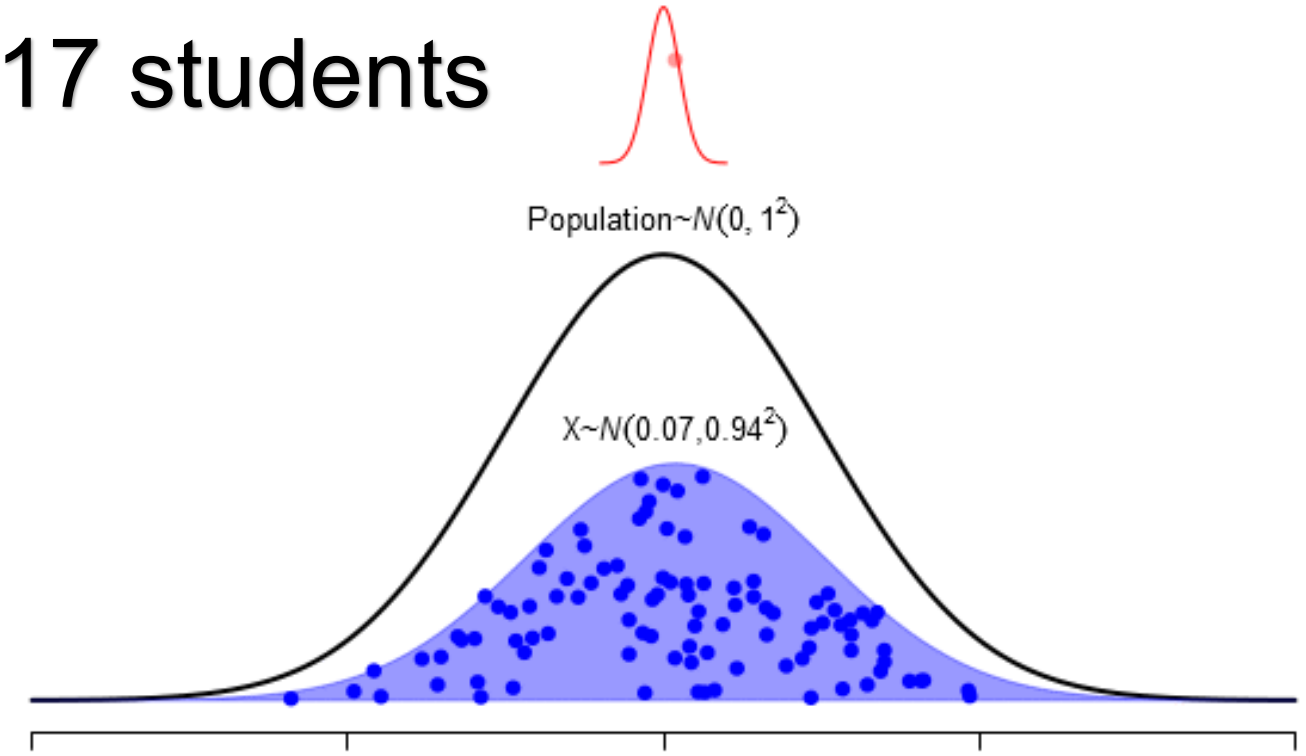
- Samples are imperfect assessments of overall probabilities
- Probabilistic trend
  - What should we expect to see based on the population?
  - Not likely to be reflected in every sample and case


Identification with Mainstream Culture
8.00
6.00
4.00
2.00
.00
10.00
20.00
30.00
40.00
50.00
Moved to Canada age

https://youtu.be/PM7z_03o_kk



All PSYC 217 students



All PSYC 217-003/004/010 students

Sample vs. Population

- Samples are imperfect assessments of overall probabilities
- Probabilistic trend
  - What should we expect to see based on the population?
  - Not likely to be reflected in every sample and case
  - E.g., regression line as the best fit to summarize all the data points

- Random Sample
  - Taking random sample from population to estimate true effect
  - As sample size decreases, estimate is less accurate

Sample vs. Population

Game: 10 players, 1 imposter

Theoretical chance of being imposter: 10%

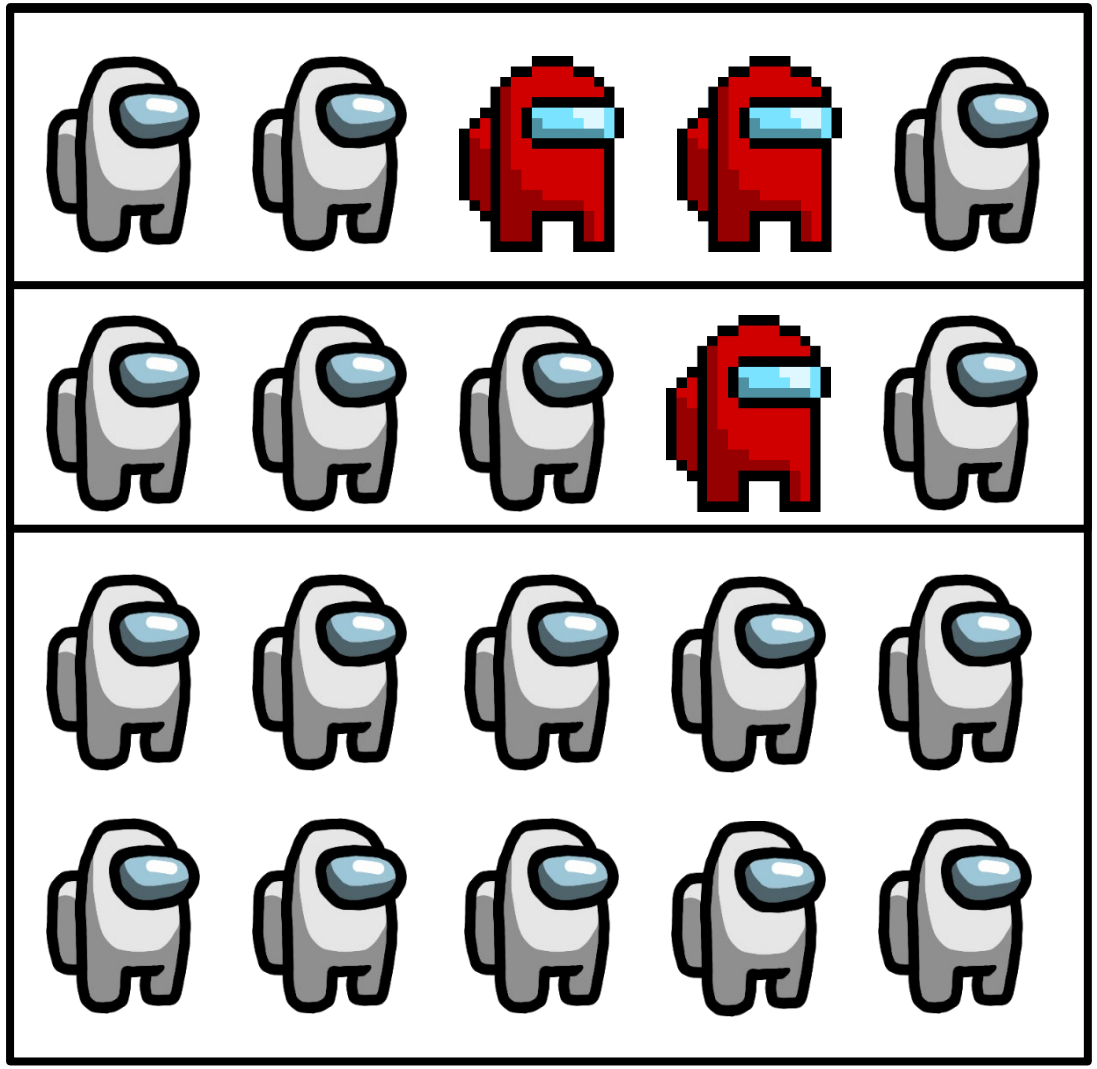
= Imposter

= Crewmate

Imposter chance: 30%

Imposter chance: 15%

Imposter chance: 40%



Infinite games: 10%

Common error in judgment 1: Failing to account for sample size

You and your friend each flip a coin for 10 days

You flip your coin 100 times a day, your friend flips their coin 20 times a day

Generally, about 50% of flips are heads; but the exact % varies per day. Sometimes it is higher than 50%, sometimes lower.

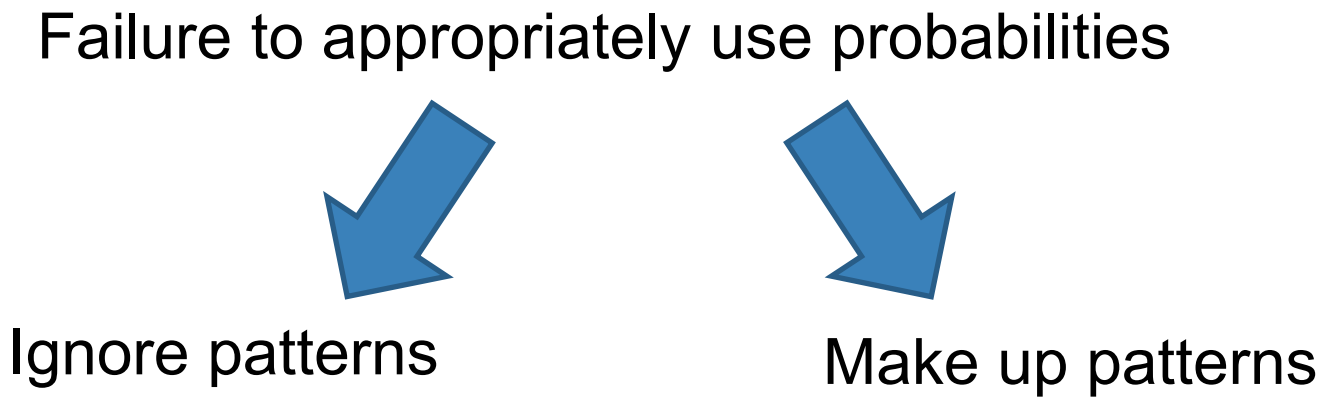
At the end of your respective set of flips, who is likelier to have more had more days with 60% or more heads?

- A) You B) Your friend C) about the same

Inferential Statistics

- Use data collected on a *sample* to infer what is happening in the *population*
- Is the effect we found in our sample due to random chance, or due to a true effect in the population?

Common error 2: Failure to appropriately use probabilities



- Focus on specific rather than general information
  - E.g., “Person-who” statistics (aka Base rate fallacy)

I know this person who…

Yeah but what about that person…

Person-Who statistic

Pancreatic cancer by 150%

Breast cancer by 50%

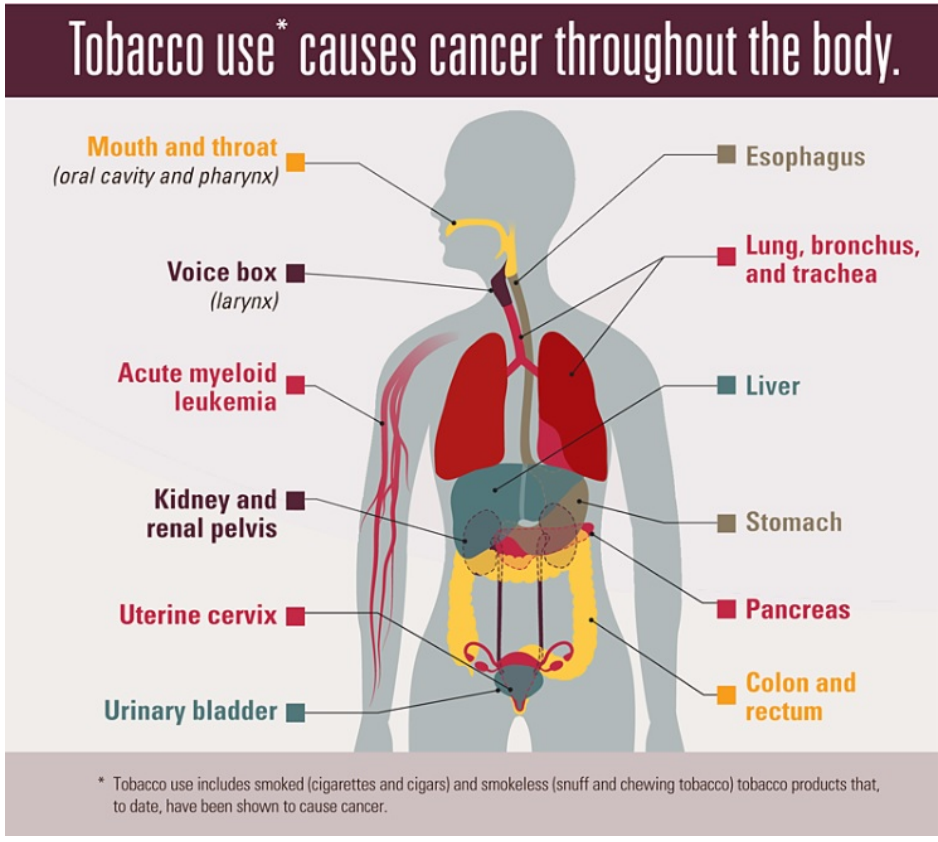
Gastric cancer by 100%

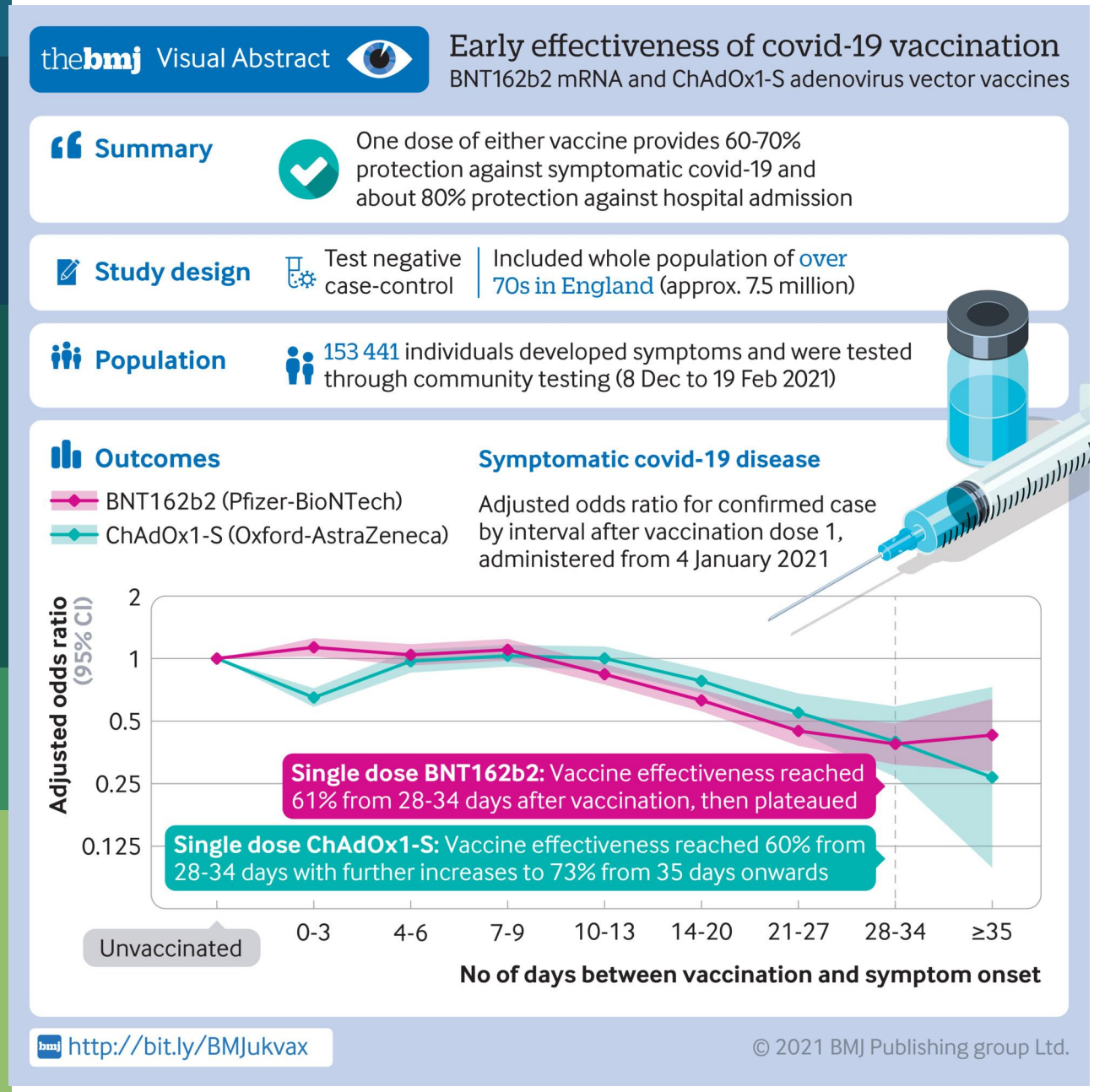
Leukemia by 40%

Lung cancer by 330%

Colorectal cancer by 70%

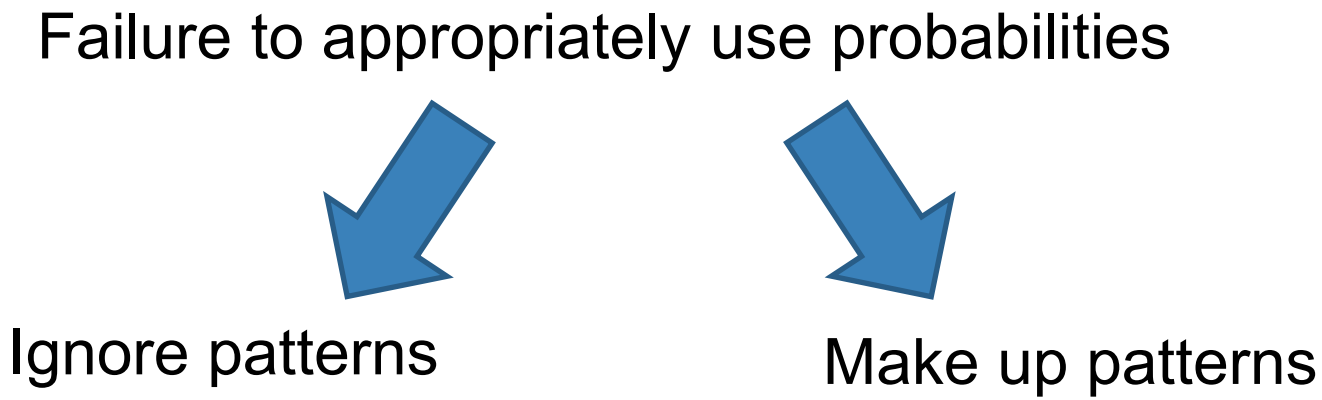
Bladder cancer by 190%



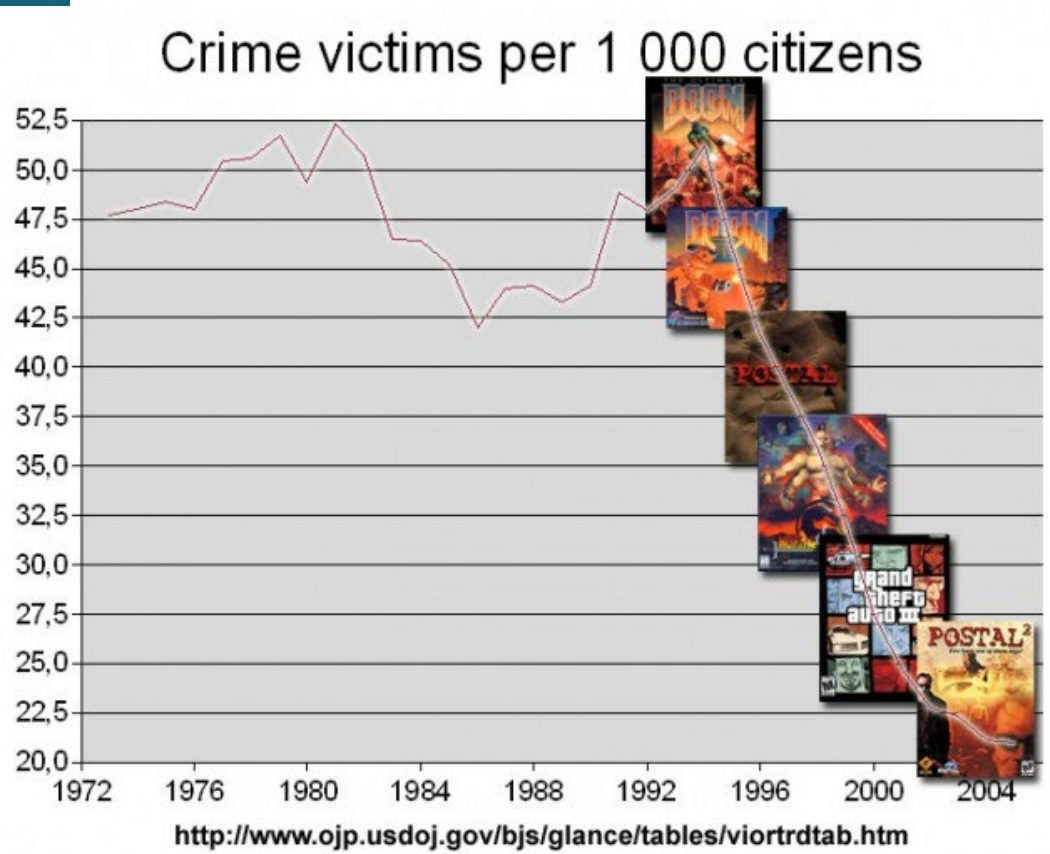


"If the vaccine is so great, then how come people are still getting COVID and spreading COVID and, unfortunately, dying from COVID?" – Aaron Rodgers

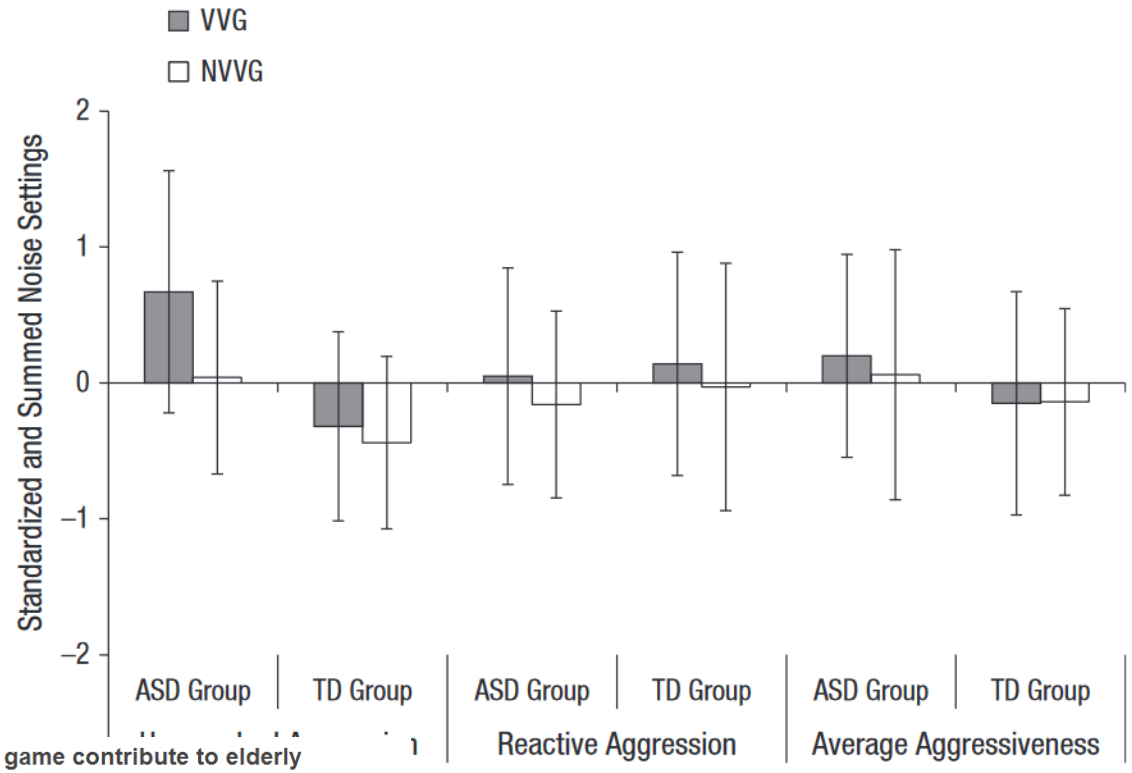
Common error 2: Failure to appropriately use probabilities



- Seeing patterns in randomness to predict future outcomes



http://www.ojp.usdoj.gov/bjs/glance/tables/viortrdtab.htm



Did popular video game contribute to elderly woman's murder?

Email 2 Share 6 Tweet 0 ShareThis 81

View Larger

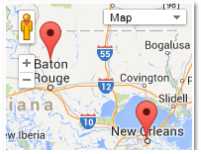
By CNN contributed to this report

Slaughter, LA — An 8-year-old Louisiana boy allegedly shot and killed his grandmother after playing a popular, and violent, video game investigators say.

87-year-old Marie Smothers was killed with a single gunshot wound to the head last week in Slaughter, Louisiana, which is about 20 miles north of Baton Rouge.

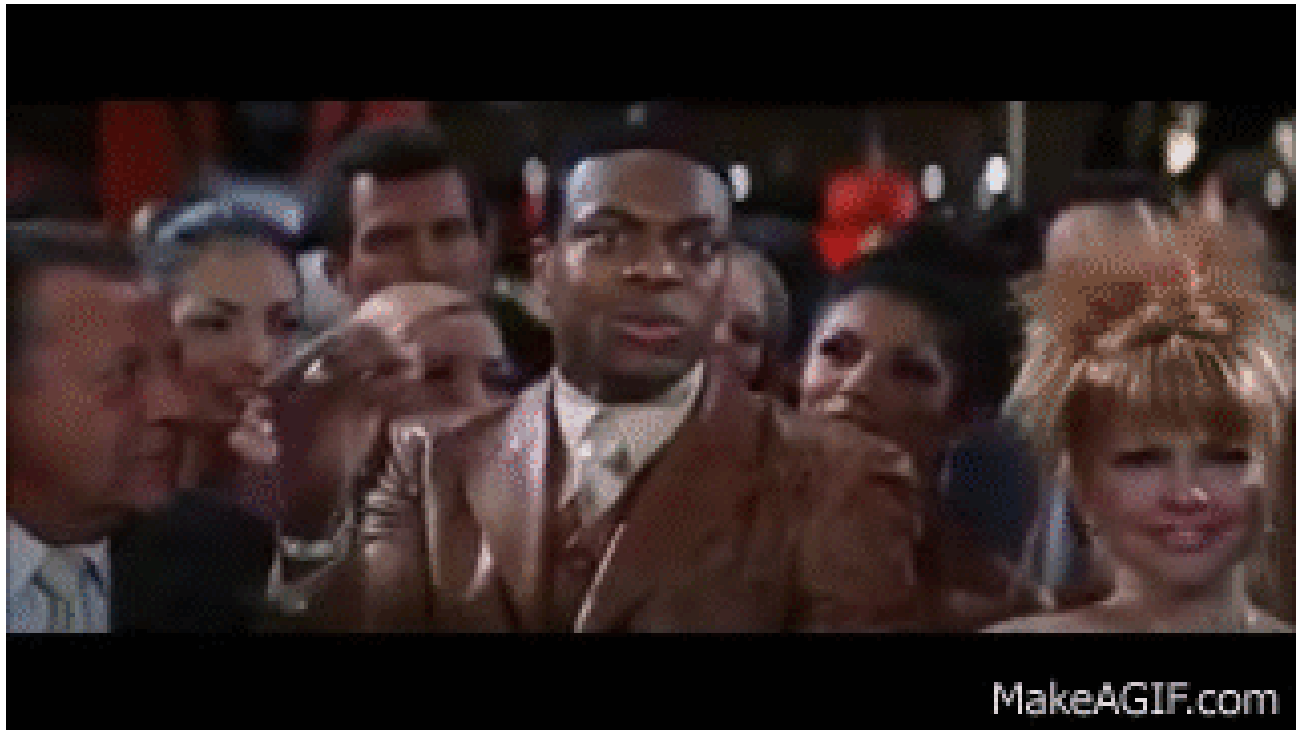
Related

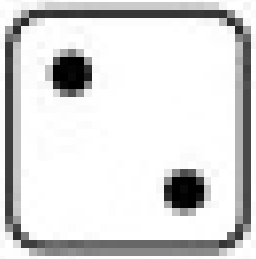
View Larger

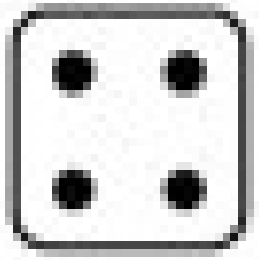


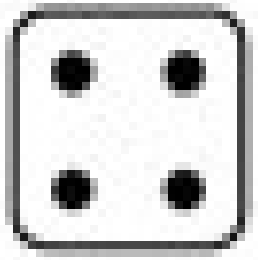
While a motive for the killing is still unknown, investigators have learned that the young boy was playing 'Grand Theft Auto IV' directly before the shooting occurred.

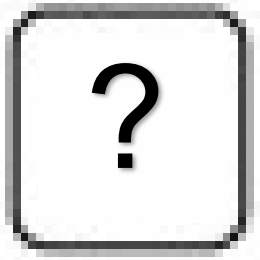
A statement sent to CNN by the video game company said this is about access to guns not video games, and "...Ascribing a connection to entertainment, a theory that has been disproven repeatedly by multiple independent studies, both minimizes this moment and sidesteps the real issues at hand."


MakeAGIF.com

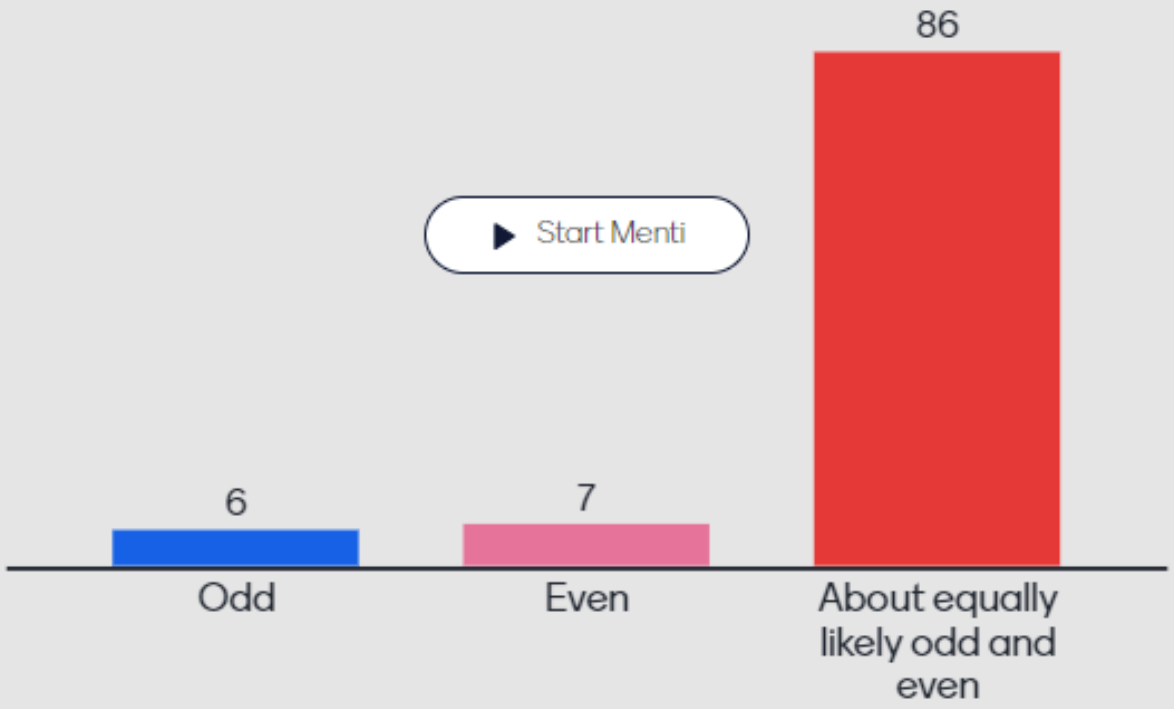


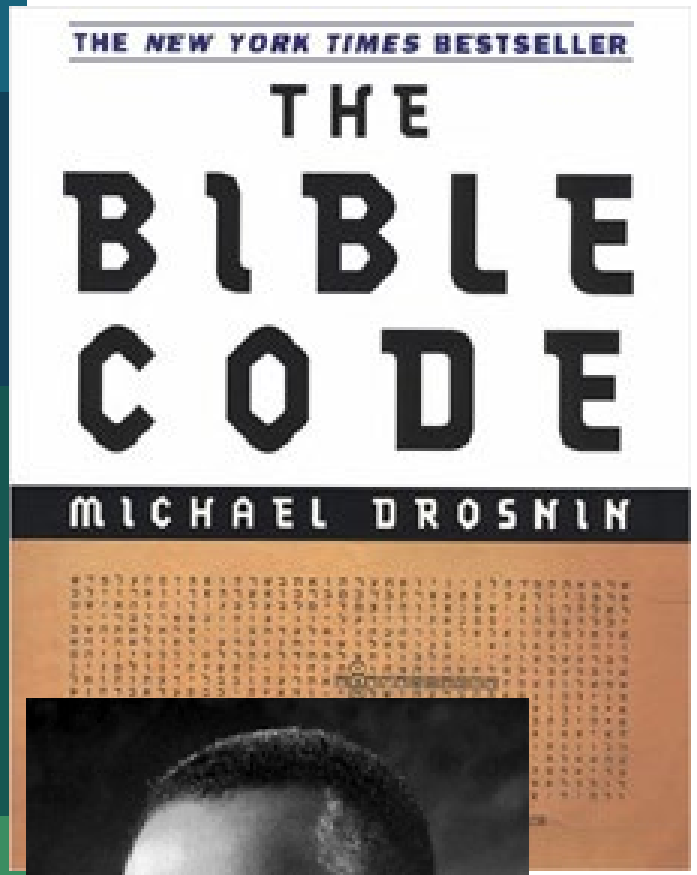


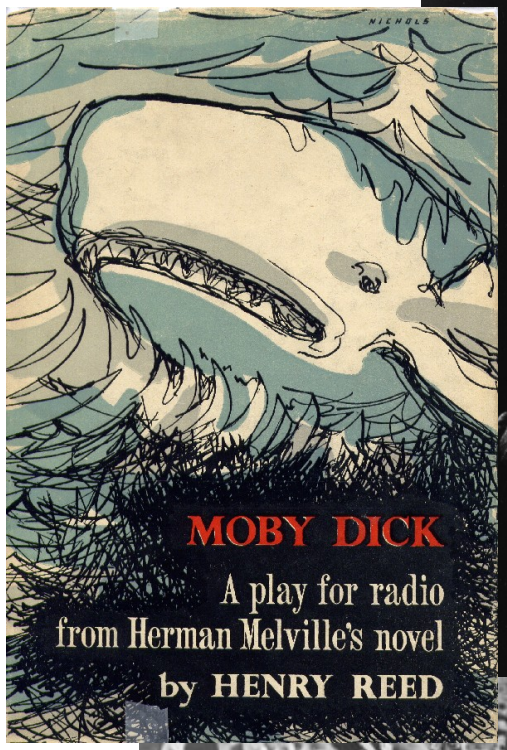


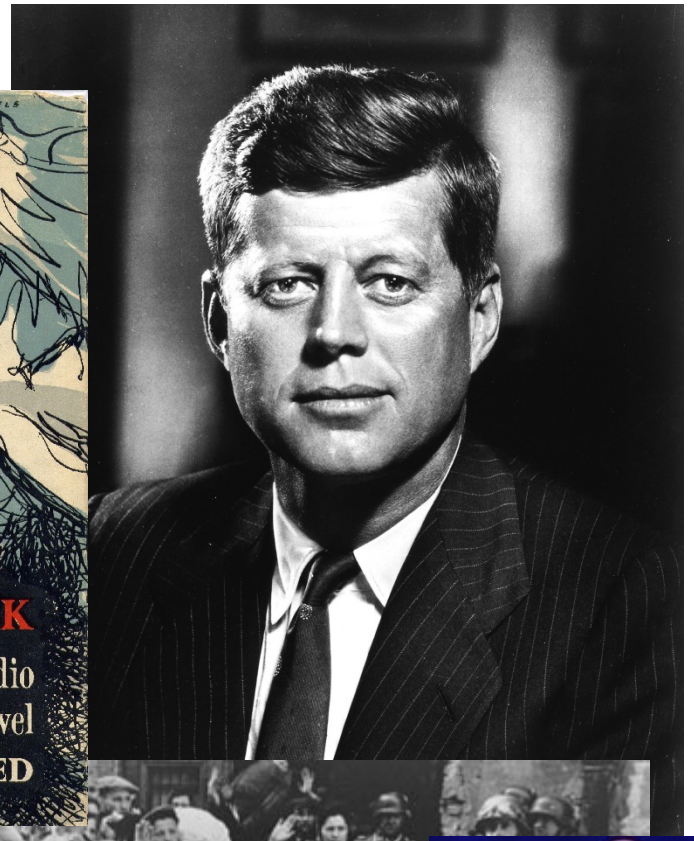

?

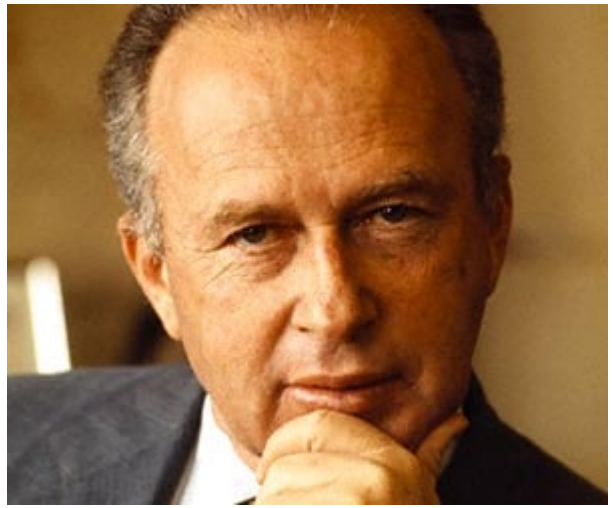
Is the next die roll likely to be an odd number or an even number?

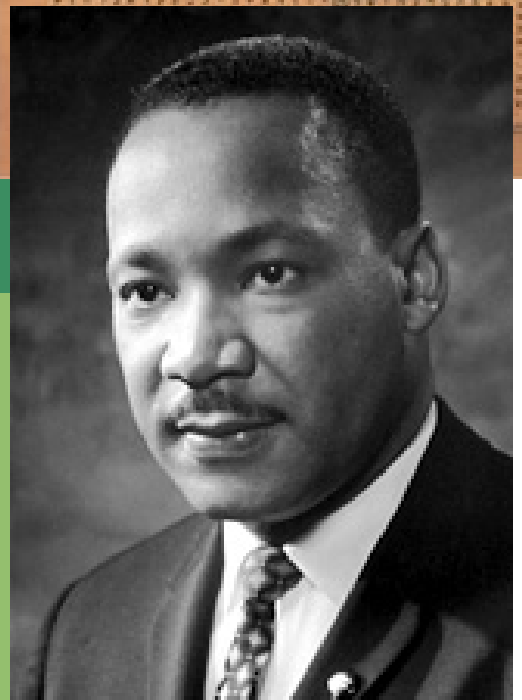



THE NEW YORK TIMES BESTSELLER
THE
BIBLE
CODE
MICHAEL DROSNIN


NICHOLS
MOBY DICK
A play for radio
from Herman Melville's novel
by HENRY REED








GETTY IMAGES

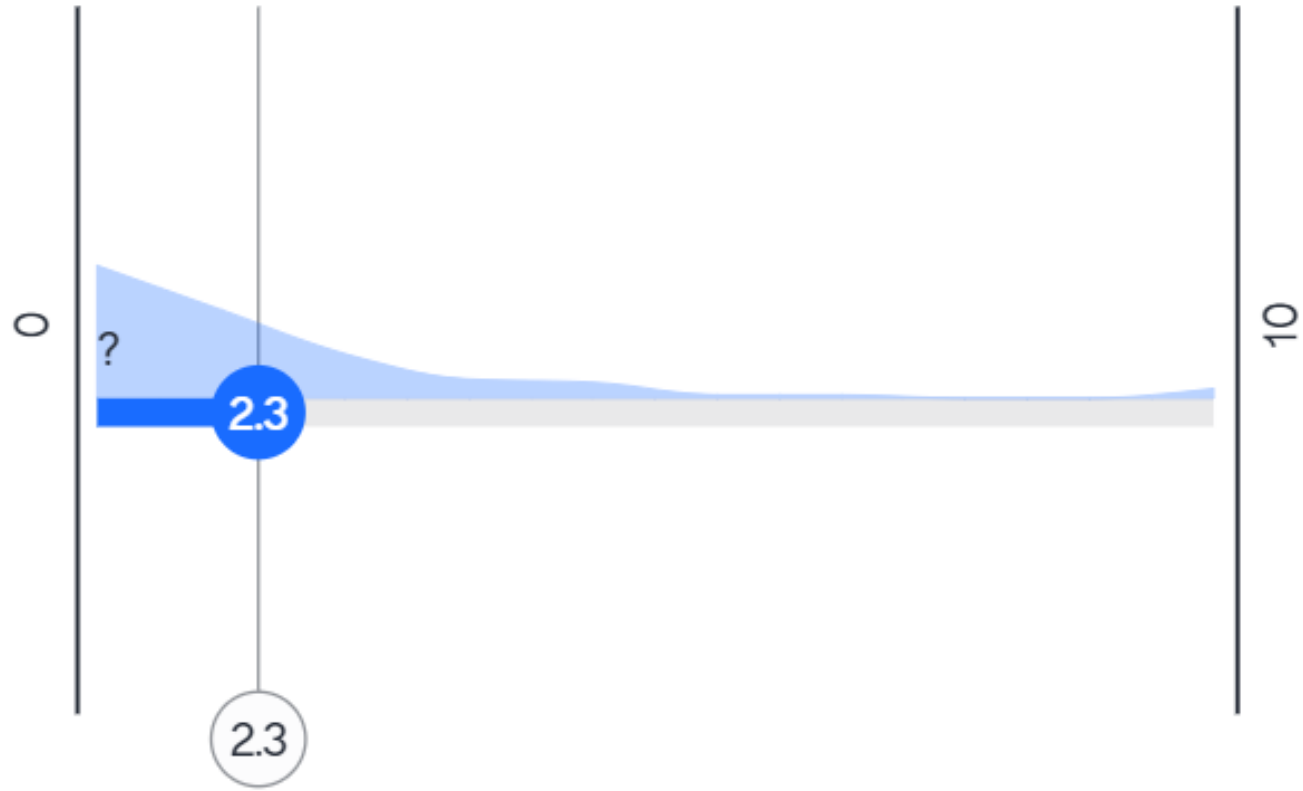
Inferential Statistics Overview

- Null & Research Hypotheses
- Sampling distribution
- *t*-test logic
- Statistically significant
- Type 1 and Type 2 errors
- Apply your understanding

Inferential statistics

- Use data collected on a *sample* to infer what is happening in the *population*
- Is the effect we found in our sample due to random chance, or due to a true effect in the population?

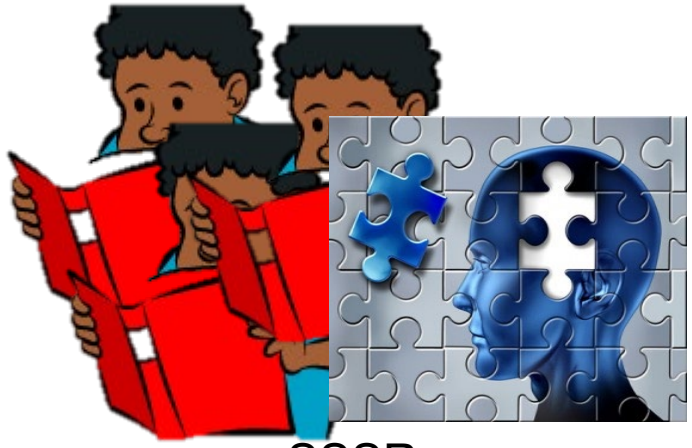
How many times do you read the textbook when you study?



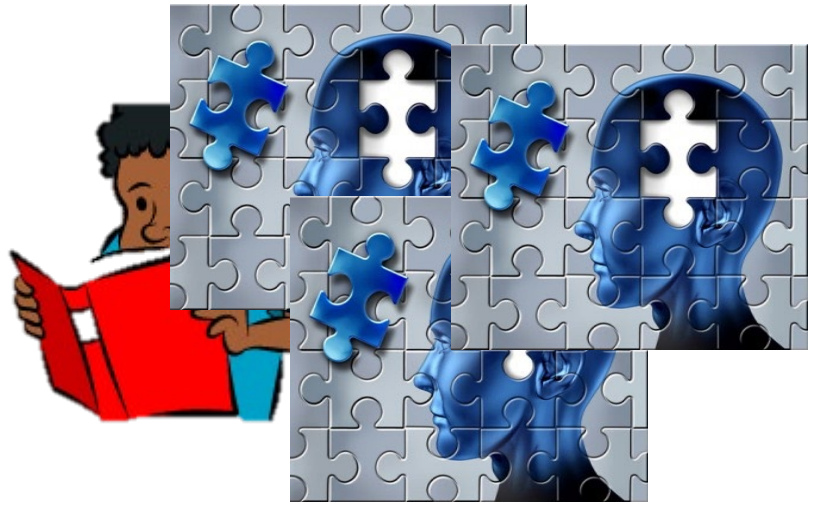
How to study?

- How does repeated reading of study materials affect learning?
  - IV: Study method

SSSS



SSSR

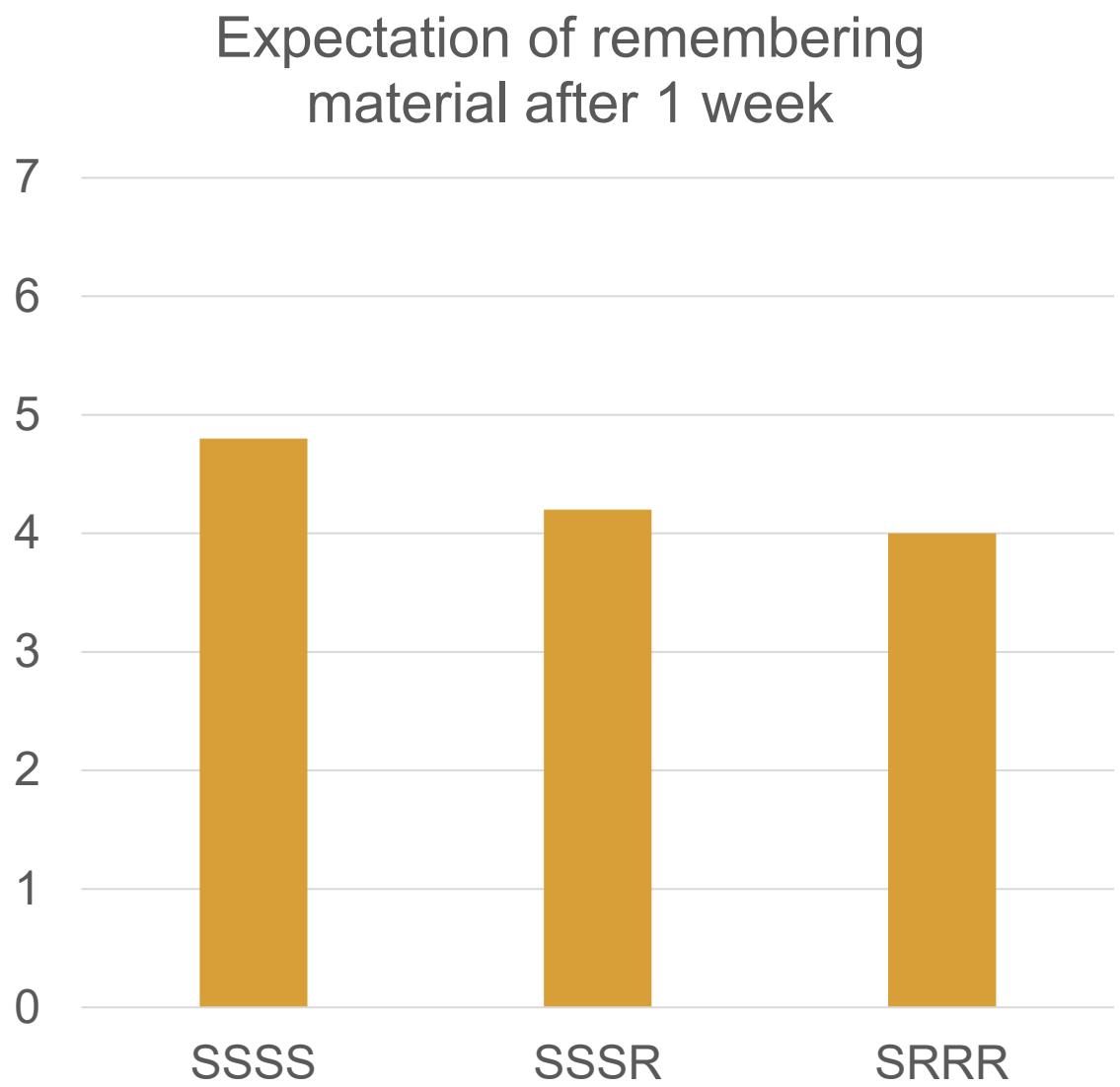


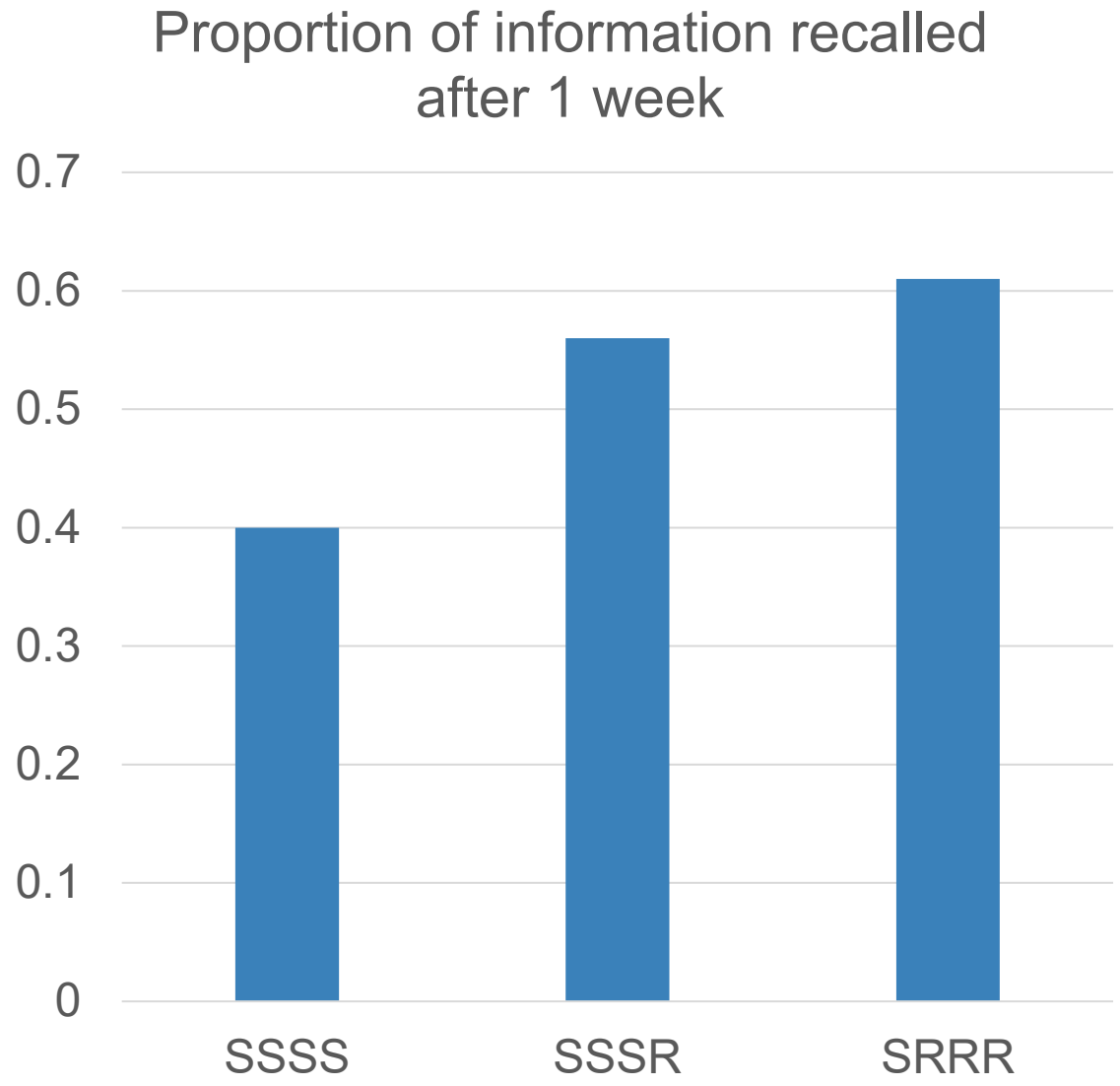
SRRR

  - DVs:
    - Predicted recall
    - Actual recall

Roediger & Karpicke (2006)

How to study?


Expectation of remembering material after 1 week
7
6
5
4
3
2
1
0
SSSS
SSSR
SRRR


Proportion of information recalled after 1 week
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
SSSS
SSSR
SRRR