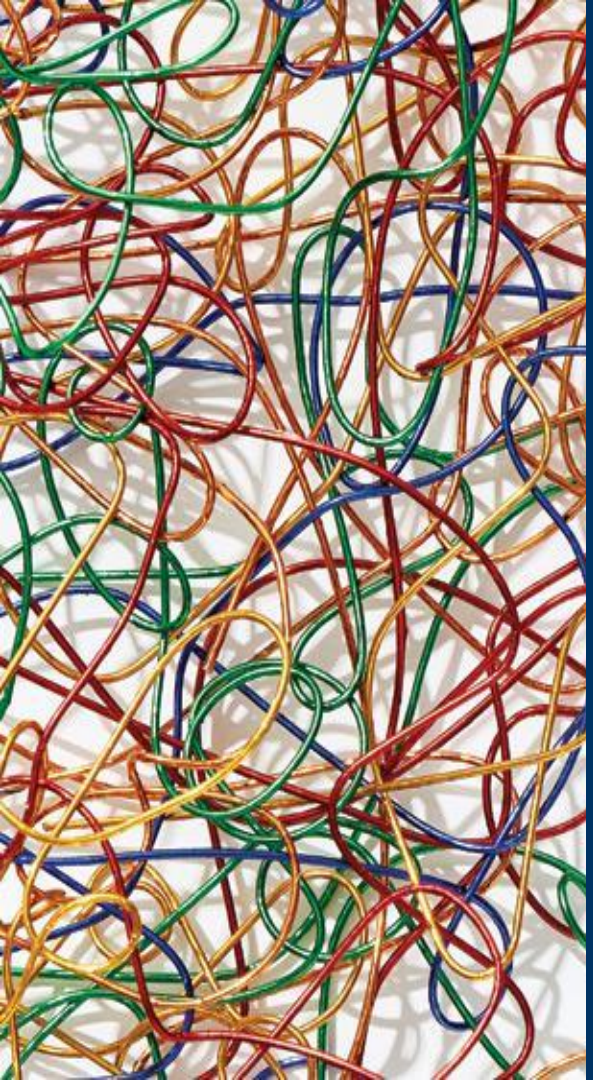
“

We know you more
than you know yourself

A population of
unsuspecting statistics

"Data_m1nefield"
Cyborg Octopus



Speaking Science

Effect of Locality in Reader Engagement

Aris Fotkatzikis, Mumin Khan, Haerang Lee

W241, Spring 2020, Final Project

Github

> Making the simple complicated is commonplace; making the complicated simple, awesomely simple, that's creativity.
>
> *Charles Mingus, Jazz Musician*


Communicating science *matters*.
Science Distrust
Misinformation
Poor Decisions
Threat to Society

Does the **locality** of a scientific article impact the reader's **engagement** with the issue discussed in the article?

Design

"engagement"?

Potential Operational Variables

Pursue career (college major/job $\in$ science)?

Museum ticket purchases?

Created a vinegar volcano?

Eco-friendly product consumption?

Jargon in Tweets?

SCRAP! 🛑

clicking ≠ engagement
Put a cat on it, people click it

Self-Reported Interest?

SCRAP! 🛑

Abstract feelings unlikely to change in short term

Locality in News?

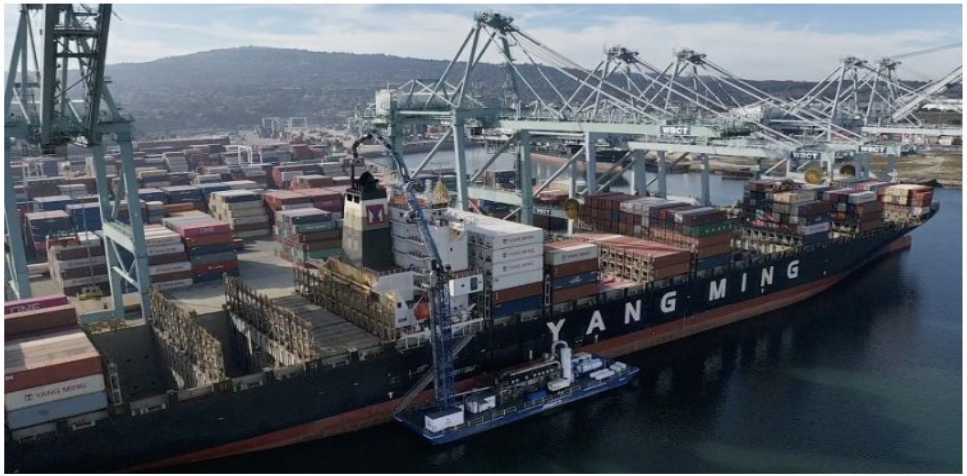
Yes, Please! ✔️

Measurable behaviors
Could vary in the short term

Los Angeles Times

CALIFORNIA

Port ships are becoming L.A.'s biggest polluters. Will California force a cleanup?



In December, a barge at the Port of Los Angeles uses a system, known as a bonnet or "sock on a stack," that's intended to scrub exhaust. (Allen J. Schaben / Los Angeles Times)

Plain text!

Treatment

Port ships are becoming **LA's** biggest polluters.
Will **California** force a cleanup?

Control

Port ships are becoming **NYC's** biggest polluters.
Will **New York** force a cleanup?

D

Does the locality of a scientific article

Y

impact the reader's engagement

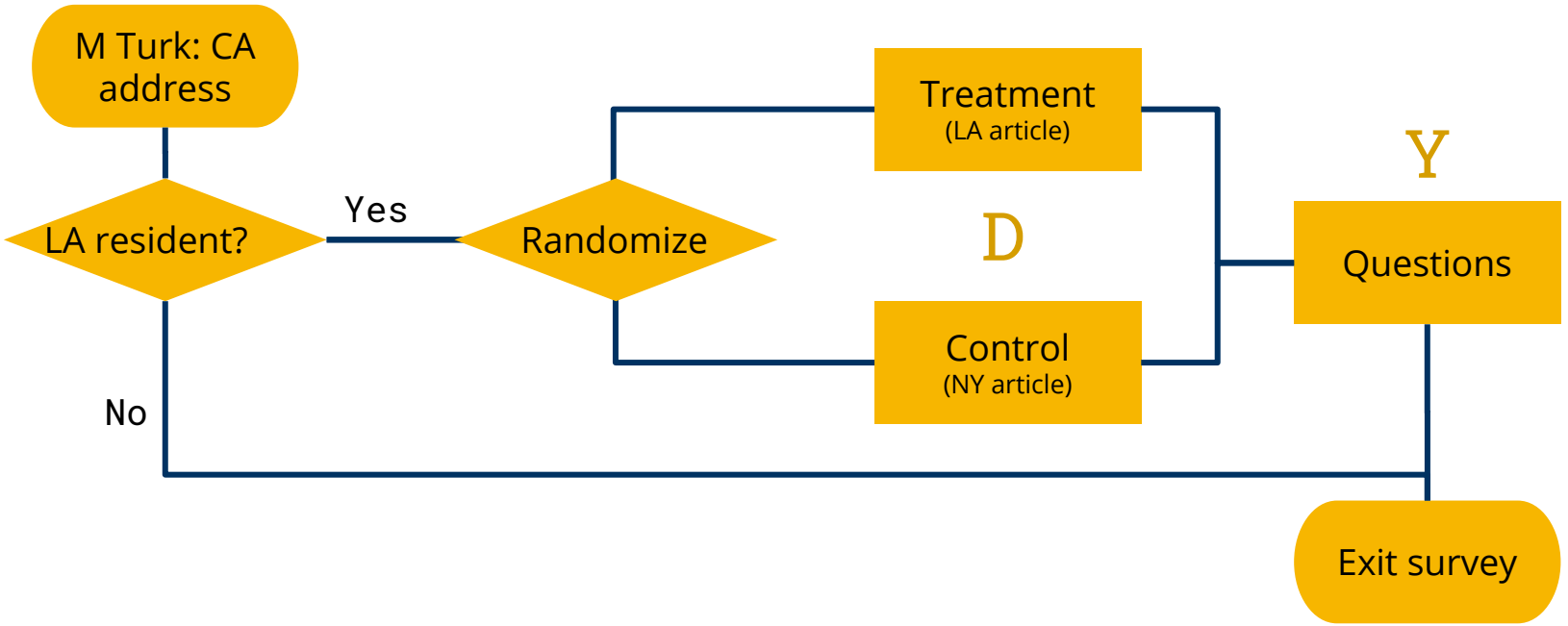
with the issue discussed in the article?

City vs. region?
City vs. remote city!

1. Reading comprehension **quiz score**

2. **Donation amount** (from the $100 raffle prize toward alleviating air pollution)

3. Article **reading time**


M Turk: CA address
LA resident?
Yes
No
Randomize
Treatment
(LA article)
Control
(NY article)
D
Y
Questions
Exit survey

Power Analysis: Quiz Score

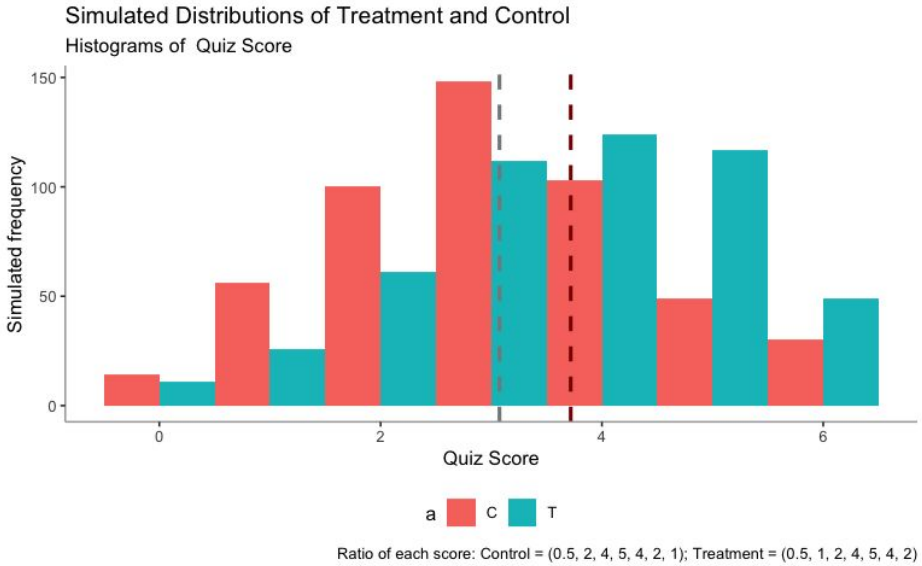
Assumed some reasonable distribution of scores 0-6

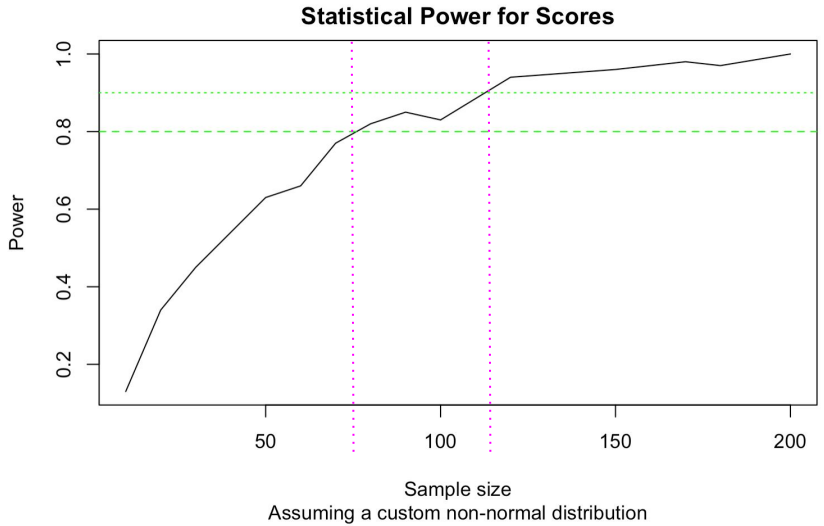
Treatment effect of 0.5 points

Sample Size Needed

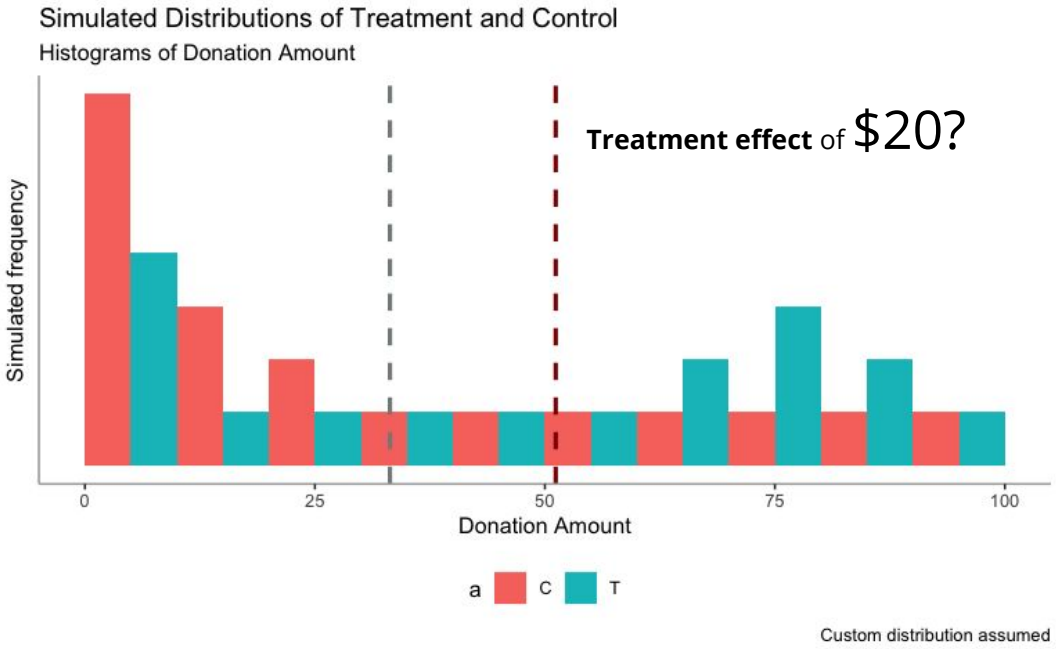
- 100 people → 80% power
- 120+ people → 90% power

(p = 0.05, using t-test and 100 simulations for each sample size)

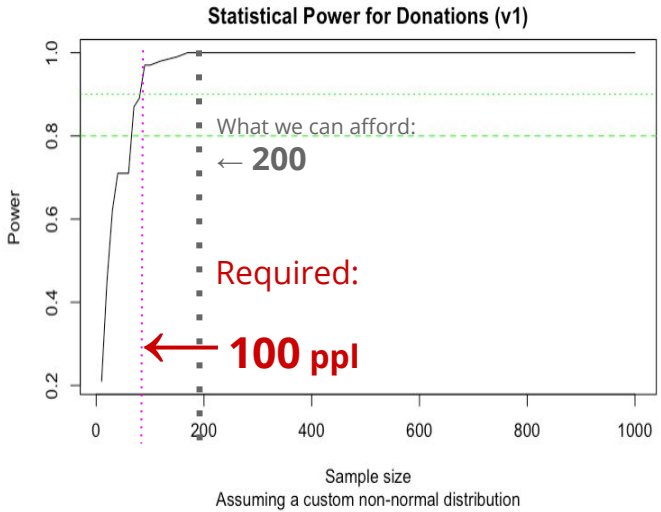




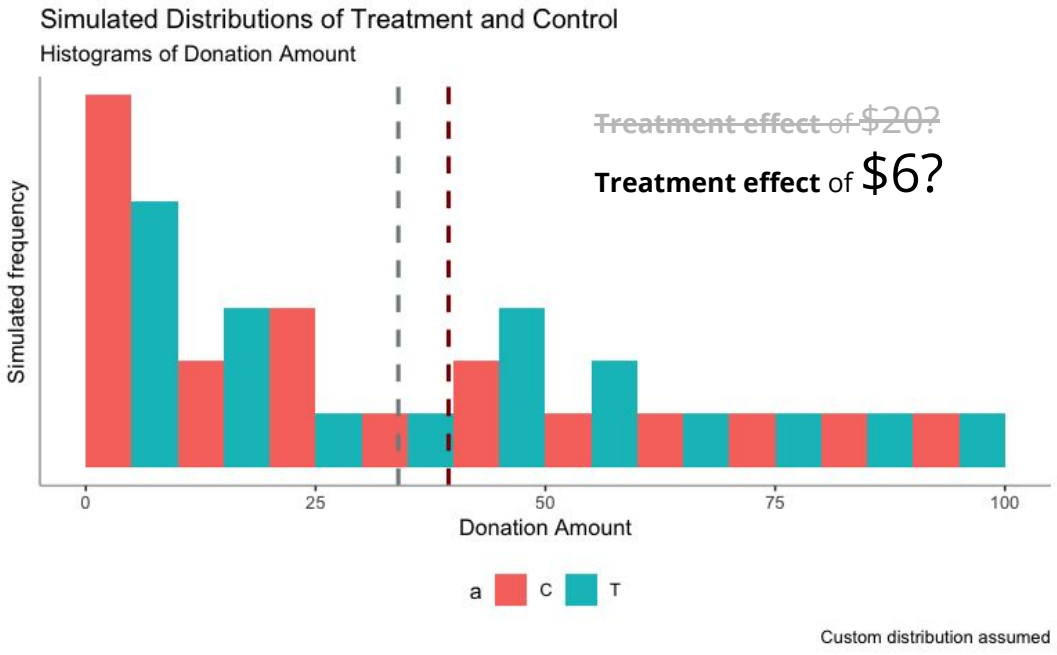
Power Analysis: Donation Amount



Treatment effect of $20?

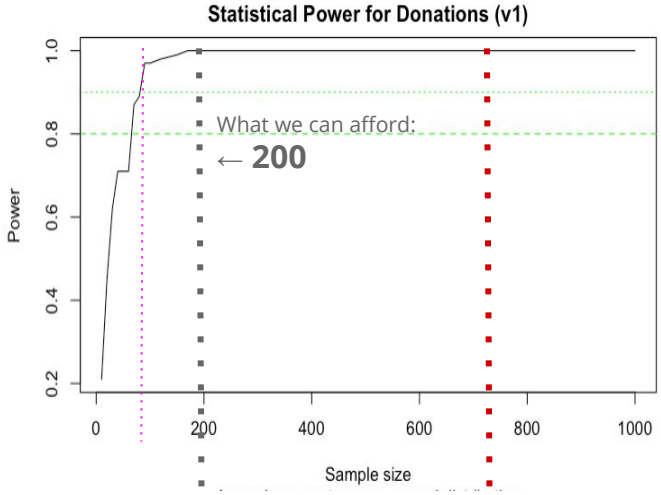


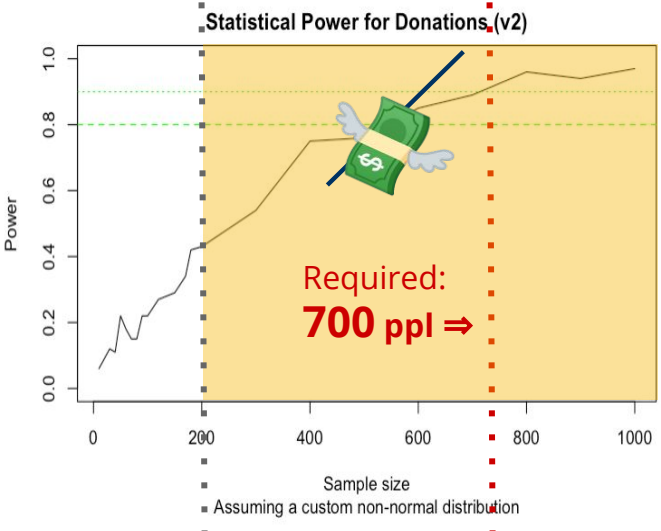
Power Analysis: Donation Amount (v2)



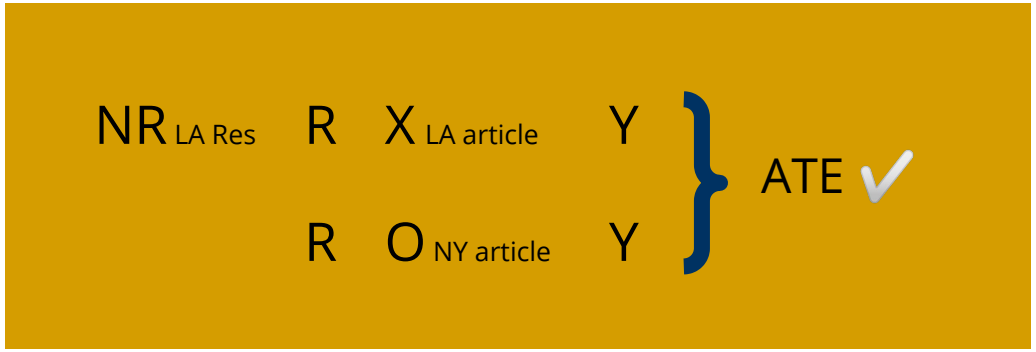
~~Treatment effect of $20?~~

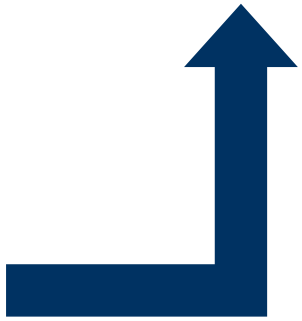
Treatment effect of $6?





200 California
M Turk Workers





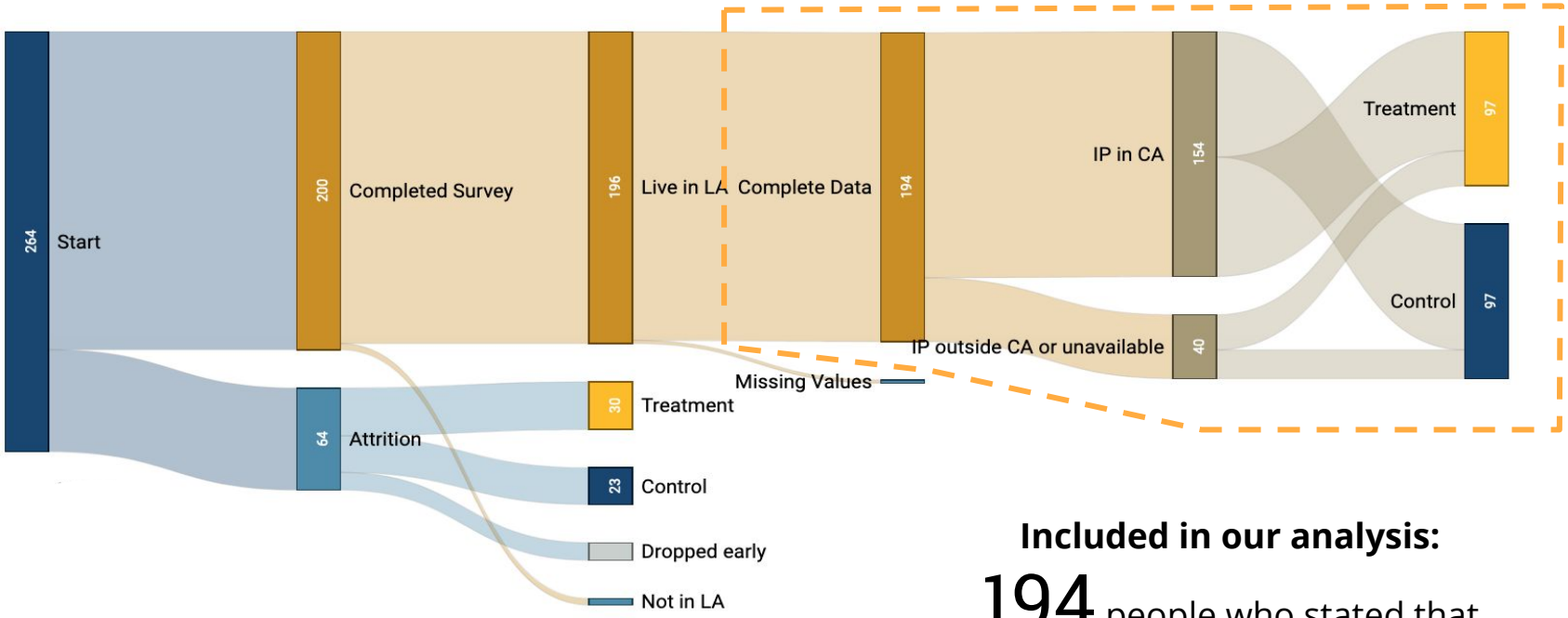
Continuous

- Quiz score
- Donation amount
- Reading time

Ordinal

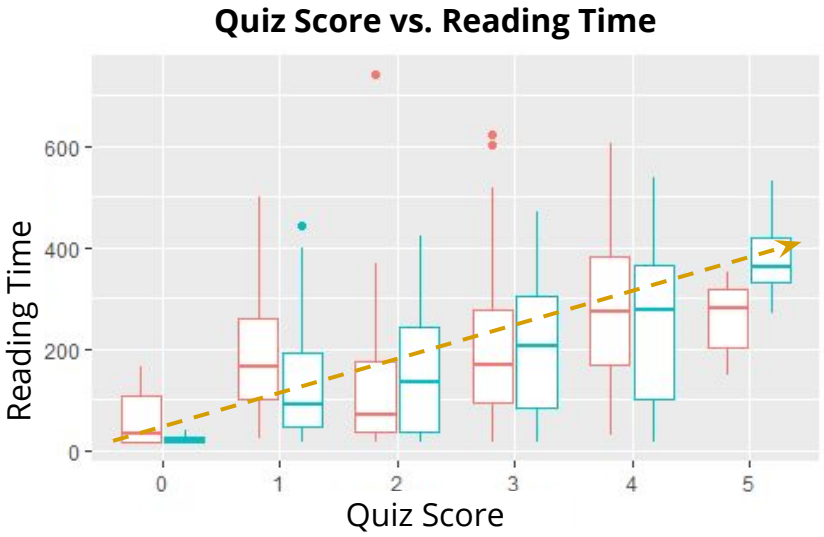
- Importance rating of issue
- Credibility rating of article

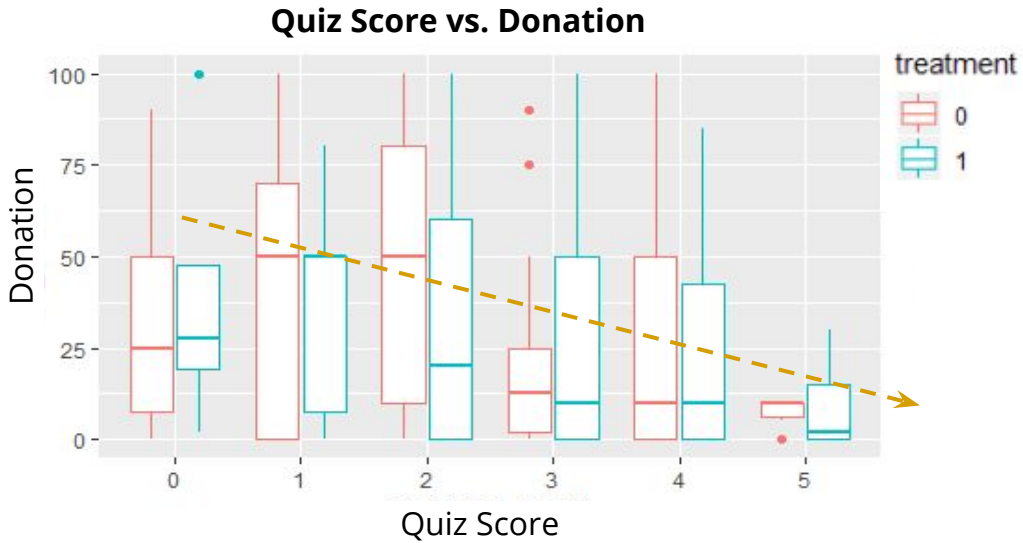
Results



Included in our analysis:

194 people who stated that they live in LA and had complete data (excl. IP address)

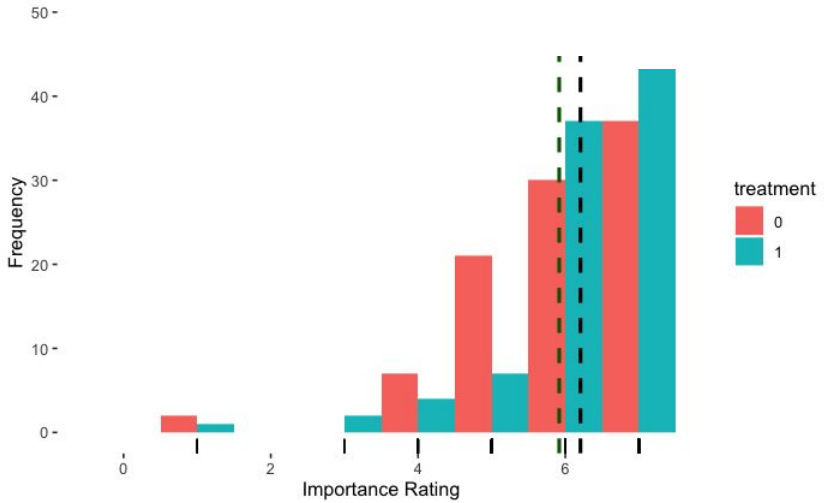




> **People who work hard want to keep their money.**

Probably one of our M Turk Workers

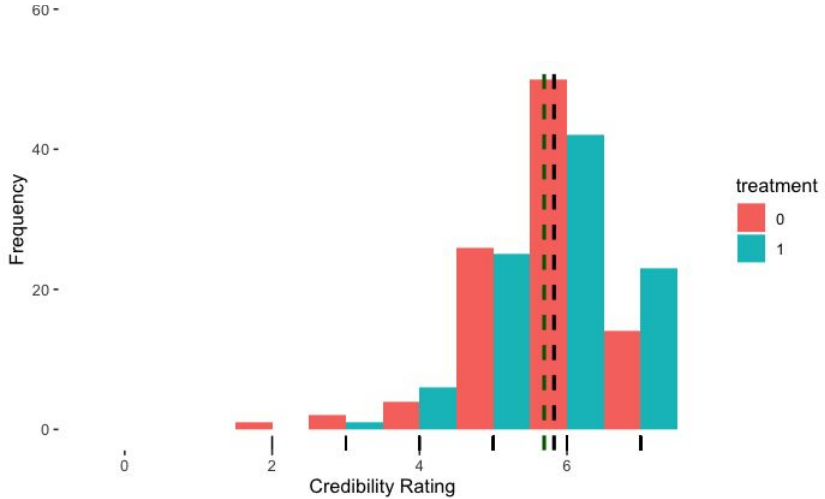
Histogram of Issue Importance




```
Wilcoxon rank sum test with continuity correction

data:  Importance_Control and Importance_Treatment
W = 3969.5, p-value = 0.04491
alternative hypothesis: true location shift is not equal to 0
```

Statistically significant!

Histogram of Article Credibility



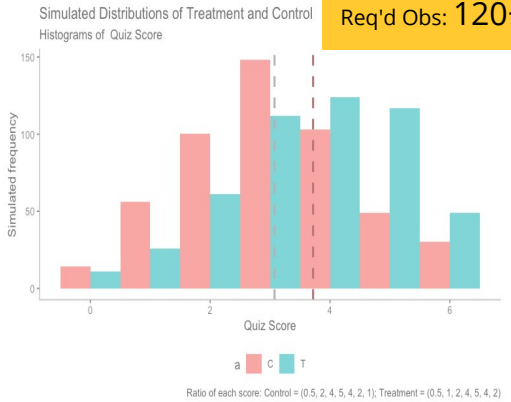

```
Wilcoxon rank sum test with continuity correction

data:  Credibility_Control and Credibility_Treatment
W = 4371, p-value = 0.3607
alternative hypothesis: true location shift is not equal to 0
```

Quiz Scores

Our assumption was right-on!

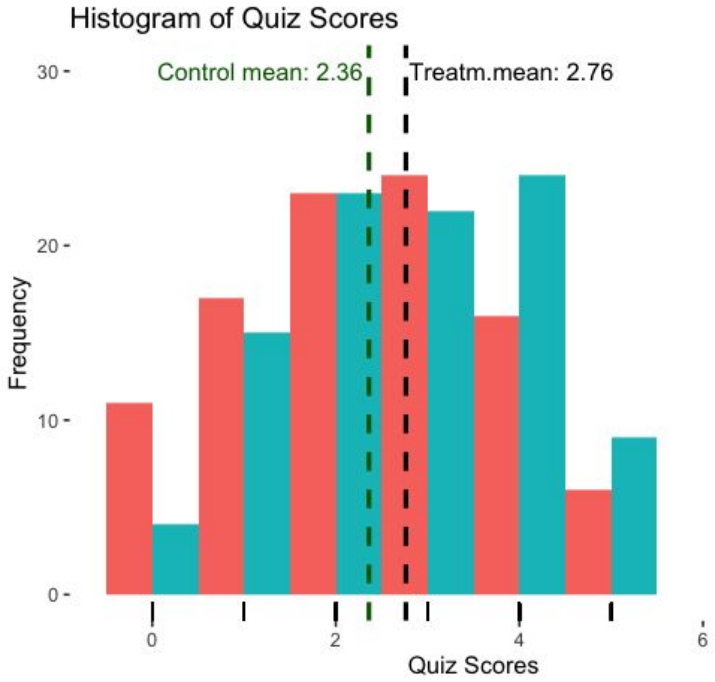
Assumed for Power Analysis

Req'd Obs: 120+



Actual

Actual Obs: 194



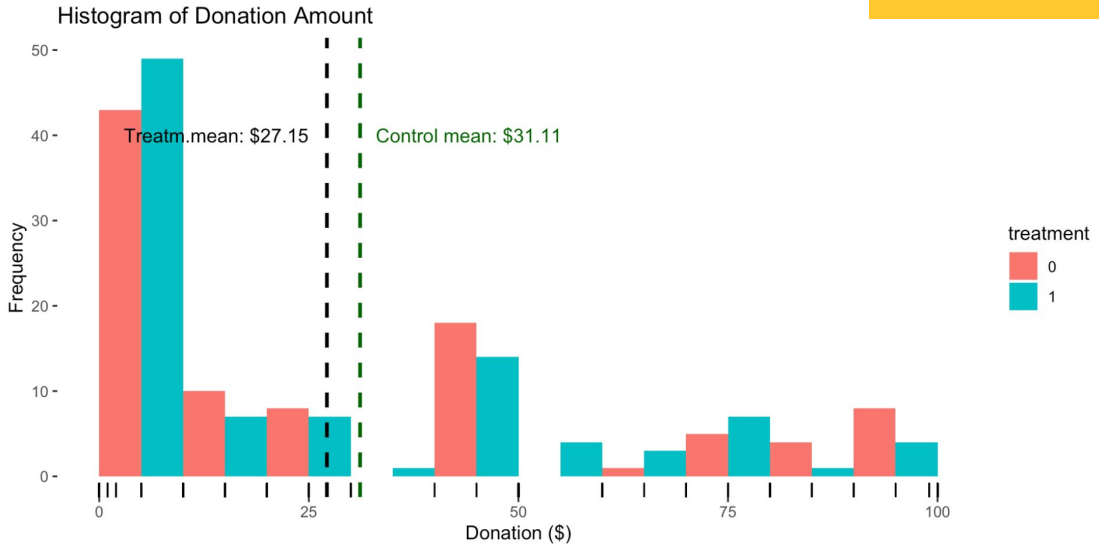
Donation Amount

We might not have enough power.

Assumed for Power Analysis

Actual

Actual Obs: 194



Interpretations

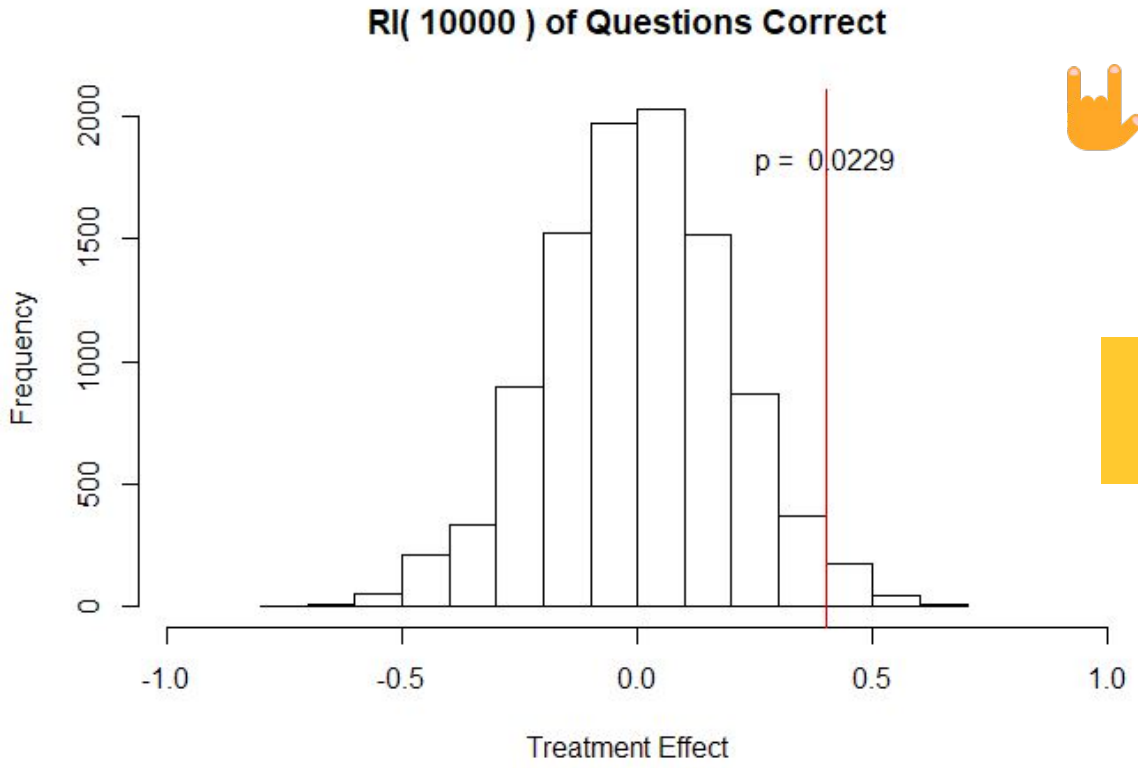
Comparing Treatment Effects

	Dependent variable:		
	Questions Correct	Article Read Time (seconds)	Donation in USD
	(1)	(2)	(3)
Treatment	0.402**	10.946	-3.959
	p = 0.044	p = 0.630	p = 0.393
Observations	194	194	194
R^2	0.021	0.001	0.004
Adjusted R^2	0.016	-0.004	-0.001
Residual Std. Error (df = 192)	1.377	157.746	32.167
F Statistic (df = 1; 192)	4.136**	0.234	0.735
Note:			*p<0.1; **p<0.05; ***p<0.01

Statistically significant!

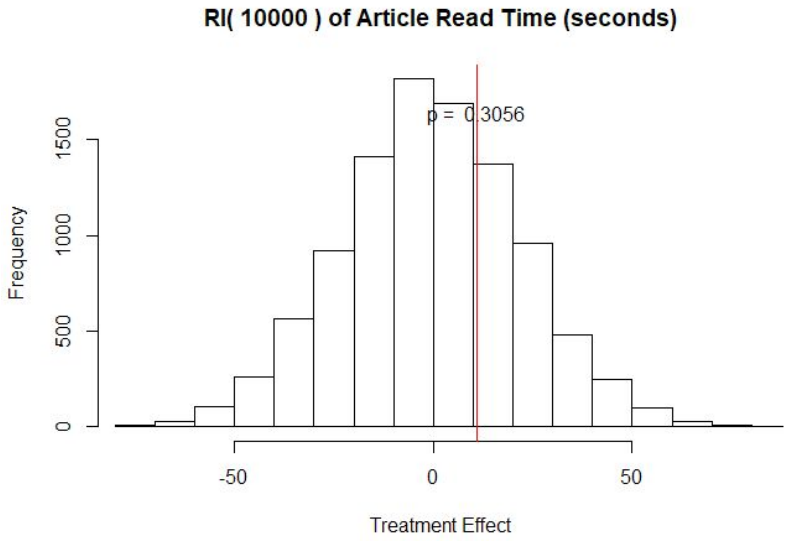
Locality is statistically significant for **quiz scores**, but NOT for the other two. R2 is abysmal.

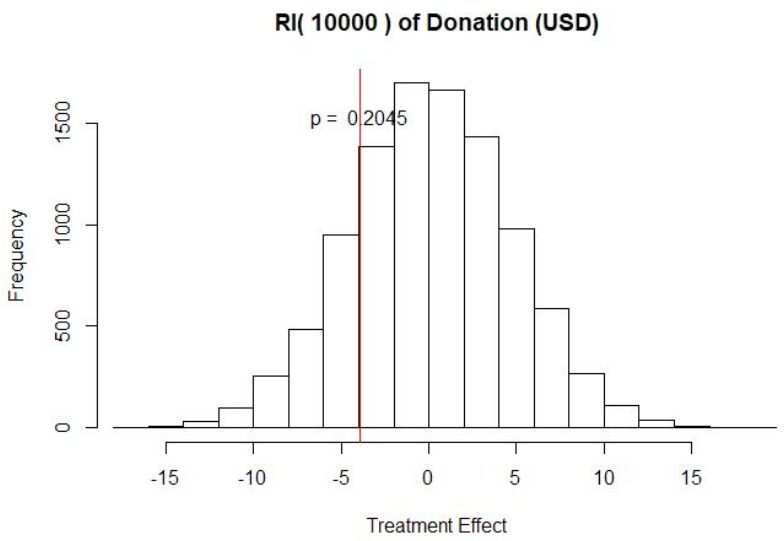
RI: Quiz Score



STILL statistically significant!

RI: Reading Time & Donation





Comparing Treatment Effects

	Dependent variable:	
	Questions Correct	
	(1)	(2)
Treatment	0.402**	0.376**
	(0.198)	(0.177)
Read time > 120		1.267***
		(0.179)
Observations	194	194
R^2	0.021	0.224
Adjusted R^2	0.016	0.216
Residual Std. Error	1.377 (df = 192)	1.229 (df = 191)
F Statistic	4.136** (df = 1; 192)	27.578*** (df = 2; 191)
Note:		*p<0.1; **p<0.05; ***p<0.01

STILL statistically significant!

Adding the **long reading time** to the regression does NOT explain the treatment effect away.

Generalizability Concerns

Science communication is broad

This single article does **not** represent all of science communications

Observed behavior might be different from organic behavior

Lab environment ≠ the real world

Comprehension is difficult to quantify

Operational variables may not be the best measurement of engagement

Surveys responses vary wildly in quality

Mechanical Turk respondents have different incentive structures

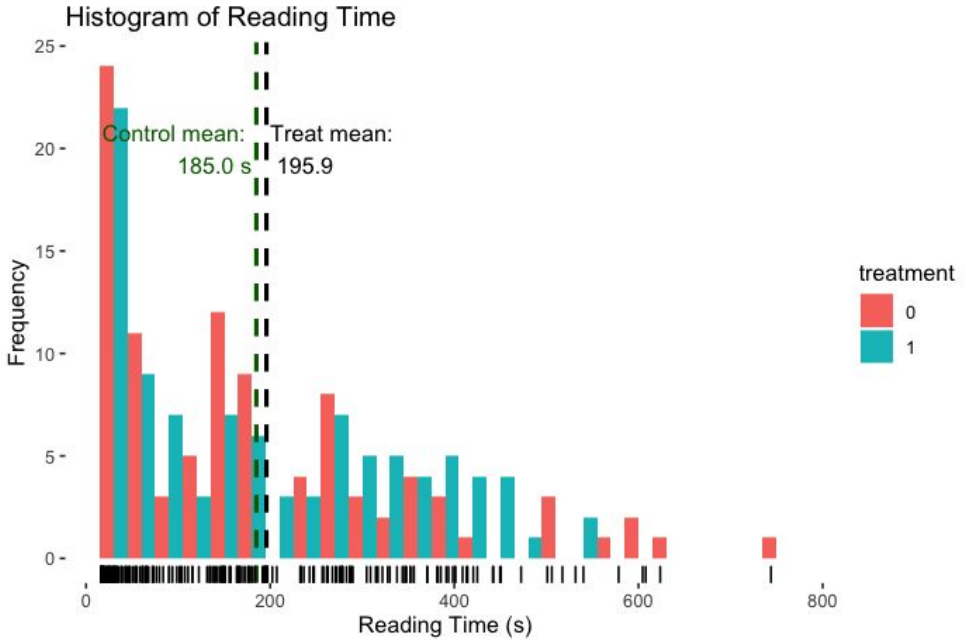
Reading Time

The monetary incentive likely affected the time that M Turk workers spend reading the article

28% under 60 seconds

41% under 120 seconds

58% under 180 seconds



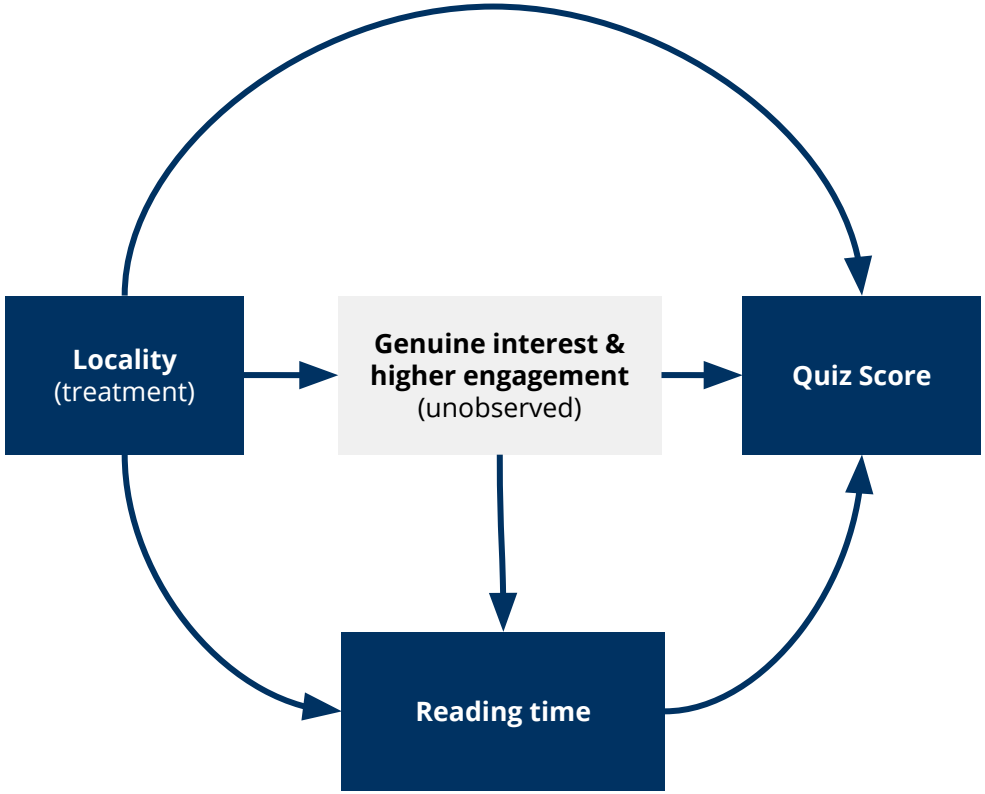
What to Make of the Reading Time?

Our question to YOU!

How should we weave reading time into the story/analysis?

1. It's not an i**nstrumental variable**
2. Not the best for measuring **compliance** either

For now, we chucked it in a regression as a covariate. Is this OK?



References

- **Theme song:** Cyborg Octopus - Data_M1nefield <https://youtu.be/mJGbgnPI-s0>
- **Images**:
  - https://towardsdatascience.com/collecting-news-articles-through-rss-atom-feeds-using-python-7d9a65b06f70
  - https://www.strategy-business.com/blog/Start-the-New-Year-with-a-simplification-month?gko=f881a
  - https://www.irishtimes.com/news/health/coronavirus-more-striking-evidence-bcg-vaccine-might-protect-against-covid-19-1.4222110
- **Slide theme & layouts**: freegoogleslidestempplates.com
- **Sankey chart**: http://sankey-diagram-generator.acquireprocure.com/

Appendix

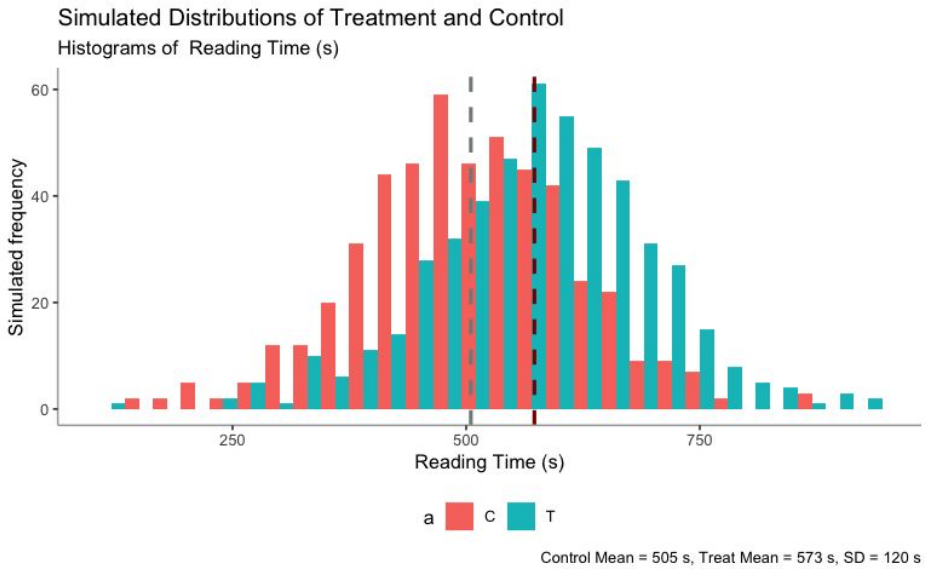
Power Analysis: Reading Time

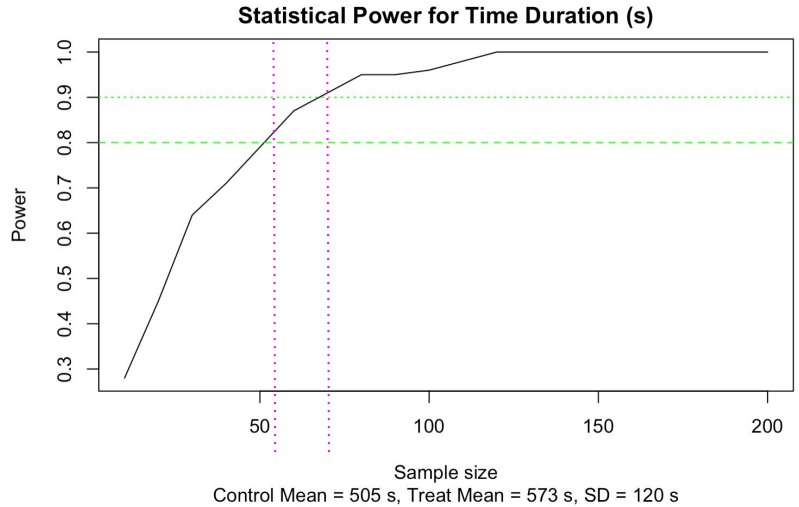
Assumed distribution

- Means: From pilot data, ATE of 70 s
- SD: 20% less than the pilot SD

Power needed

- 100 would comfortably get us above 90% power (p = 0.05, using t-test and 100 simulations for each sample size)

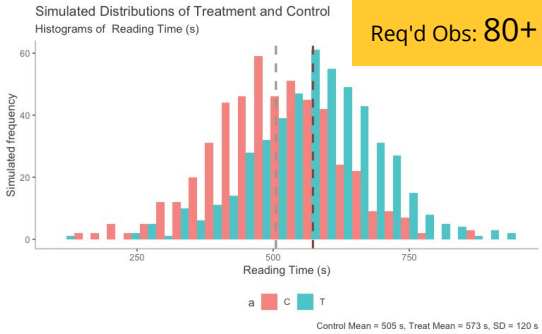




Reading Time

Assumed vs. Actual Distributions

The actual distribution is completely different from our prediction. Might be due to the monetary incentive.

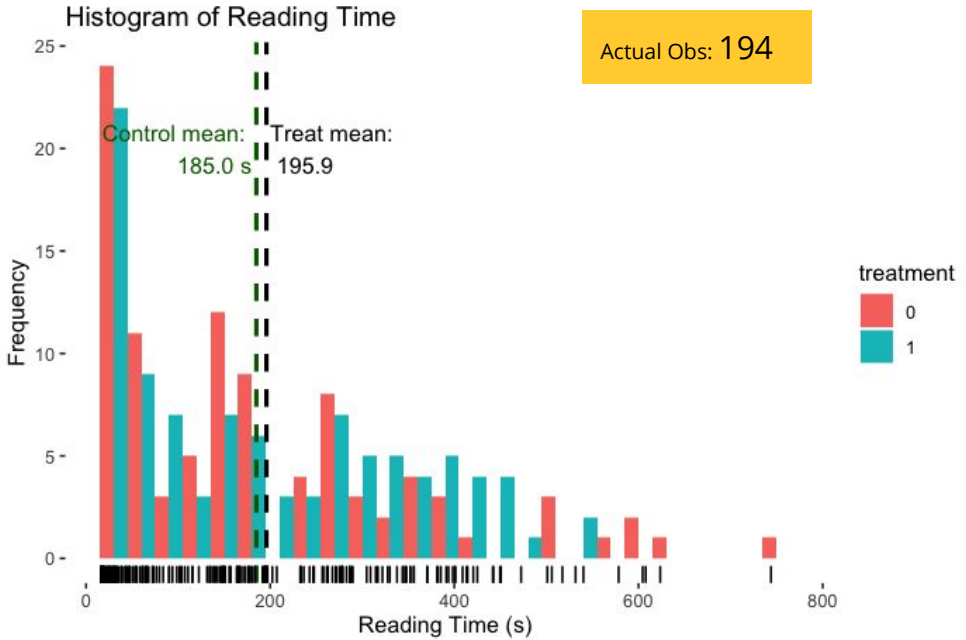


Req'd Obs: 80+

28%
Read the article in under 60 seconds

41%
Read the article in under 120 seconds

58%
Read the article in under 180 seconds



Actual Obs: 194