Randomized Trials (SOC 412)

Week 1 Lecture 2

Sherrerd Hall 306

J. Nathan Matias

@natematias

civilservant.io

jmatias@princeton.edu

Department of Psychology
PRINCETON UNIVERSITY

CITP

mit media lab

What we will cover today

Discuss reading

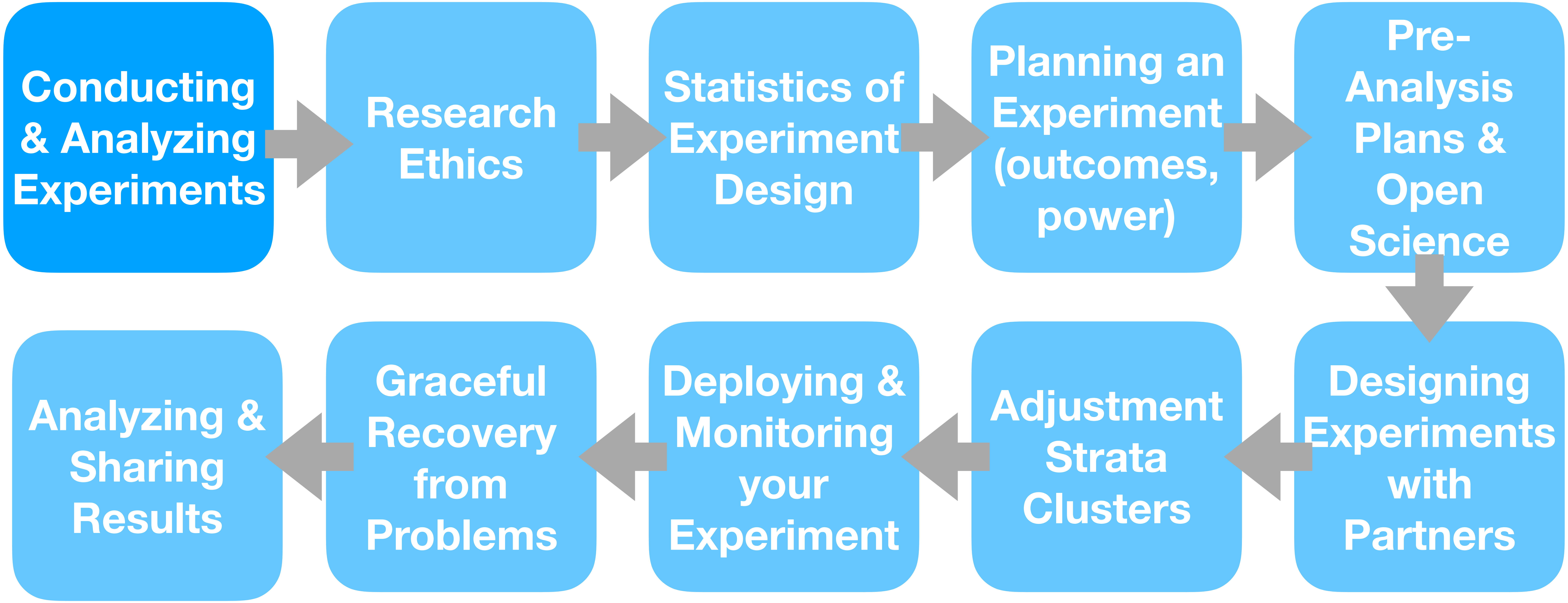
Introduction to randomized trials

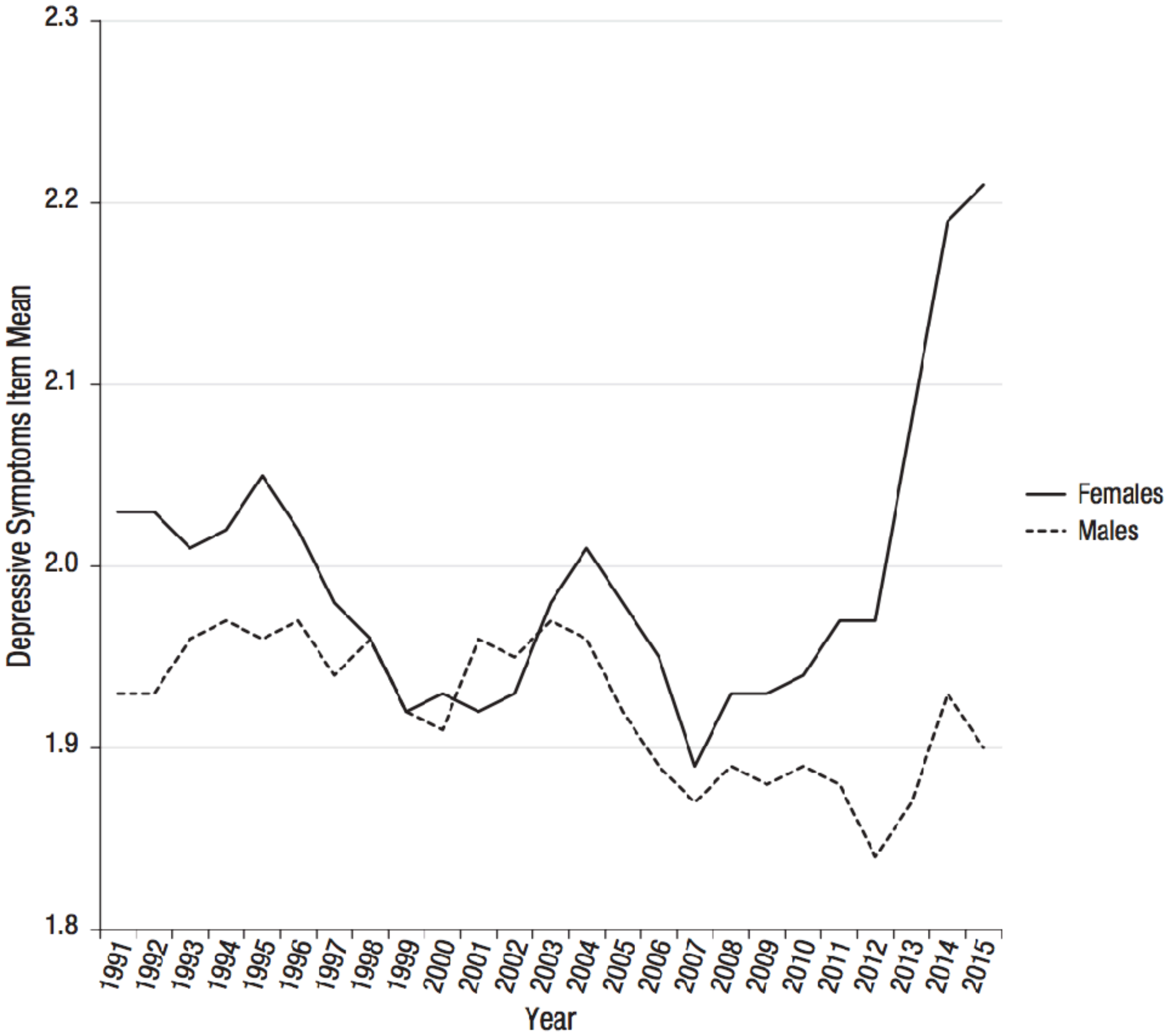
Discuss assignment

Discuss first full field experiment

Questions about Assignments

- Project Assignments: Due Monday evening?
- Reflective Writing Due Friday midday?


Conducting & Analyzing Experiments
Research Ethics
Statistics of Experiment Design
Planning an Experiment (outcomes, power)
Pre-Analysis Plans & Open Science
Designing Experiments with Partners
Adjustment Strata Clusters
Deploying & Monitoring your Experiment
Graceful Recovery from Problems
Analyzing & Sharing Results



Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

“ Next, we examined **possible causes of the increase in depressive symptoms** and suicide-related outcomes among adolescents

When examined individually, the four **suicide-related outcomes** were **all significantly correlated with electronic device use**

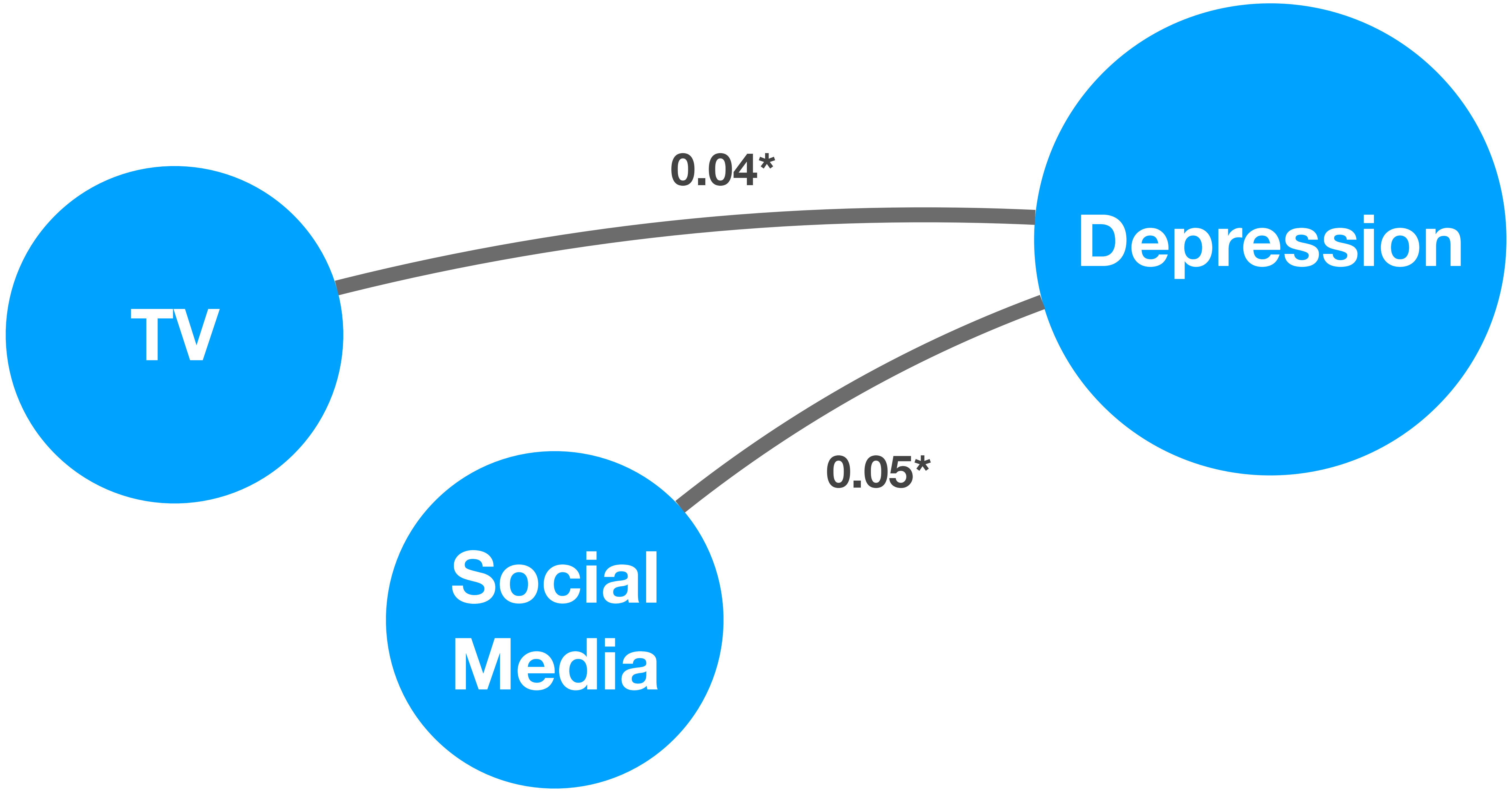
Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

Table 2. Correlations Between Screen and Nonscreen Activities and Depressive Symptoms (8th and 10th Graders, MtF) and Suicide-Related Outcomes (9–12th Graders, YRBSS), 2009–2015 *p < .001.

	Bivariate *r*	Controlled for sex, race, SES, grade, and region	Controlled for sex, race, SES, grade, and region and in-person social interaction	Girls (controlled for race, SES, grade, and region)	Boys (controlled for race, SES, grade, and region)
MtF (correlations with depressive symptoms)					
Screen activities					
Social media use	.05*	.03*	.06*	.06*	.01
TV viewing	.04*	.02*	.03*	.03*	.02*
Internet news use	.00	.00	.01*	.01	–.02
Nonscreen activities					
In-person social interaction	–.07*	–.08*	–.09*	–.08*	–.09*
Religious services attendance	–.15*	–.14*	–.14*	–.16*	–.13*
Sports or exercise	–.22*	–.19*	–.18*	–.20*	–.19*
Homework hours	–.06*	–.05*	–.06*	–.06*	–.04*
Print media use	–.11*	–.10*	–.09*	–.12*	–.08*
Having a paid job	.00	.01	.02*	–.01	.02*

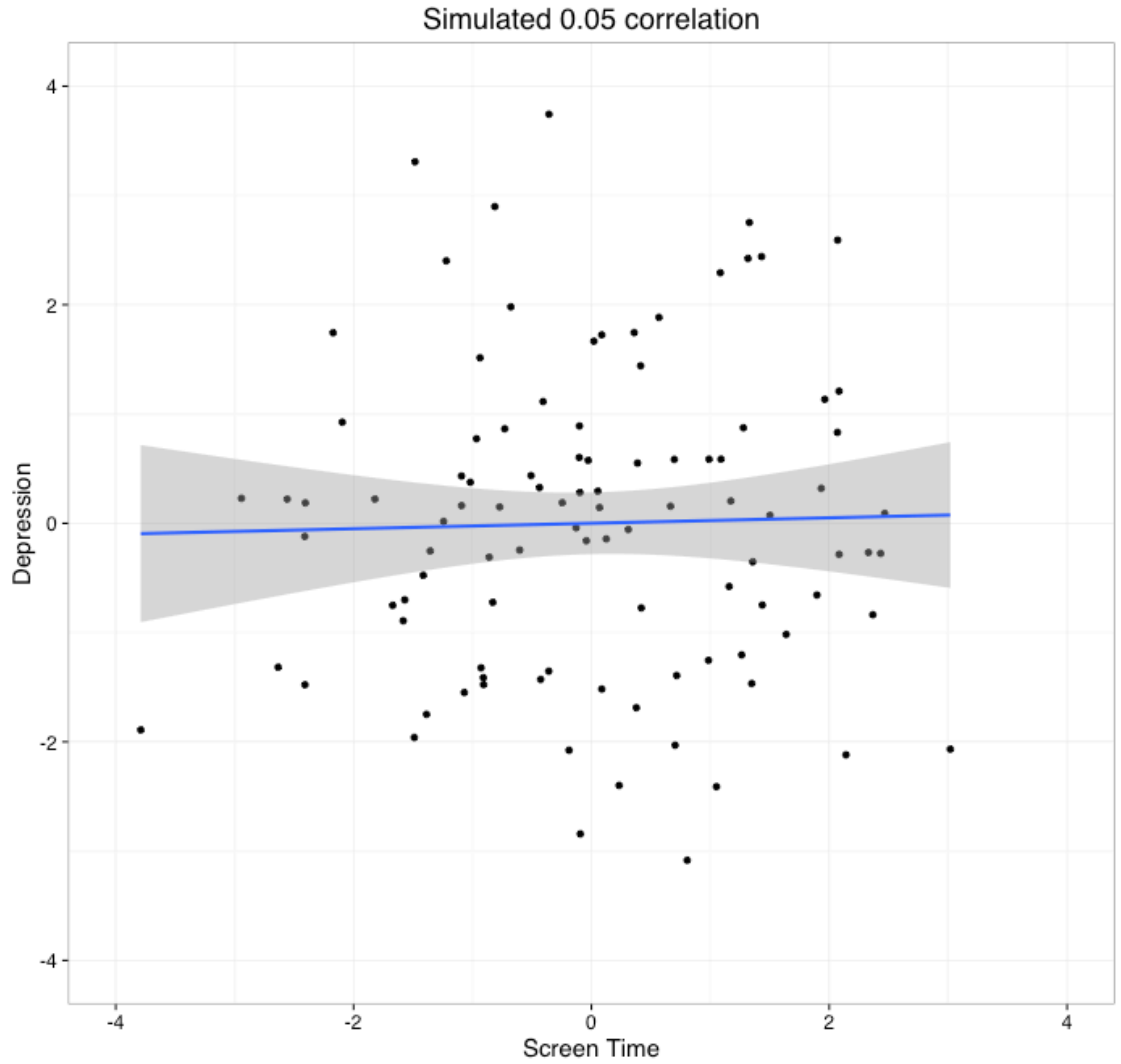
Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

Correlation? Association? Description?



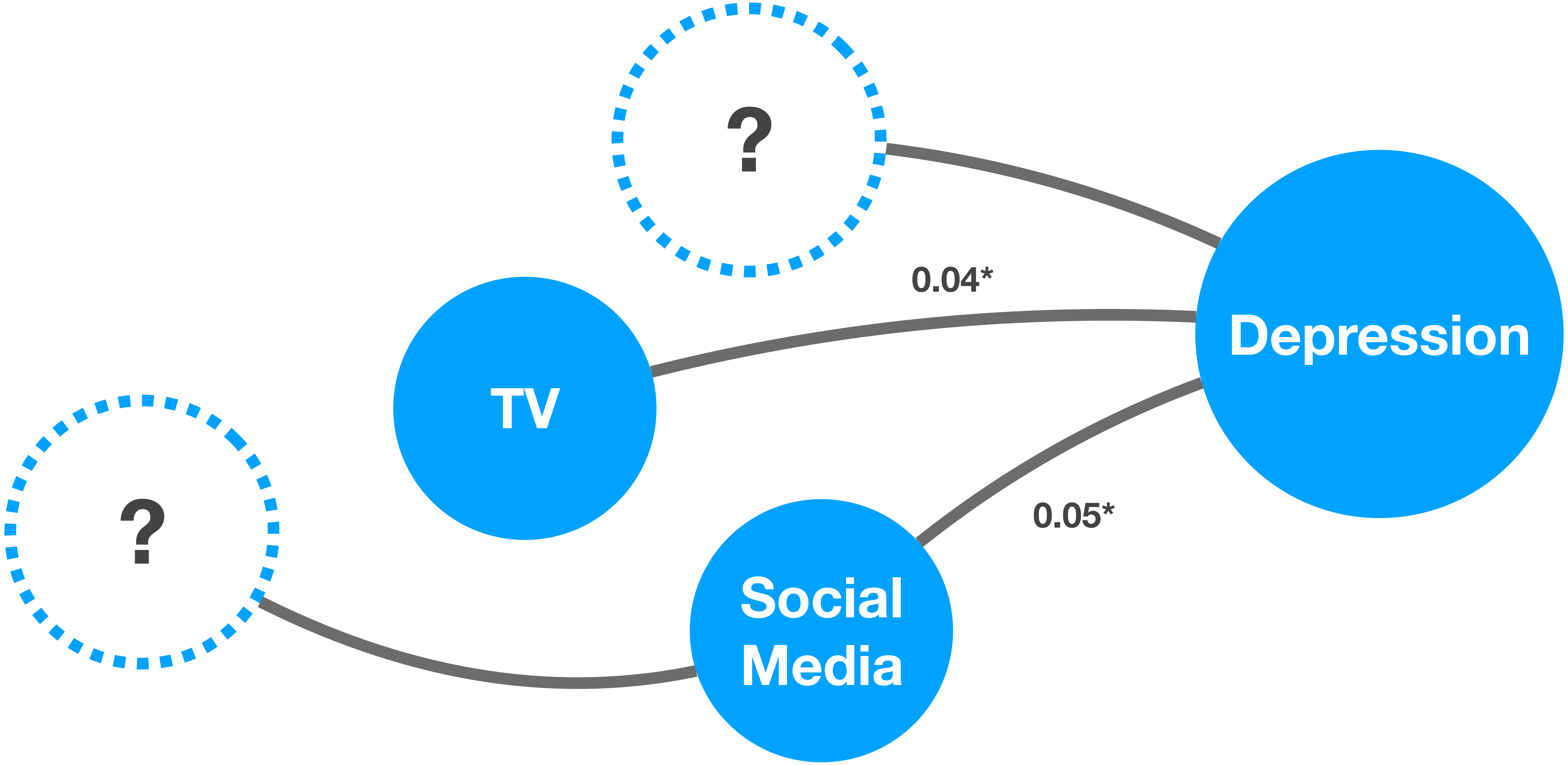
Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

Correlation? Association? Cause?



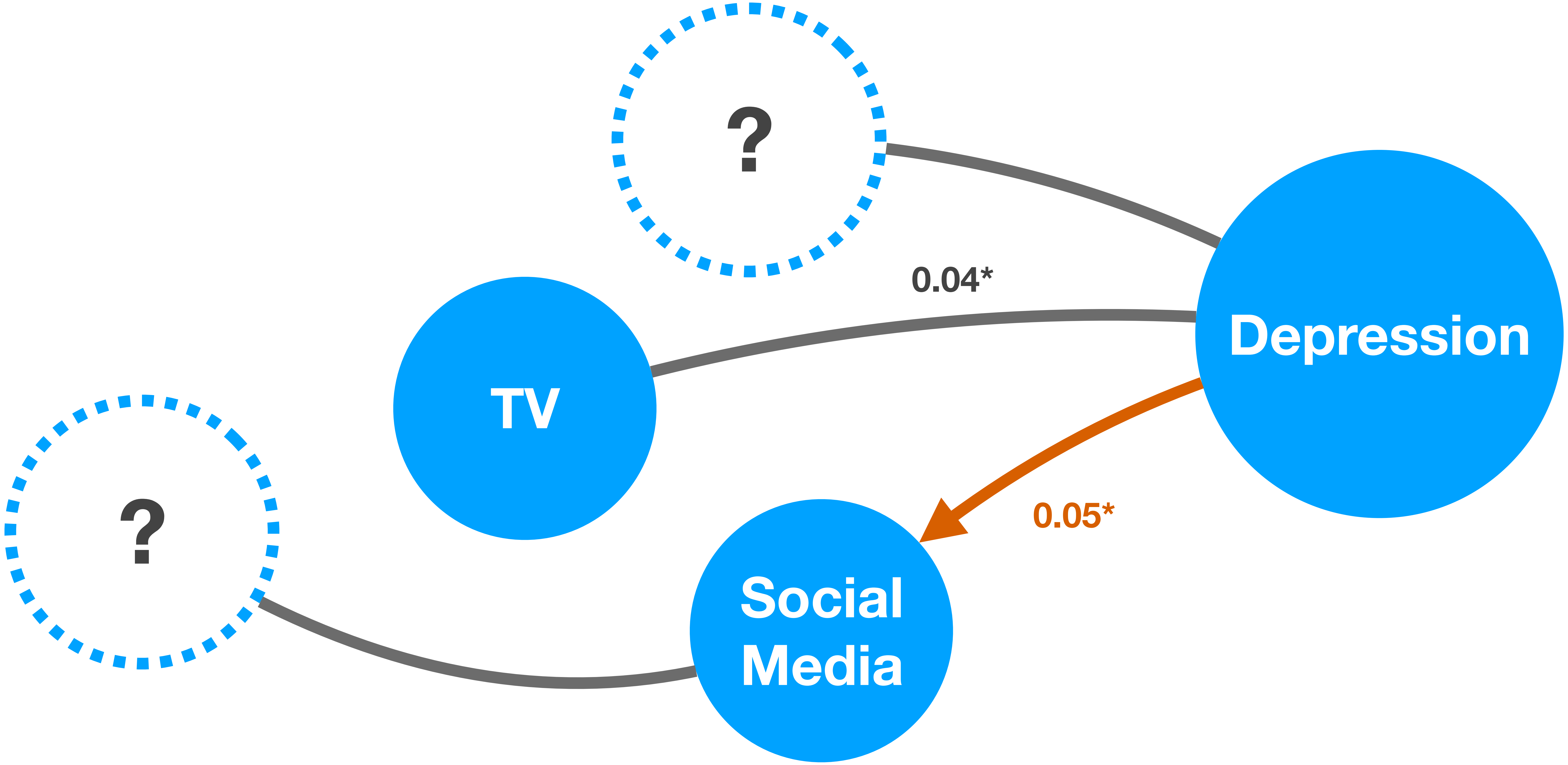
Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

Unobserved Confounders



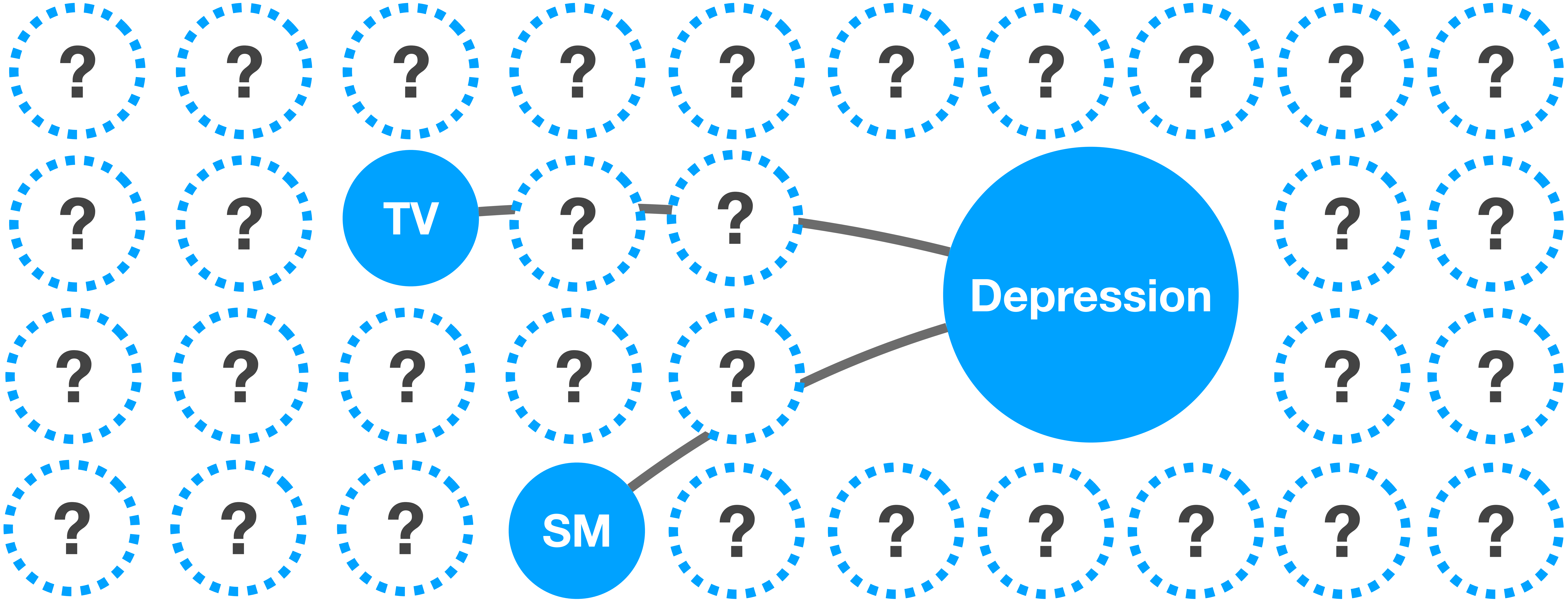
Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

Direction of Causality (Selection Bias)



Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

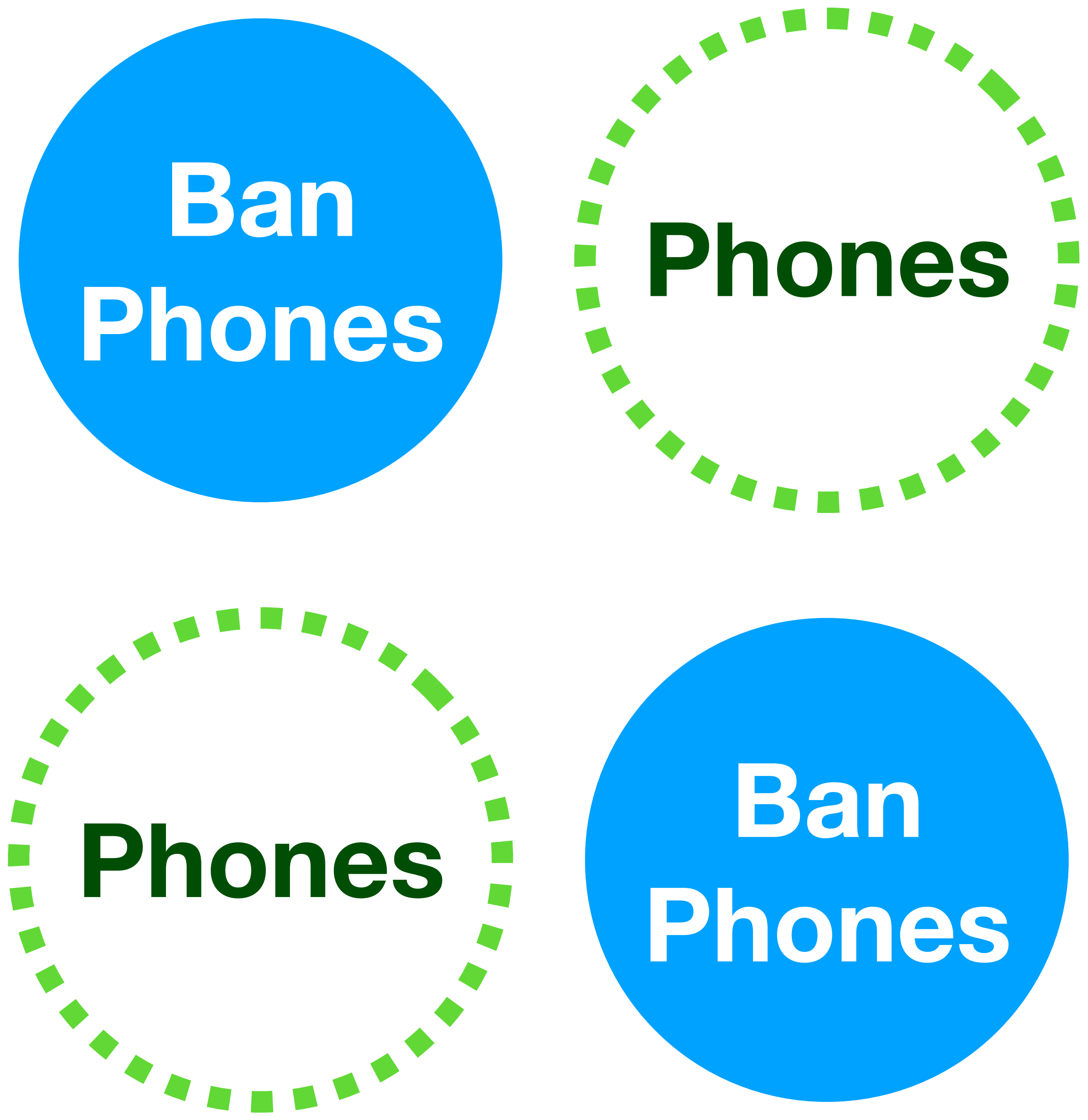
Multiple Comparisons

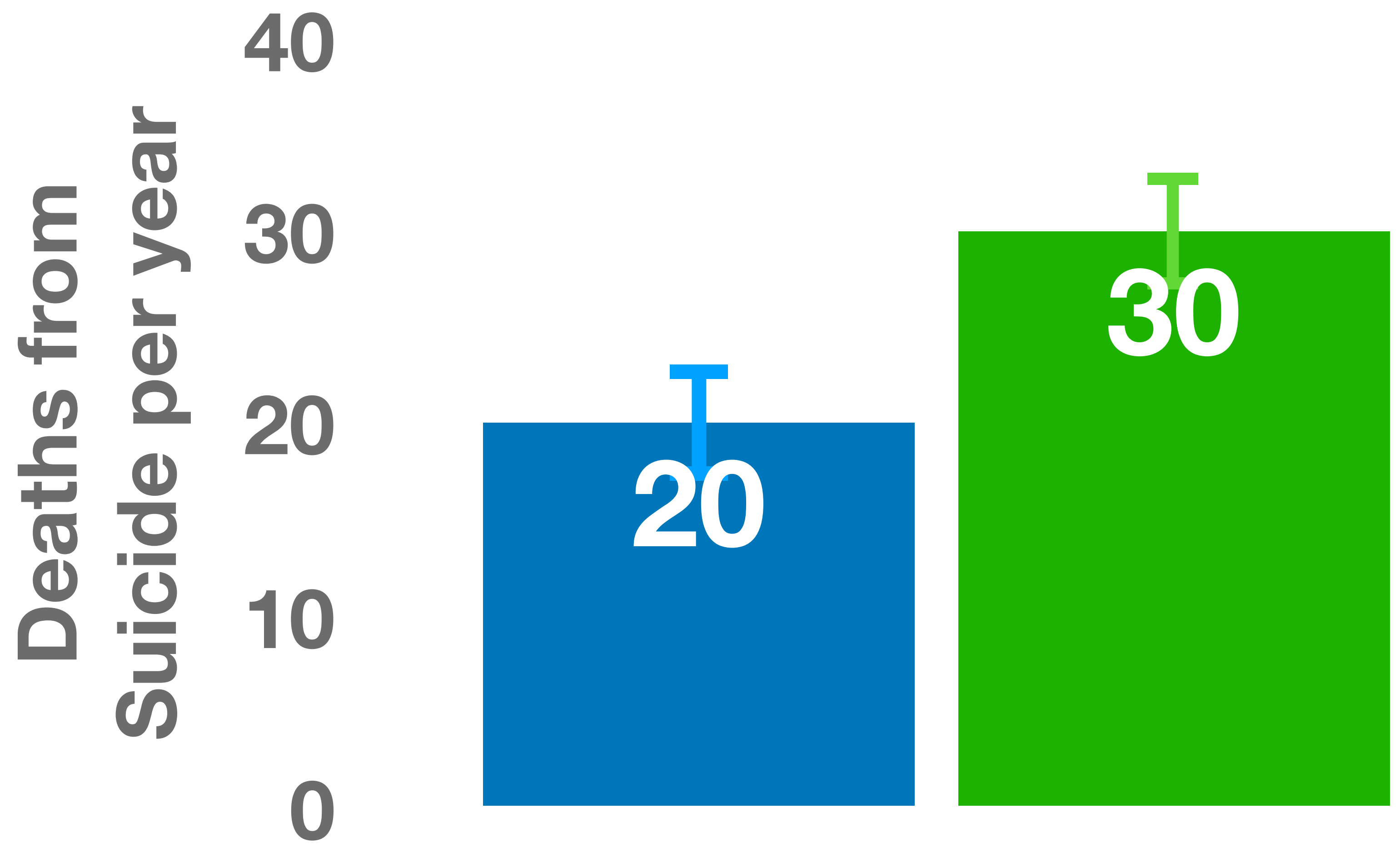


Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, *6*(1), 3-17.

Observational Data Analysis

(**imagine** a dataset of depression cases)
(for schools that banned phones and those that didn't)





Why Experiment? (methodology)

- **Causal explanation** vs **description**
- Addresses **unobserved confounders**
- **Unbiased inference** (e.g. if the experiment were replicated an infinite number of times, our model would generate the right answer on average)
- Up-front design (“ex ante”) in principle **limits the analyst’s discretion**

Parts of an Experiment (example)

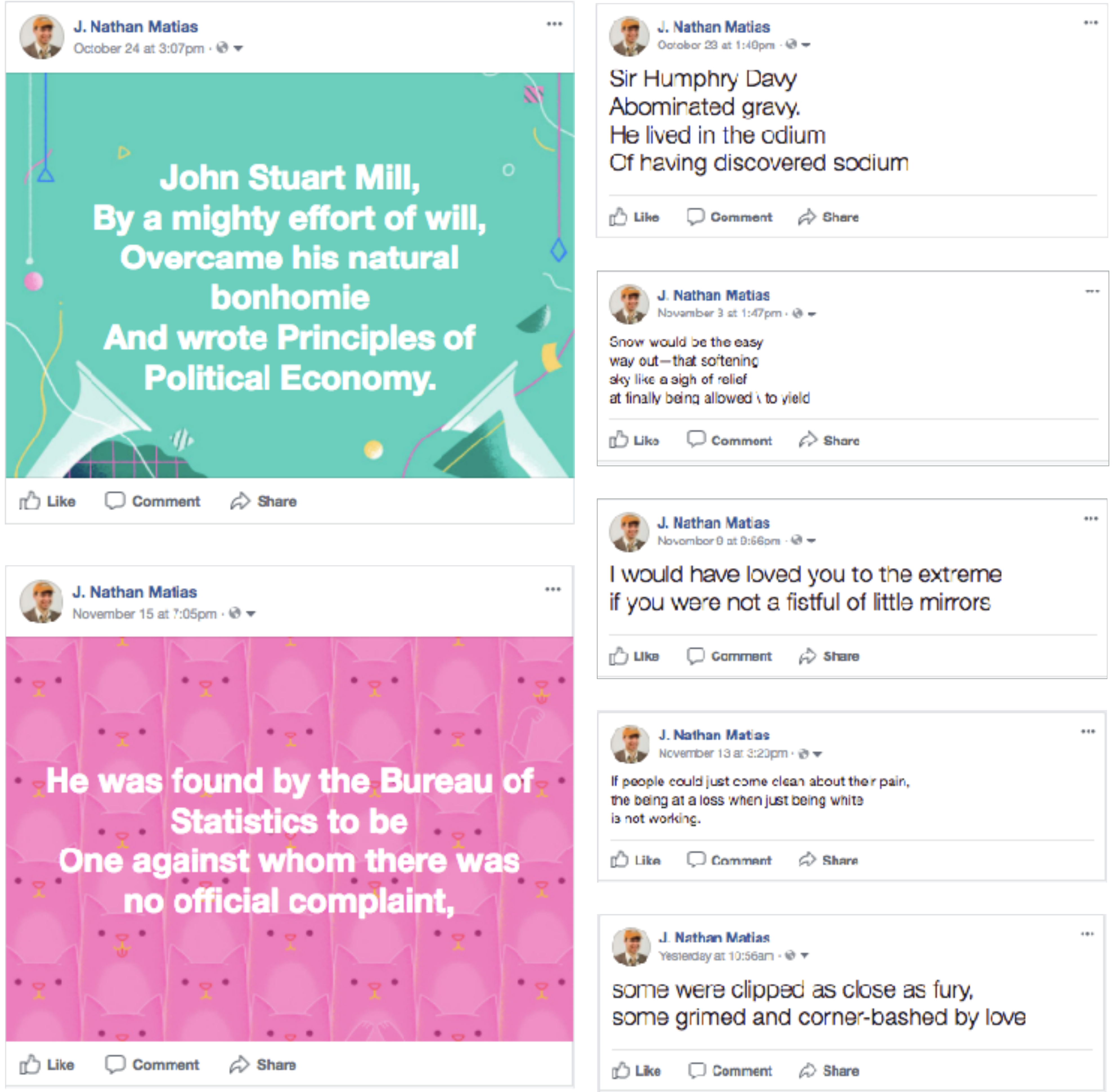
> When the skies opened and **dumped $185 million on an obscure Chicago poetry journal** in 2002, it was as much a burden as a gift…
>
> Barr still speaks in unpoetic terms like "cost per impact"

Fisher, Daniel. ***No Rhyme or Reason***. Forbes, Jan 7, 2011.

Parts of an Experiment

1. Does A Colored Background Increase Facebook Interactions on average?

2. If so, **by how much** on average?



Parts of an Experiment

Background Color

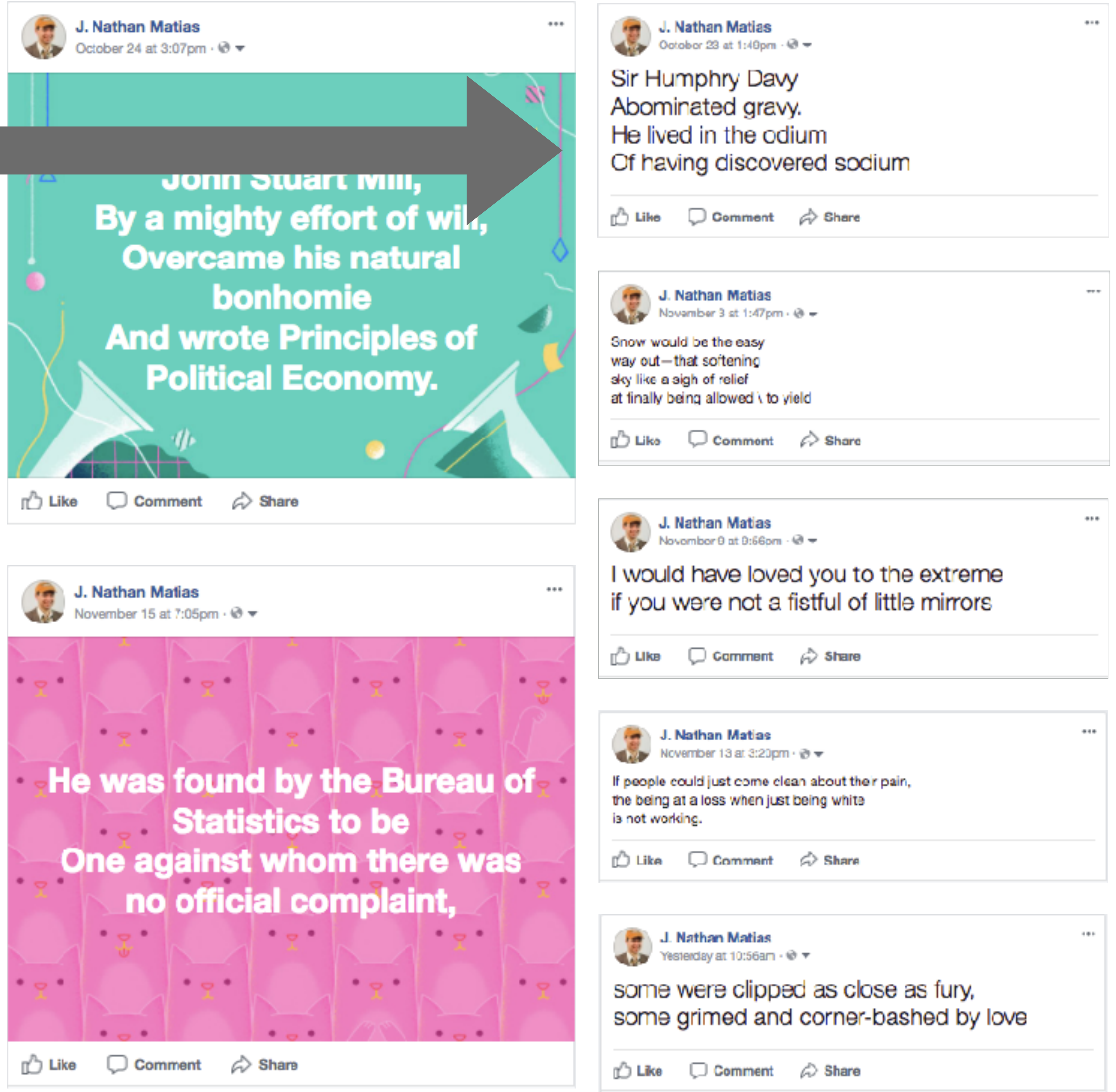
?

Comparison
(control)
(placebo)

1. Does A Colored Background Increase Facebook Interactions on average?

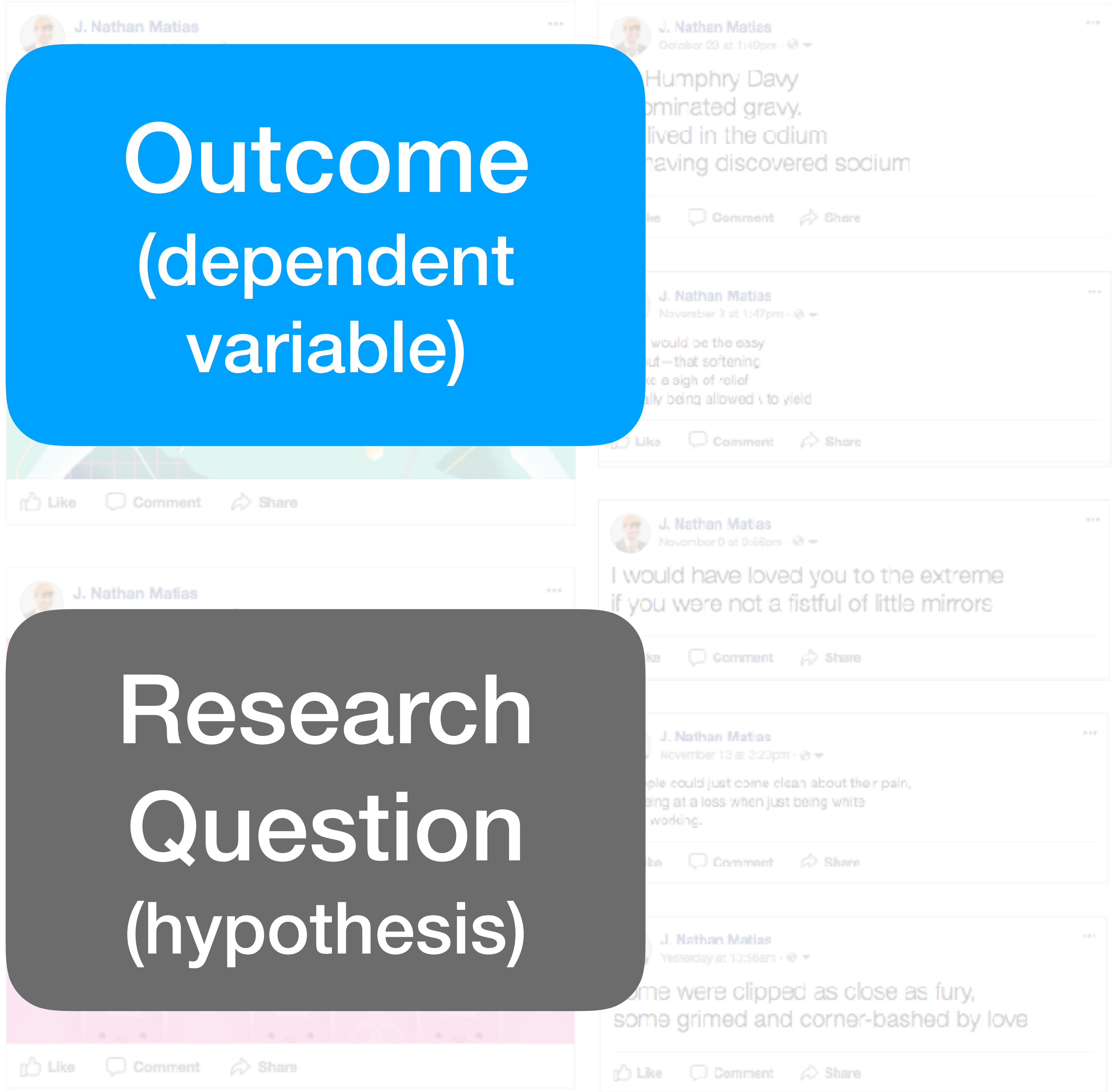
2. If so, by how

Intervention
(treatment)



Parts of an Experiment

Background Color ? Facebook Interactions



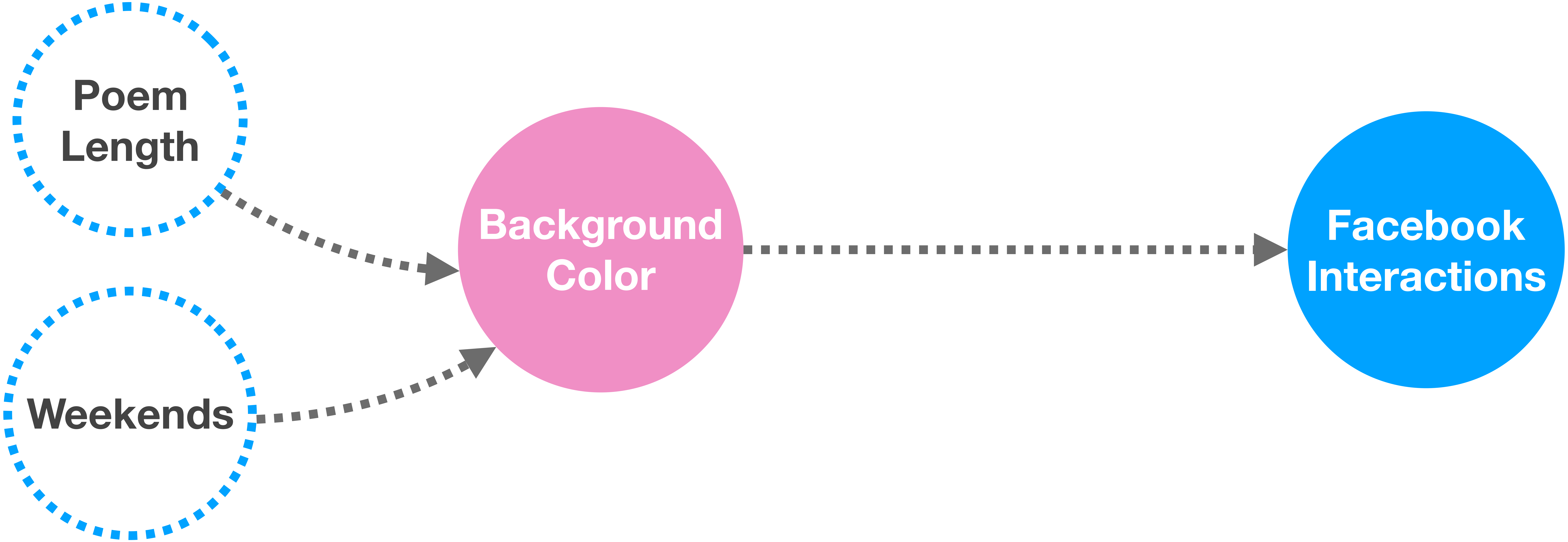
Outcome (dependent variable)

Research Question (hypothesis)

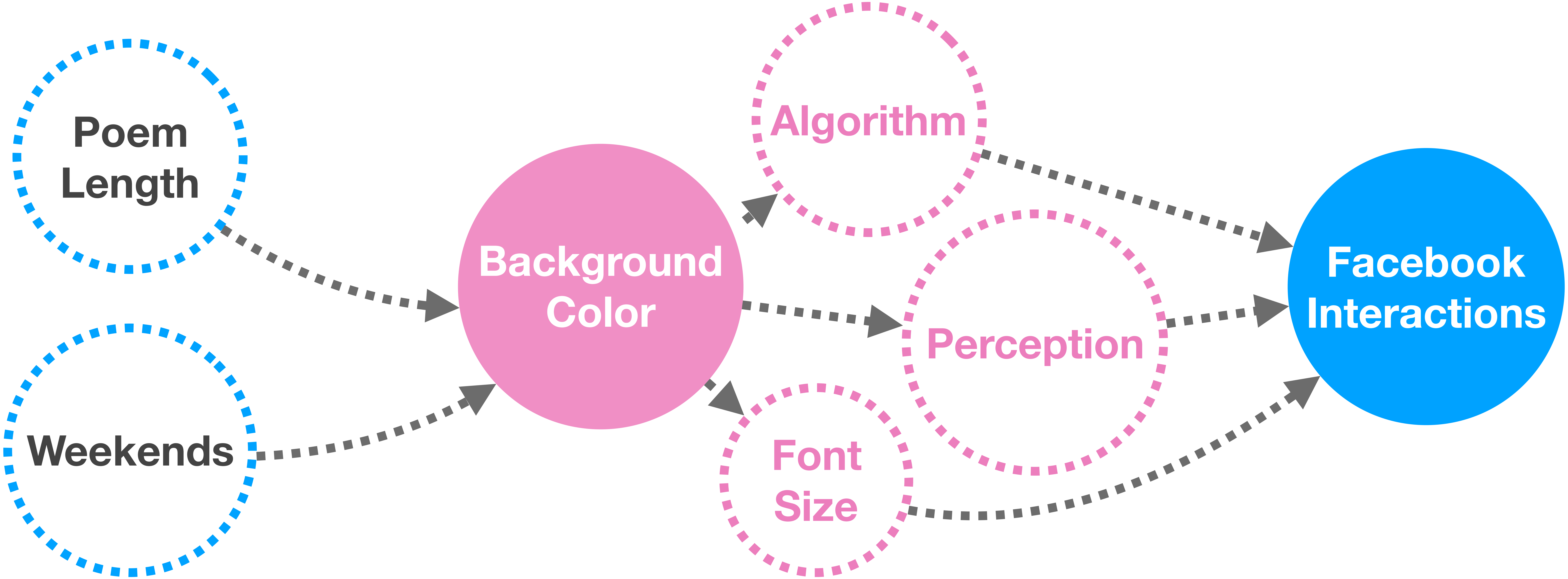
1. Does A **Colored Background** **Increase** **Facebook Interactions** on average?
2. If so, **by how much** on average?

Unobserved Confounders

(by the way, this graph is called a path diagram)

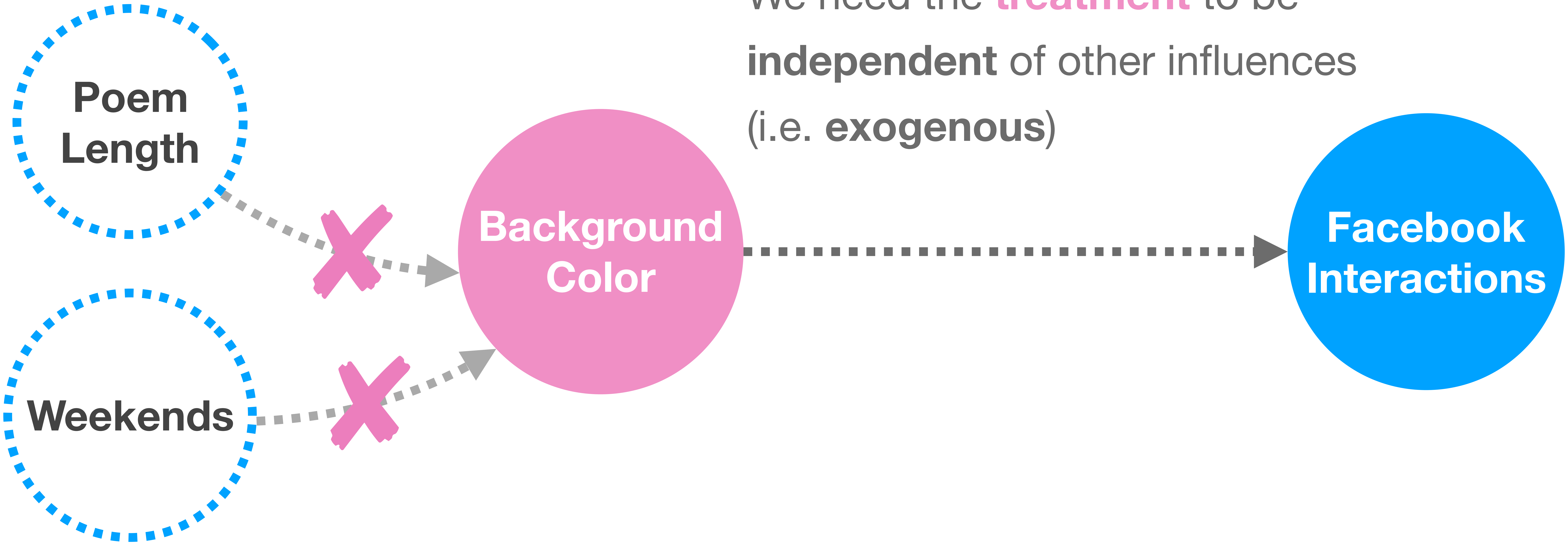


Unobserved Confounders



Independence of the Treatment

We need the **treatment** to be **independent** of other influences (i.e. **exogenous**)



Random Assignment

By using random assignment to choose between the **treatment** and **control** (conditions)

we ensure that the treatment is **independent** (exogenous) of outside influence, and that any **difference in outcomes** is due to **the effect of the intervention**

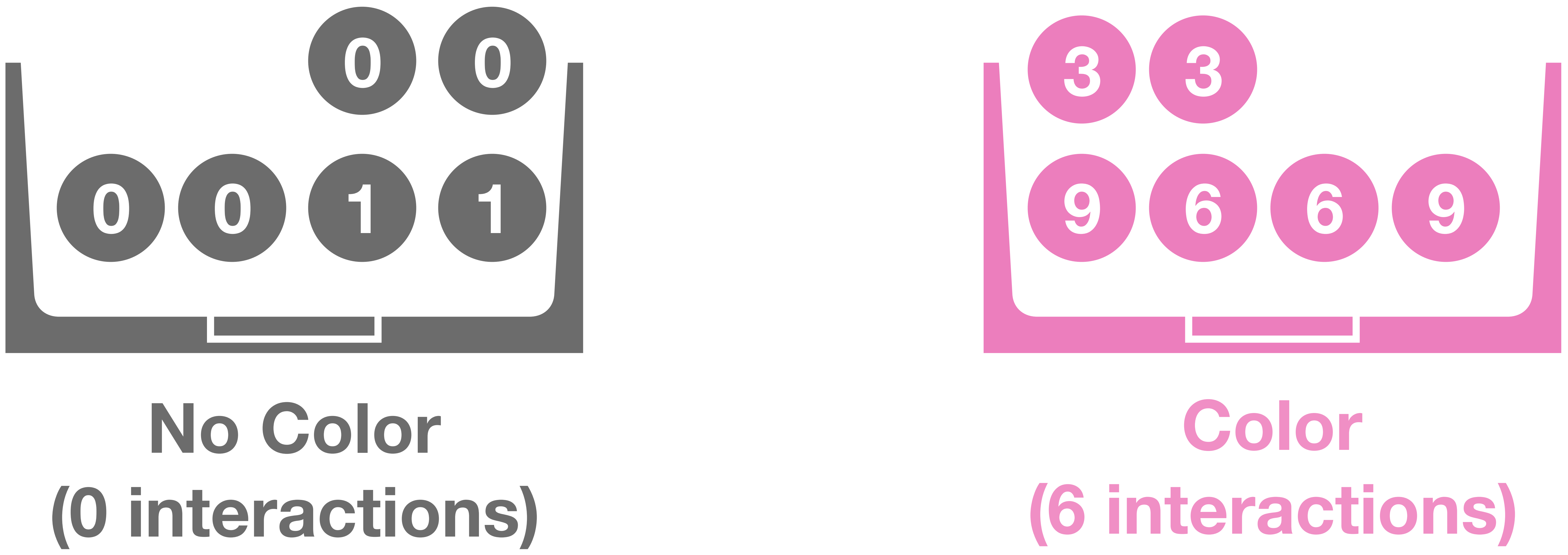
Understanding Randomization via Sampling

Imagine each poem has a basic potential “interactability” and that the poem-poster unconsciously allocates colors based on that factor

0 6 3 0 1 1 3 6 9 9 0 0

Understanding Randomization via Sampling

Imagine each poem has a basic potential "interactability" and that the poem-poster unconsciously allocates colors based on that factor



Understanding Randomization via Sampling

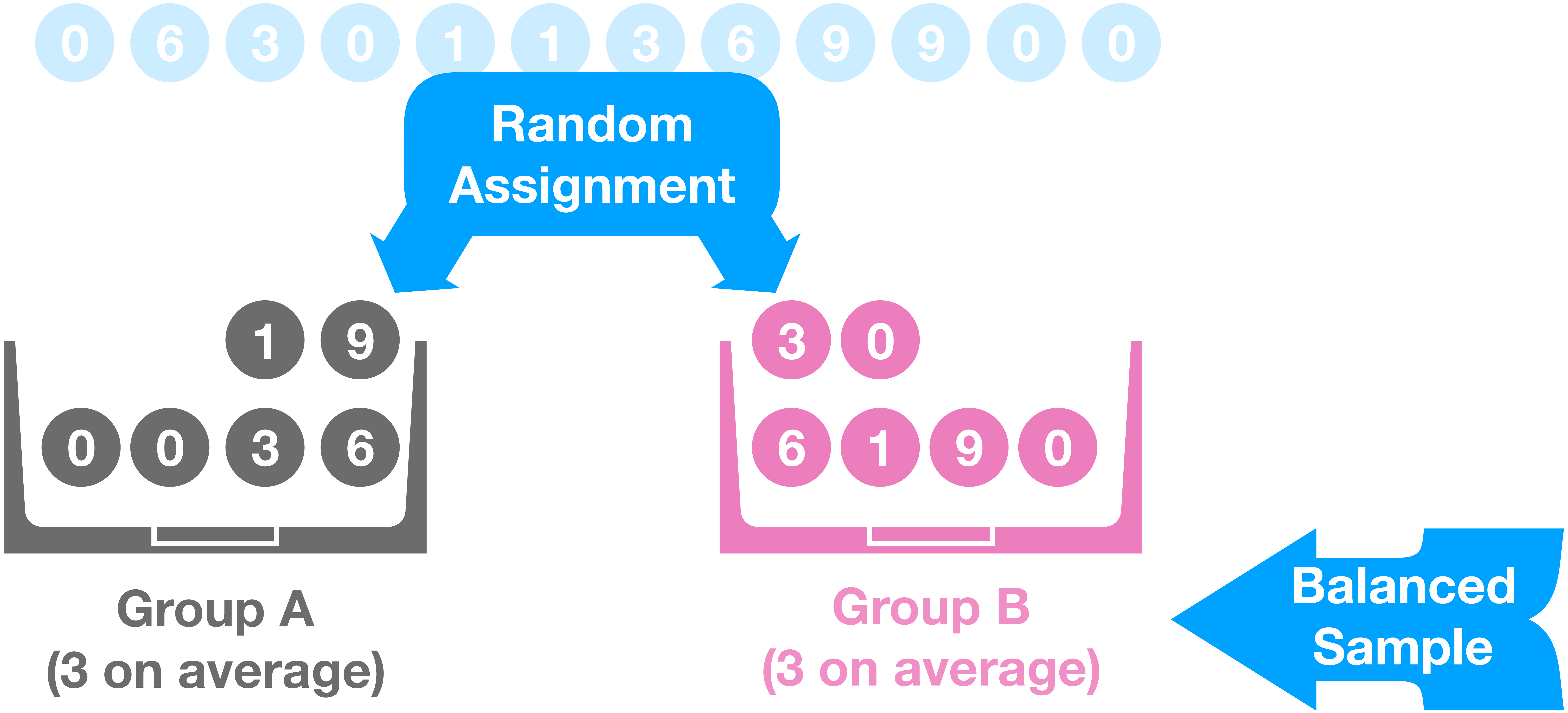
Now imagine that we allocate poems into groups based on a random sample

0 6 3 0 1 1 3 6 9 9 0 0

3 interactions on average

Understanding Randomization via Sampling

Now imagine that we allocate poems into groups based on a random sample



Understanding Randomization via Sampling

Now imagine that posting the colored background has an effect

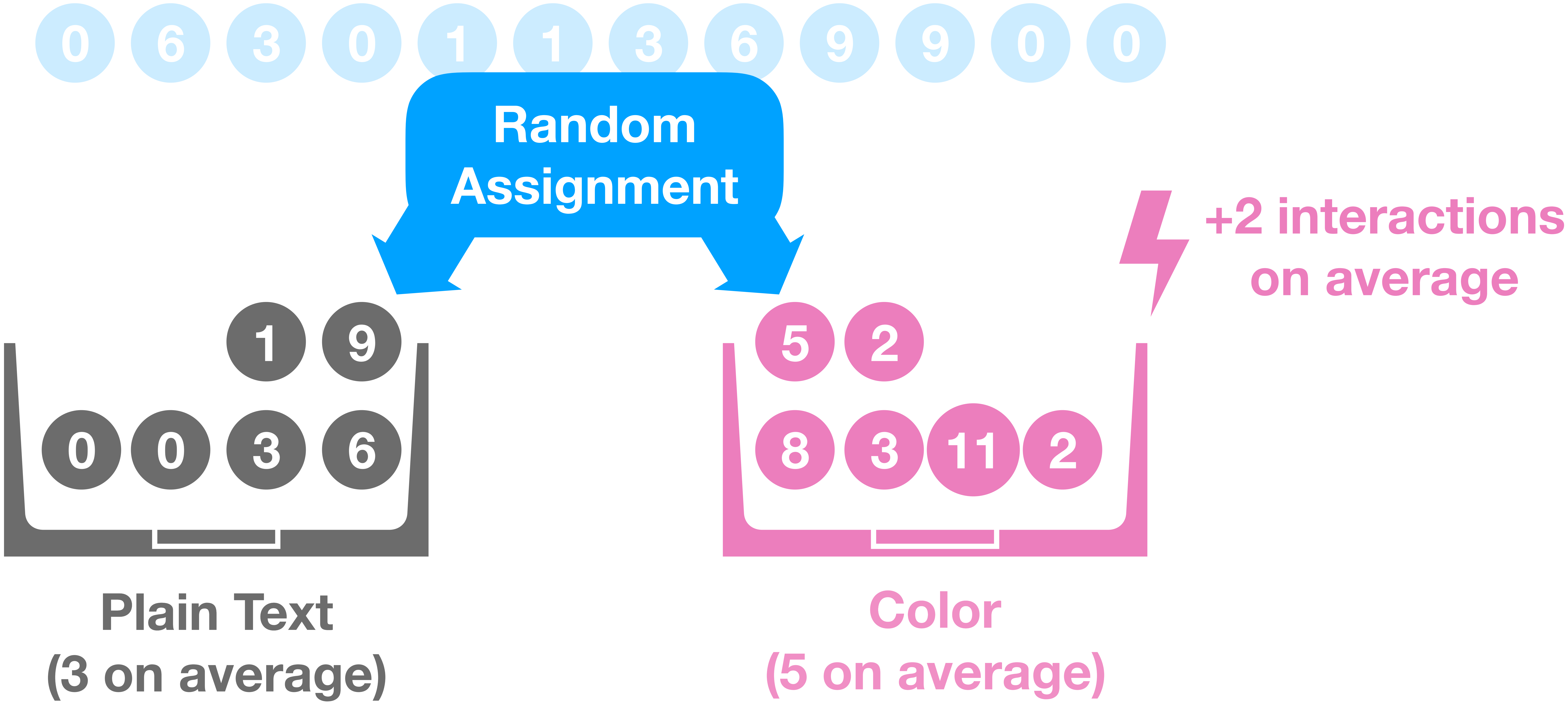
0 6 3 0 1 1 3 6 9 9 0 0

+2 interactions on average

Plain Text

Understanding Randomization via Sampling

Now imagine that posting the colored background has an effect

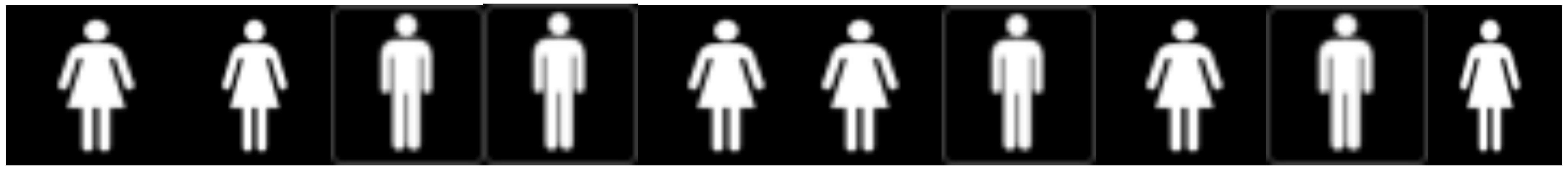


Common Methods of Random Assignment

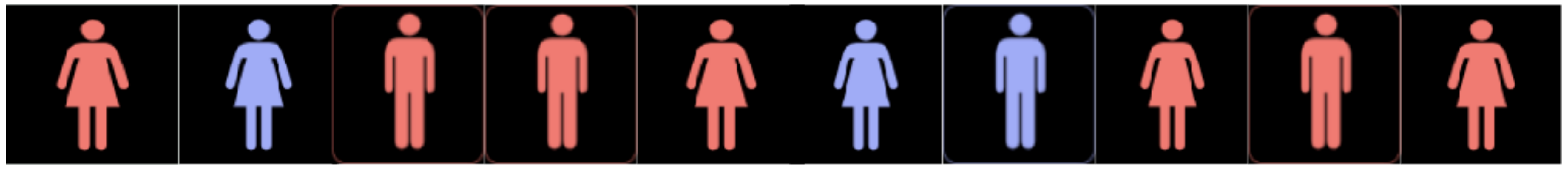
- **Simple randomization** (coin flips)
  - Problem: it's hard to get equal groups
- **Complete** (equal groups)
  - Example: sorted lists
- **Clustered**: (by group)
  - Randomized students by randomizing schools
- **Blocked**: (within groups)

Simple Randomization

Sample:



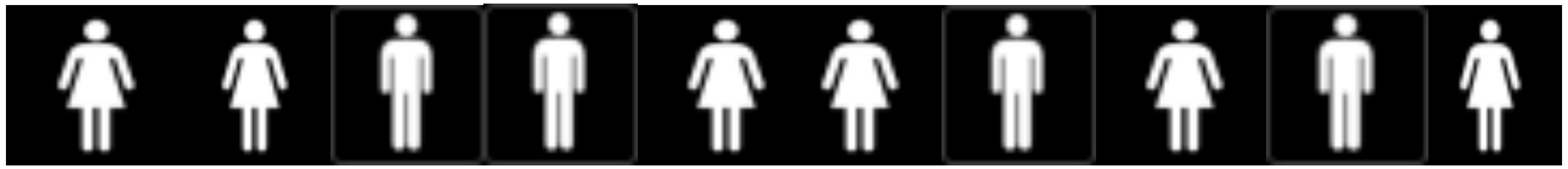
Iteration 1:



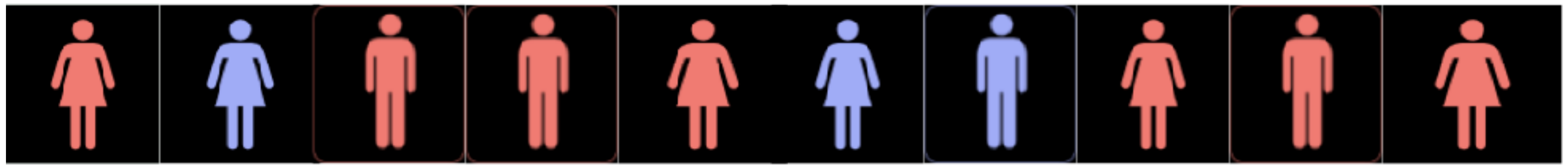
Simple Randomization

May not allocate the expected number of participants
to treatment and control

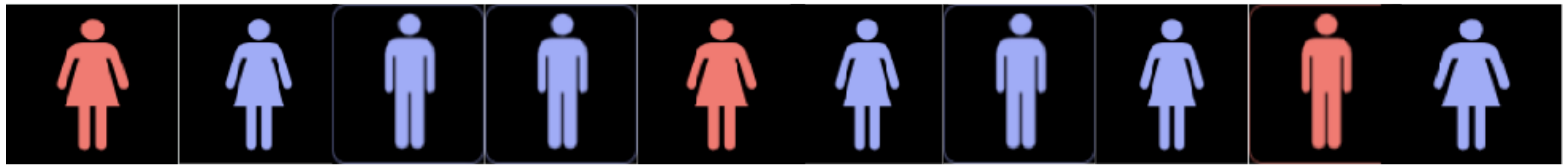
Sample:



Iteration 1:

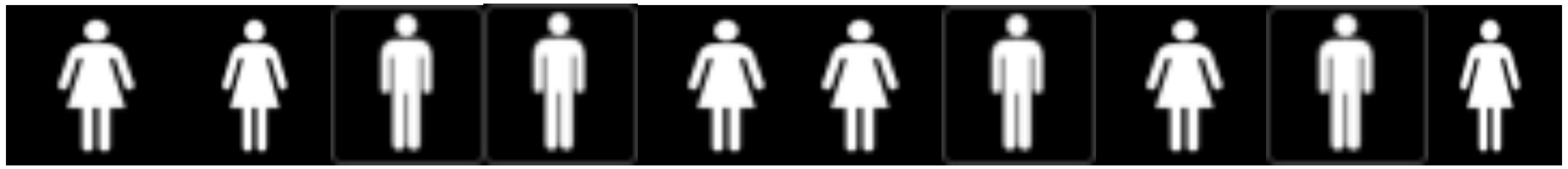


Iteration 2:

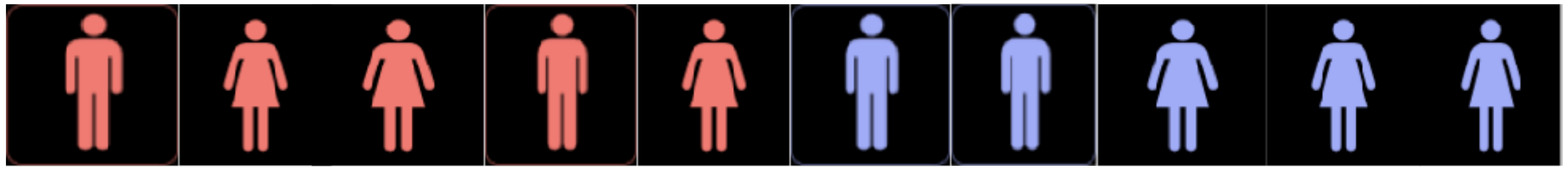


Complete Randomization

Sample:



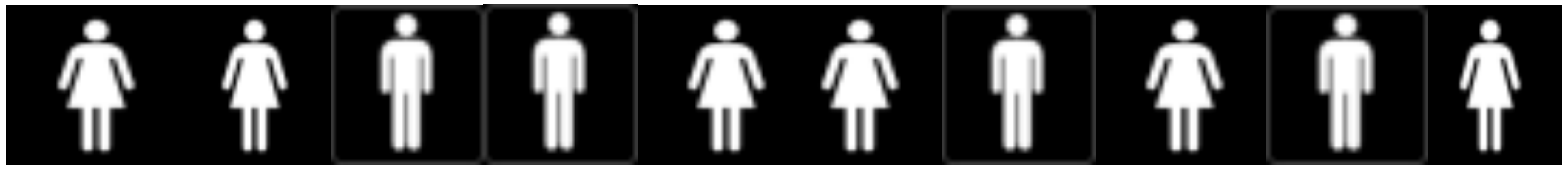
Iteration 1:



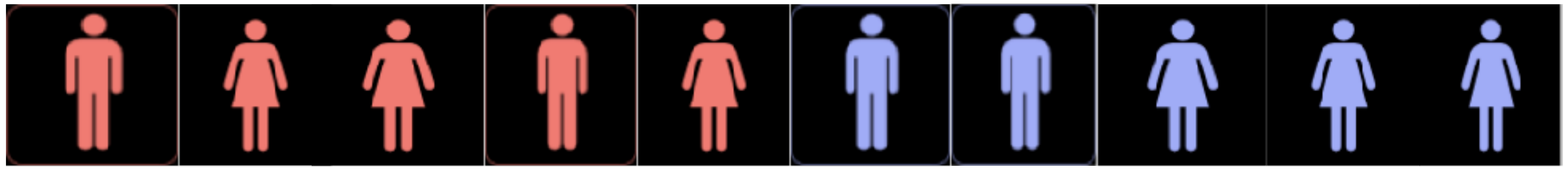
Complete Randomization

Always allocates the expected number of participants to treatment and control

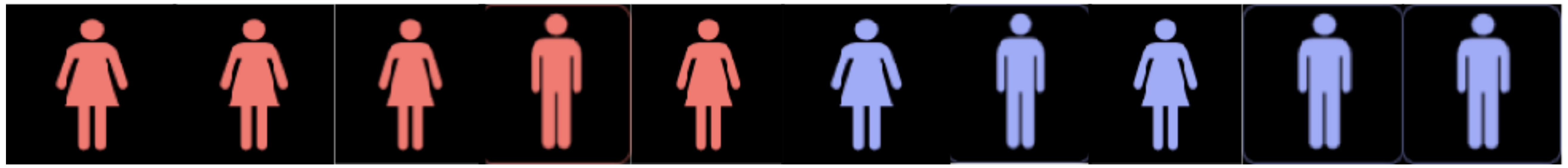
Sample:



Iteration 1:

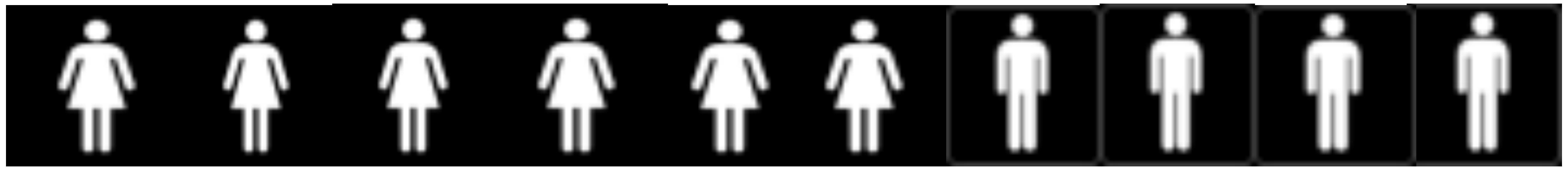


Iteration 2:



Blocked Randomization

Sample:

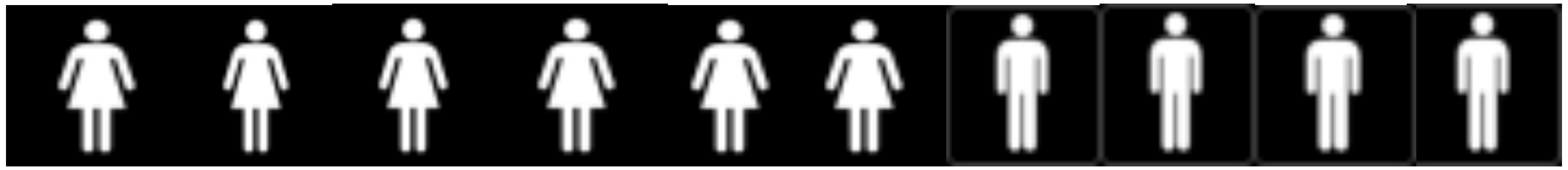


Iteration 1:

Blocked Randomization

Always maintains balance between characteristics (such as gender)

Sample:

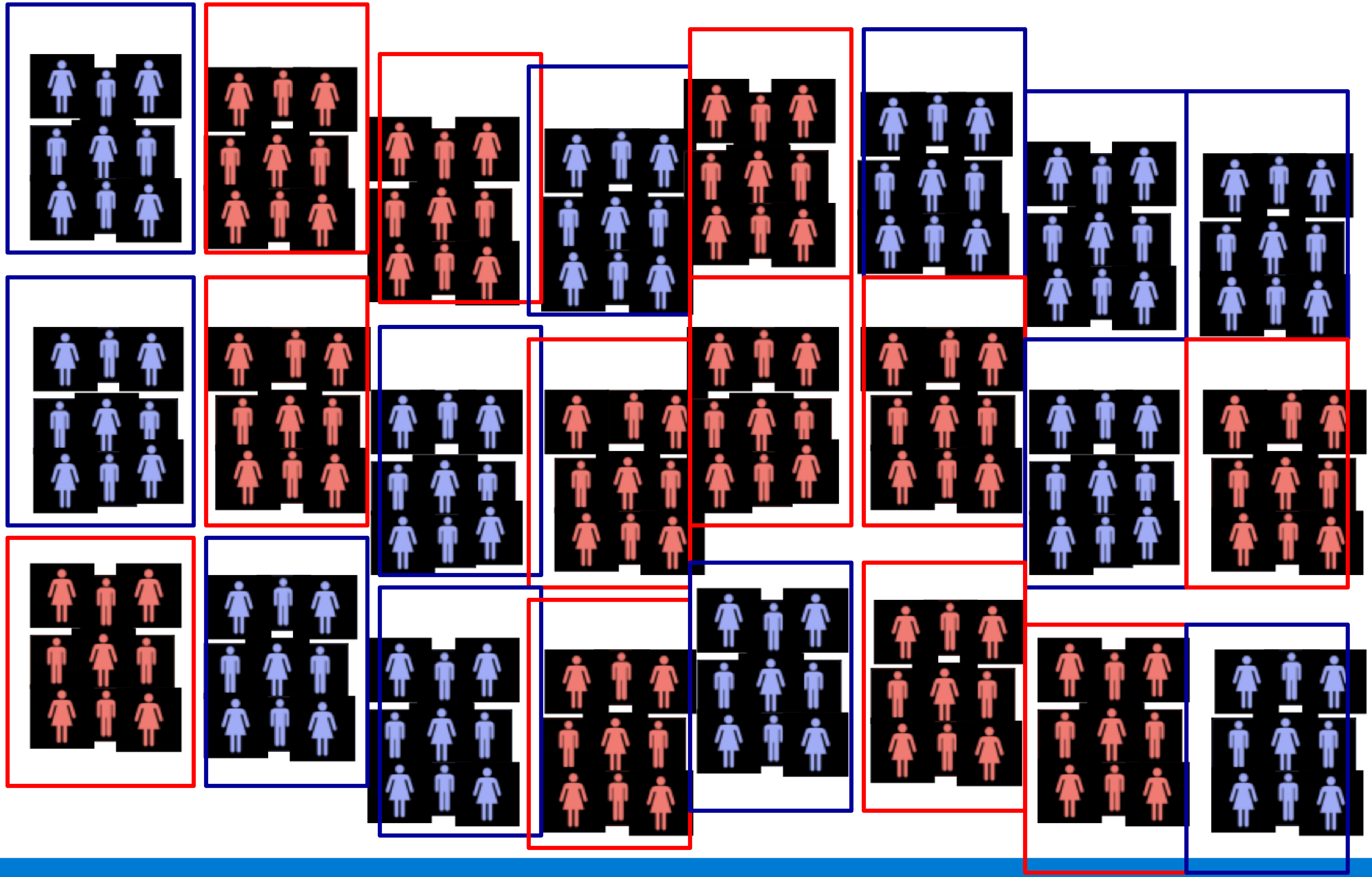


Iteration 1:

Iteration 2:

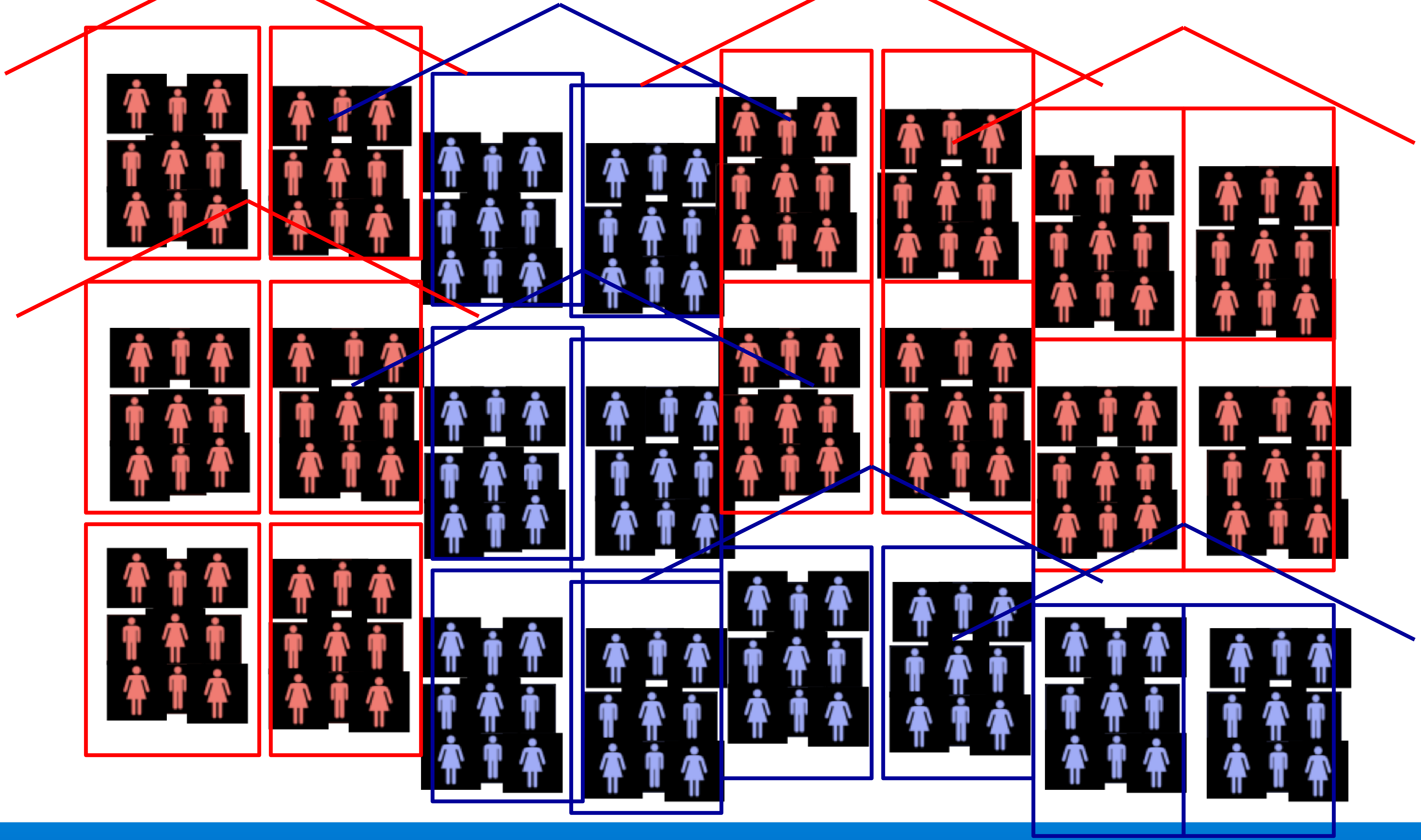
Clustered Randomization: Class

Clustered Randomization: Class

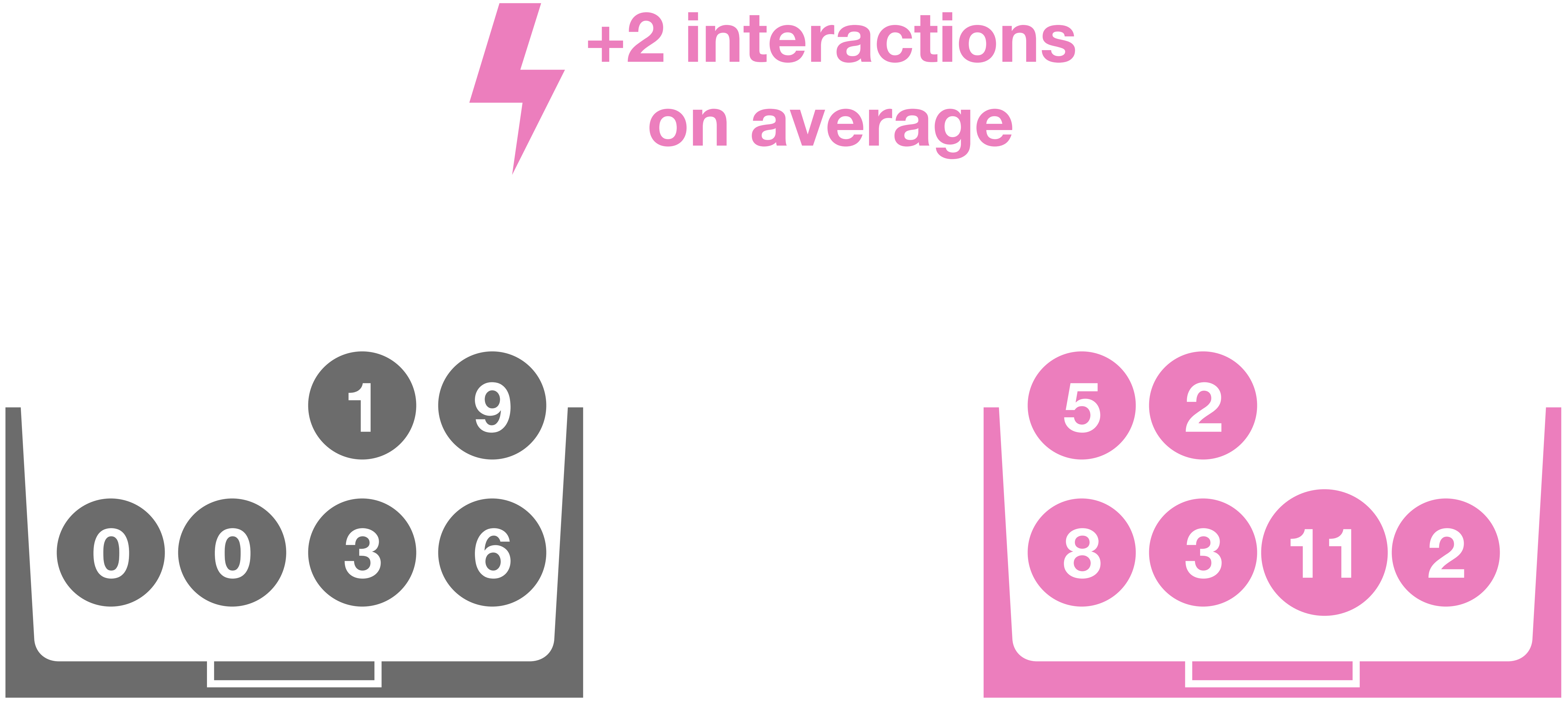


Clustered Randomization: School

Clustered Randomization: School



Average Treatment Effect (ATE)



Plain Text
(3 on average)

Color
(5 on average)

Average Treatment Effect (ATE)

$$Y = \alpha + \beta_1 X + \epsilon$$

$$Interactions = \alpha + \beta_1 Background + \epsilon$$

lm(interactions ~ condition, data=poems)

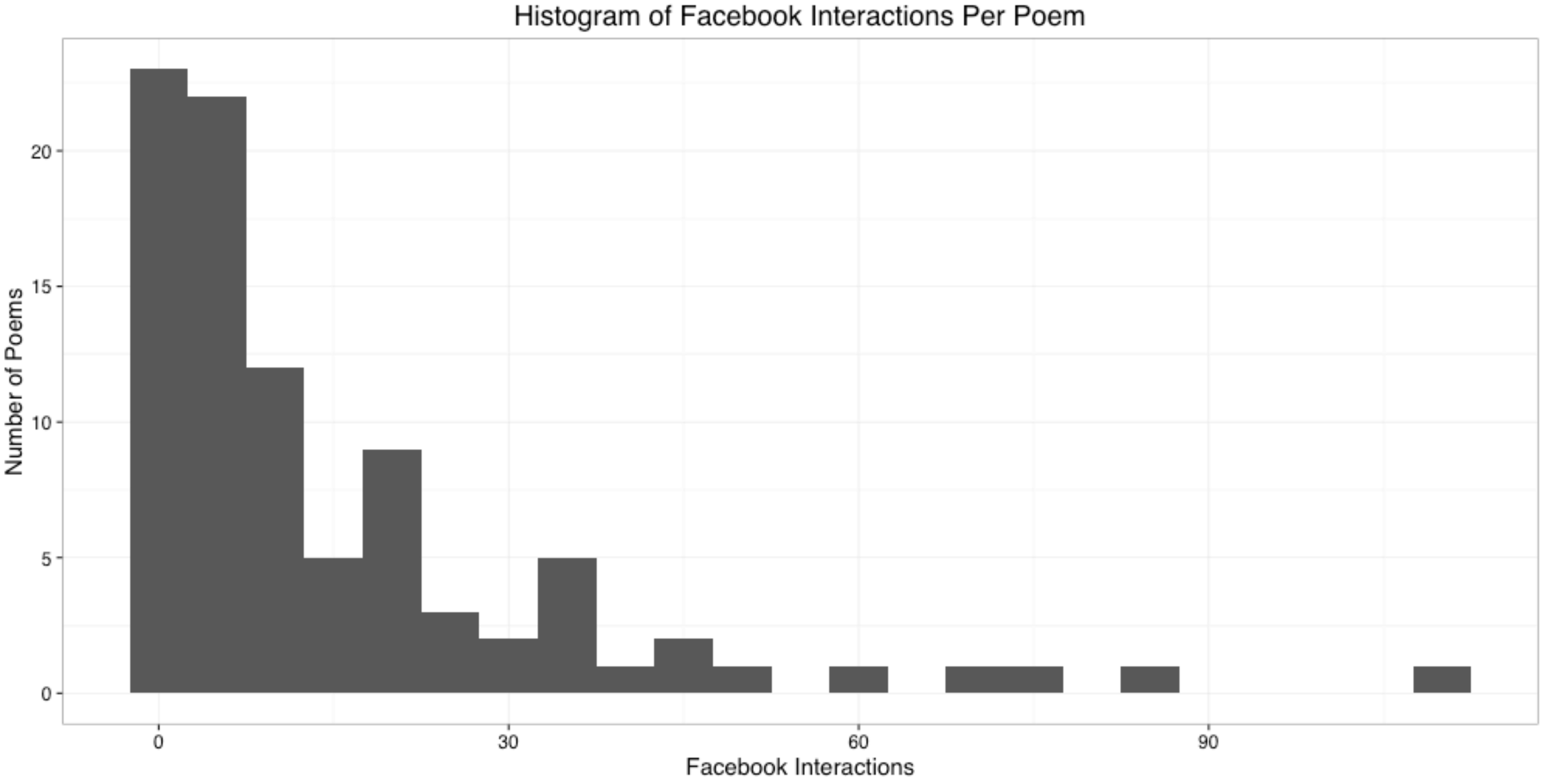
Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)

```
Call:
lm(formula = interactions ~ condition, data = poems)

Residuals:
    Min      1Q  Median      3Q     Max
-17.578 -11.578  -8.578   3.922  99.289

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)      11.711      3.008   3.893 0.000192 ***
conditionColor    7.867      4.254   1.849 0.067785 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.18 on 88 degrees of freedom
Multiple R-squared:  0.03741,	Adjusted R-squared:  0.02647
F-statistic:  3.42 on 1 and 88 DF,  p-value: 0.06778
```

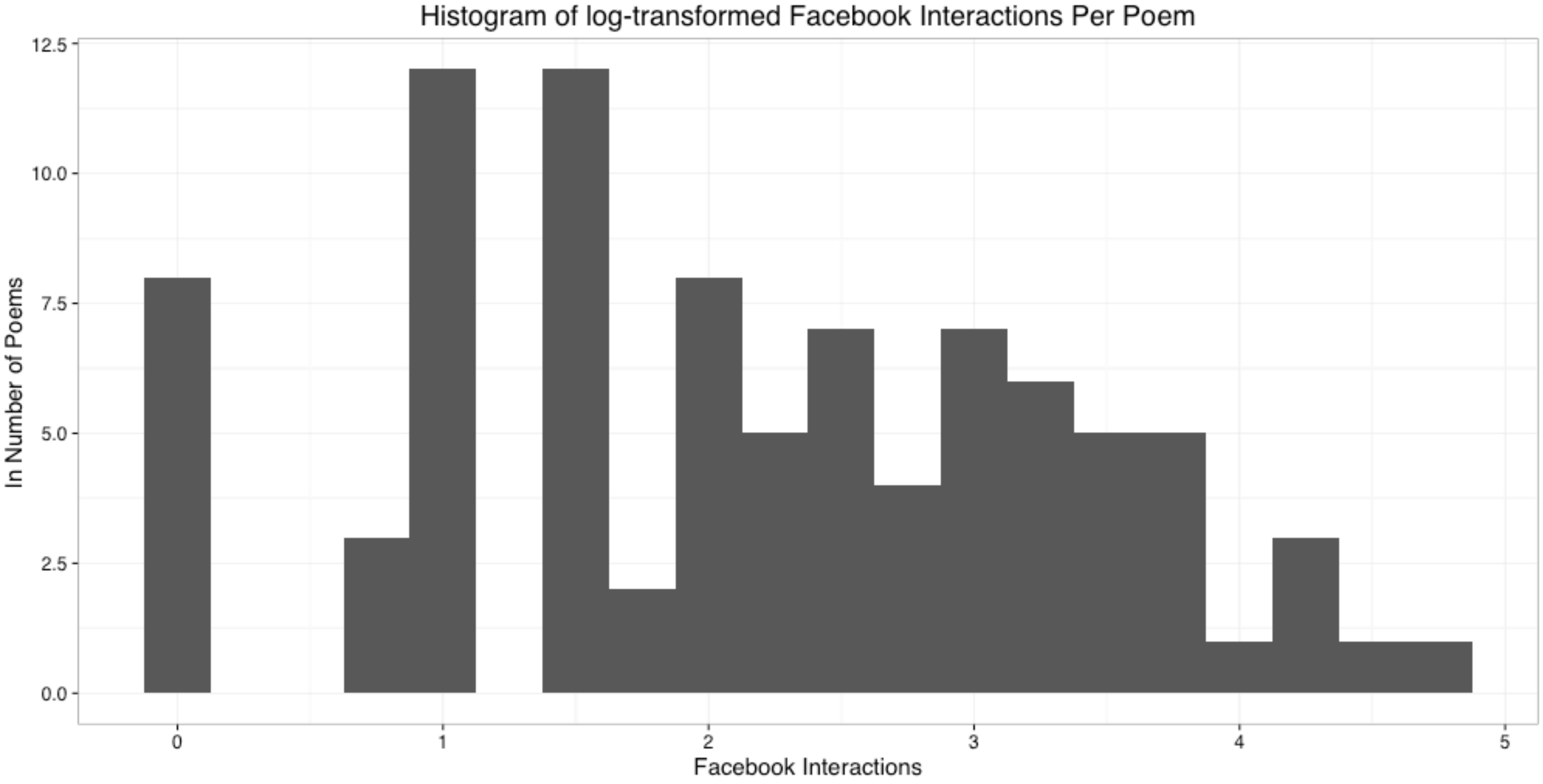
Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)



Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)



Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)

(log-transformed dependent variable)

$$\ln(Interactions + 1) = \alpha + \beta_1 Background + \epsilon$$

lm(*log1p*(interactions) ~ condition, data=poems)

Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)

(log-transformed dependent variable)

```
Call:
lm(formula = log1p(interactions) ~ condition, data = poems)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.6856 -0.5869 -0.1596  0.7456  3.0329 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)      1.6856     0.1635  10.309  < 2e-16 ***
conditionColor   0.9589     0.2312   4.147 7.74e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.097 on 88 degrees of freedom
Multiple R-squared:  0.1635, Adjusted R-squared:  0.154 
F-statistic:  17.2 on 1 and 88 DF,  p-value: 7.737e-05
```

Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)

(log-transformed dependent variable)

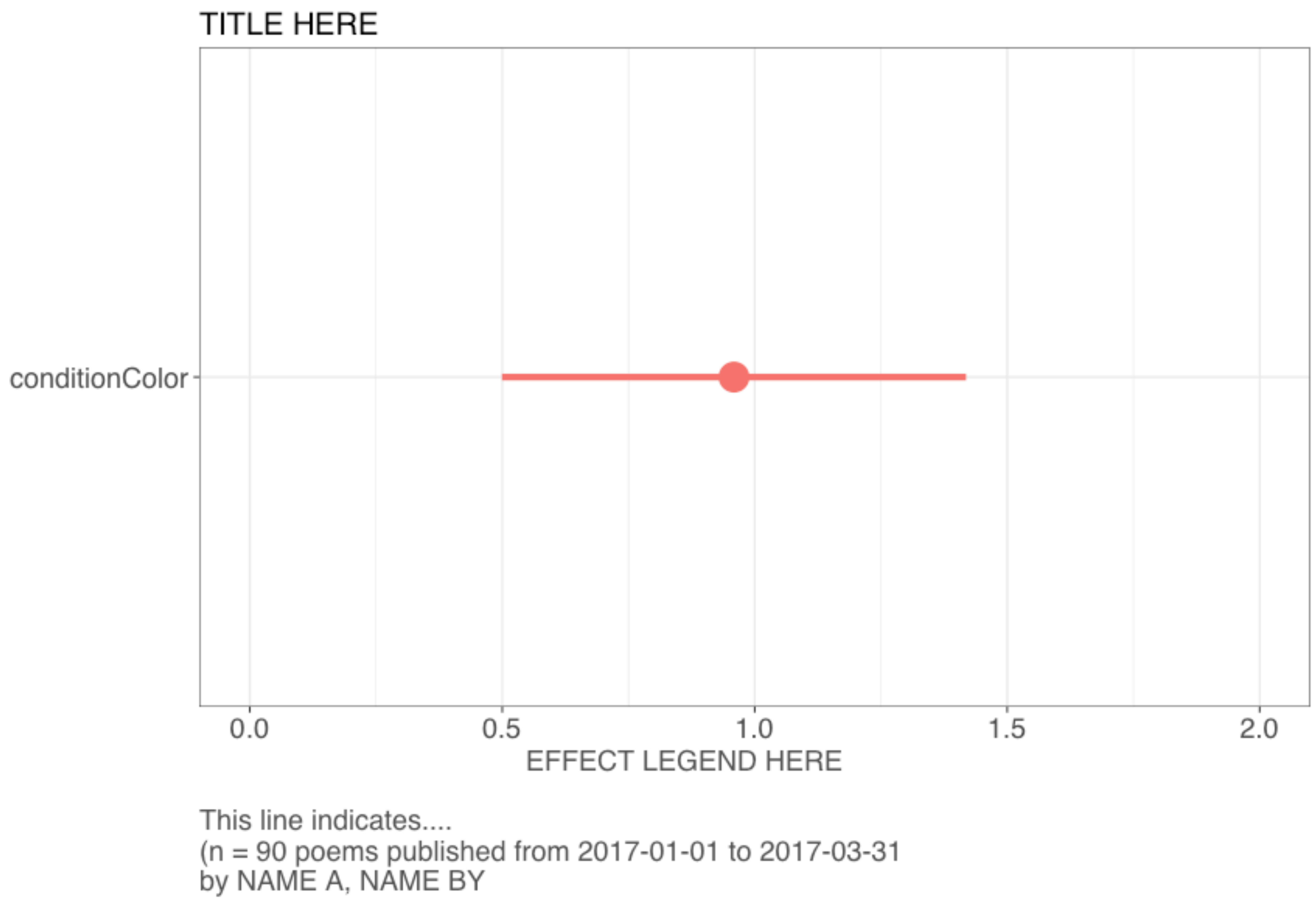
	Linear	Log-Transformed
Color	**7.87**	**0.96 *** **
	(4.25)	(0.23)
(Intercept)	11.71 ***	1.69 ***
	(3.01)	(0.16)
R^2	**0.04**	**0.16**
Num. obs.	90	90
RMSE	20.18	1.10

Linear models estimating log-transformed likes, comments, and shares per poem

Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

Average Treatment Effect (ATE)

(log-transformed dependent variable)

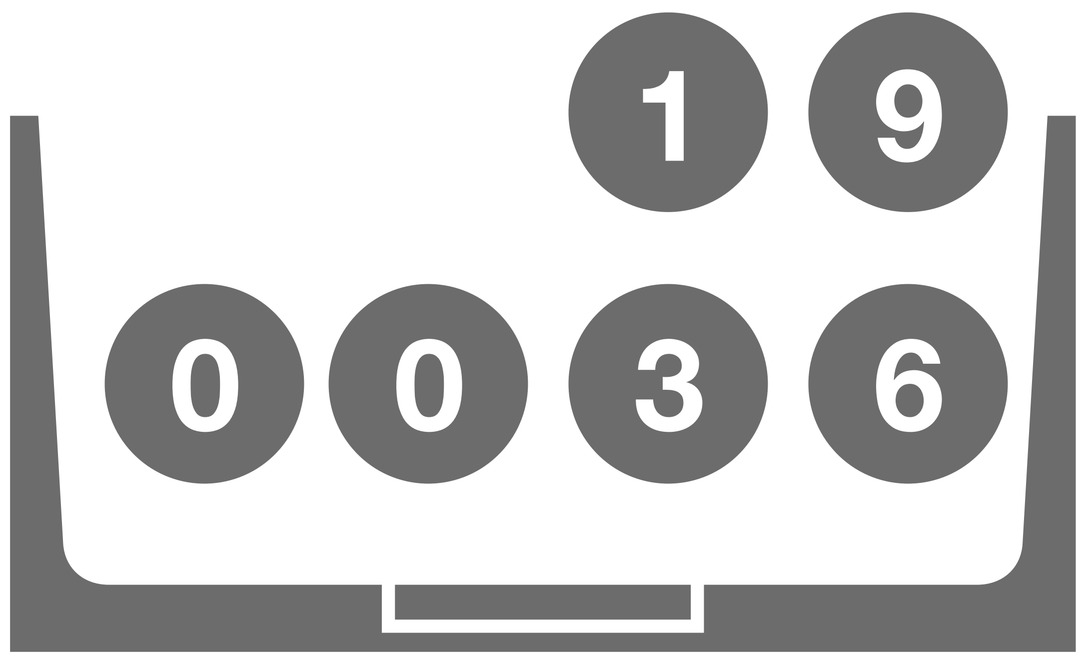


Example at github.com/natematias/SOC412/tree/master/1-facebook-poem

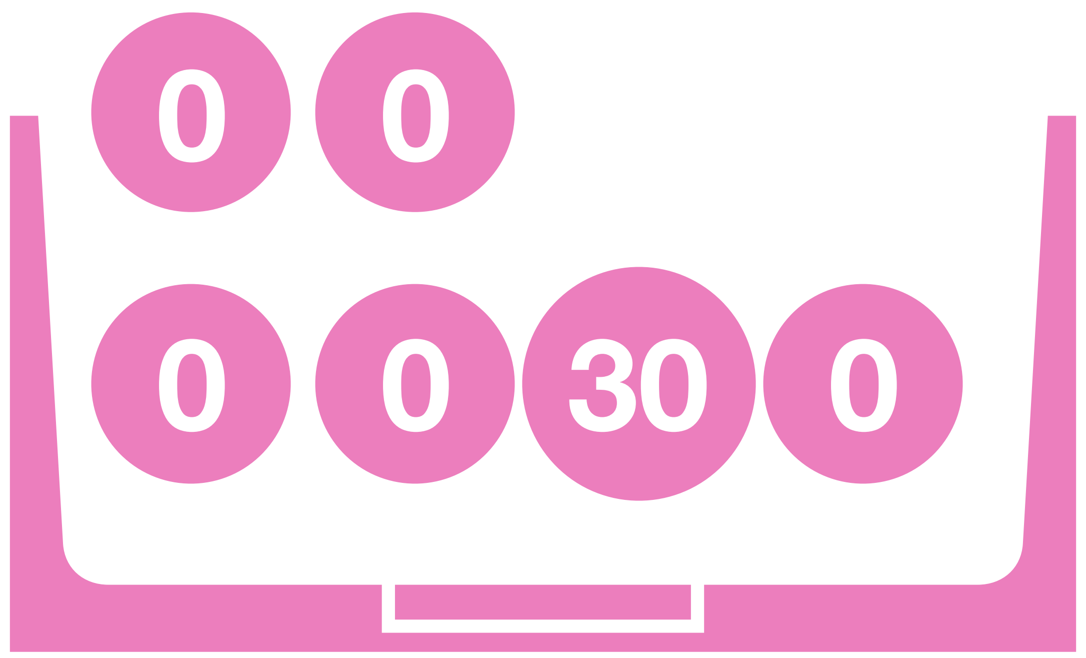
Average Treatment Effect (ATE)

APPLIES TO POPULATIONS **NOT TO INDIVIDUALS**

+2 interactions on average



Plain Text
(3 on average)



Color
(5 on average)

Assumptions of ATE

- **Random assignment** of participants to treatment
  - implies that receiving the treatment is statistically independent of participants' potential outcomes
- **Non-interference**: a participant's potential outcomes reflect only whether they receive the treatment themselves (not by others receiving it)

Assumptions of ATE

- **Excludability** a participant's potential outcomes respond only to the defined treatment, not other other extraneous factors that may be correlated with treatment
  - importance of defining the treatment precisely and maintaining symmetry between treatment and control groups (e.g. through blinding)

Conspicuously Absent Assumptions

- Random sampling of subjects from a larger population is not a core assumption
- The issue of "external validity" is a separate question that relates to the issue of whether the results obtained from a given experiment apply to other subjects, treatments, contexts, and outcomes
- For now, we aim only to estimate the ATE in our subject pool