

How Experiments Work

COMM 4940

Kennedy Hall 213

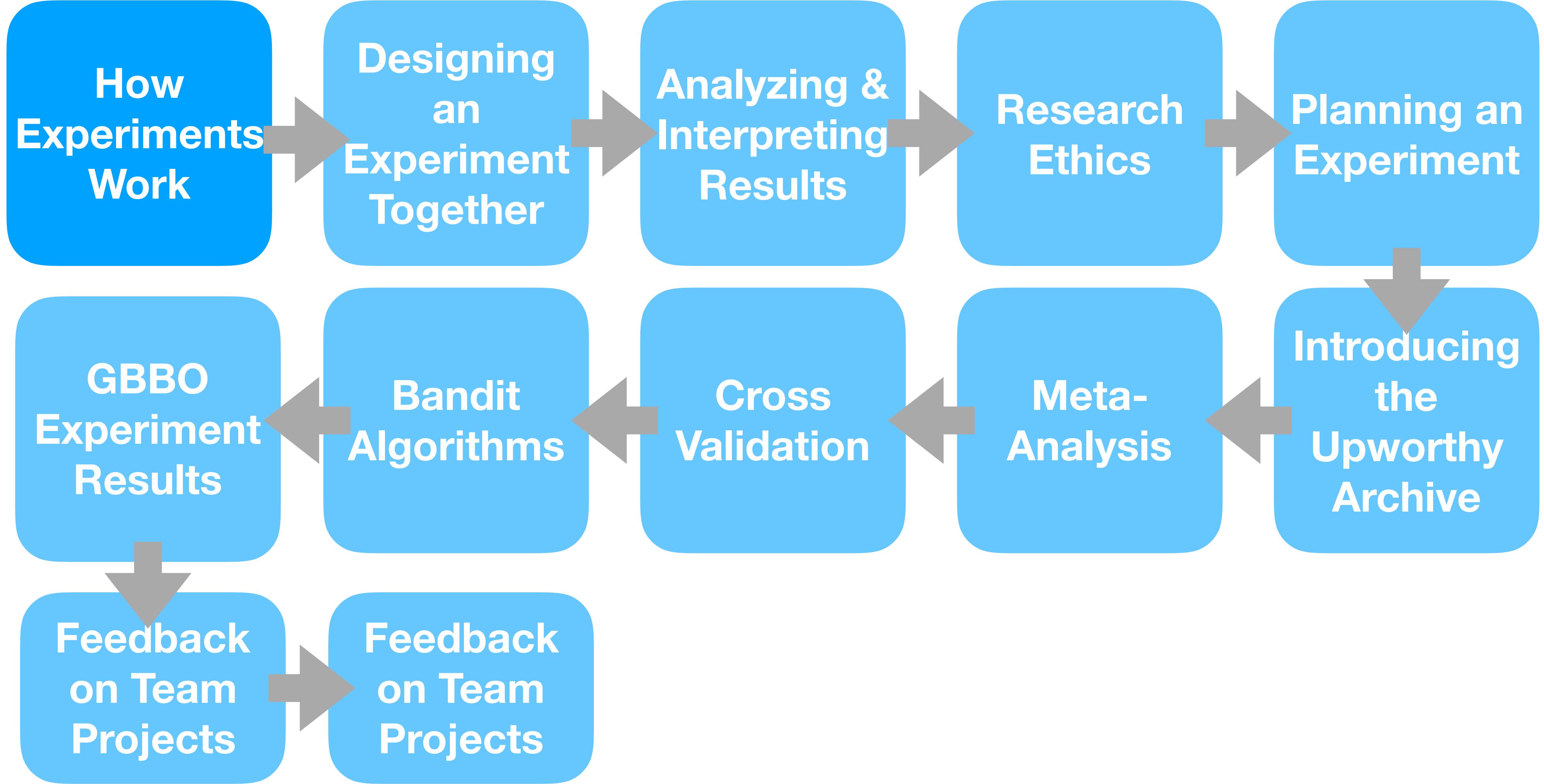


J. Nathan Matias
@natematias
citizensandtech.org
natematias.com



Goals for Today

- Review how experiments work
- Introduction to CoCalc
- Discuss First Assignment



06.07.15

How Upworthy Used Emotional Data To Become The Fastest Growing Media Site of All Time

If it makes an editor cry, it's Upworthy.





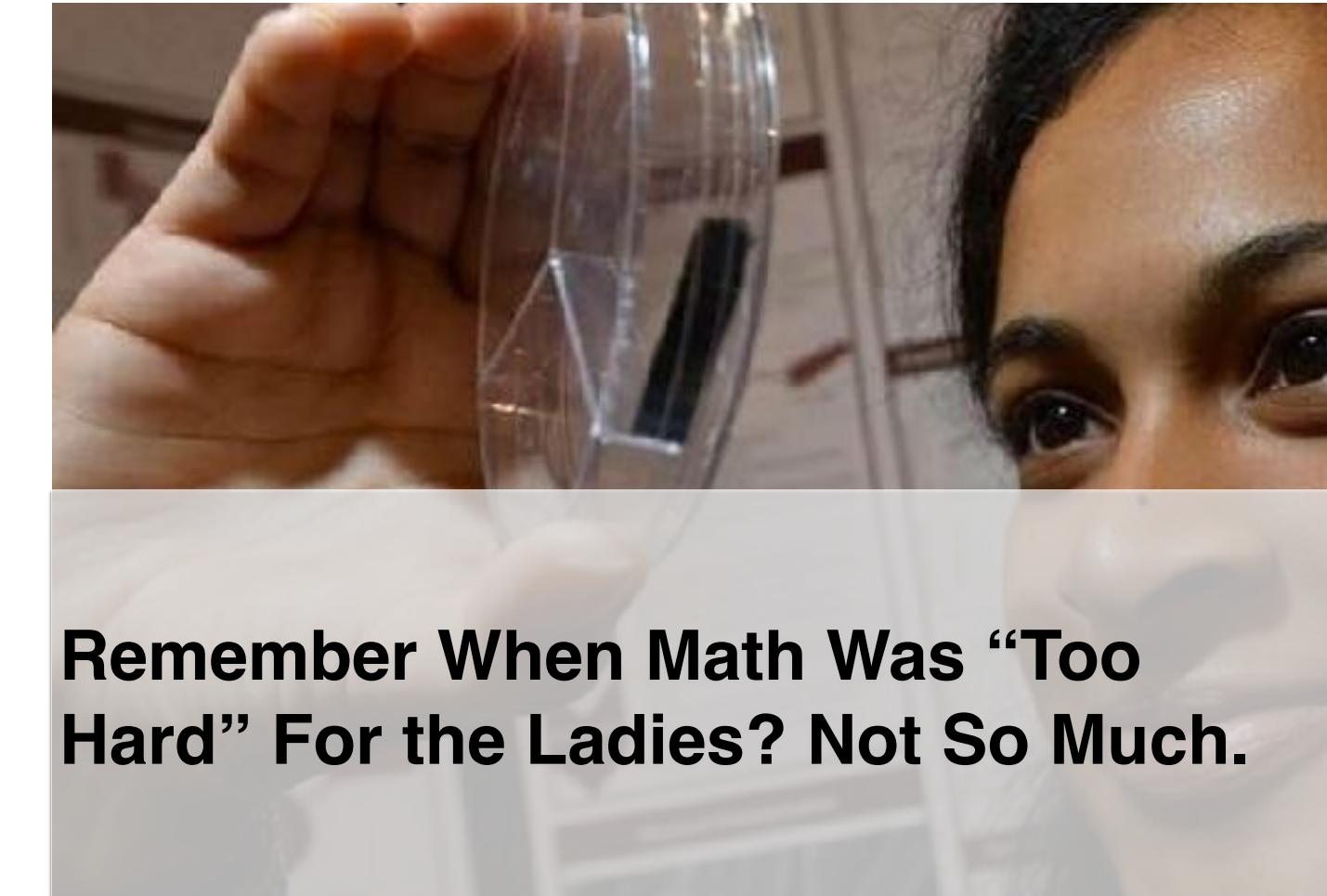
How Should Upworthy Amplify Media?

A



**She's Not Just Destined For
Greatness, She's Destined To Do
Great Things For Women**

B



**Remember When Math Was “Too
Hard” For the Ladies? Not So Much.**

Group Identity?

Overcoming stereotypes?

Reviewing Historical Data

group	headline	impressions	clicks	click-through-rate
Group	Local Witch Stands up for Muggles	938	6	0.6%
Group	Jeremy Clarkson Smashes Barriers for Auto Drivers	950	11	1.2%
Group	This Courtroom Speech is Changing How the World Understands Bakers	1097	9	0.8%
Group	Teenage Poet Reminds us what it means to be American	995	14	1.4%
Group	Vegetarians Celebrate Olympics Champion's Diet	1011	3	0.3%
Stereotype	Area Toddler wins New York Mathematics Competition	1168	26	2.2%
Stereotype	Tech Bro Wins Beauty Pageant	915	30	3.3%
Stereotype	This Woman Broke the Glass Ceiling in Journalism	1054	18	1.7%
Stereotype	Football Player Wins Neuroscience Award	992	10	1.0%
Stereotype	Mathematicians Sweep National Wrestling Competition	1056	19	1.8%

Reviewing Historical Data

```
1 aggregate(click.through.rate ~ group, data = headlines, mean)
```

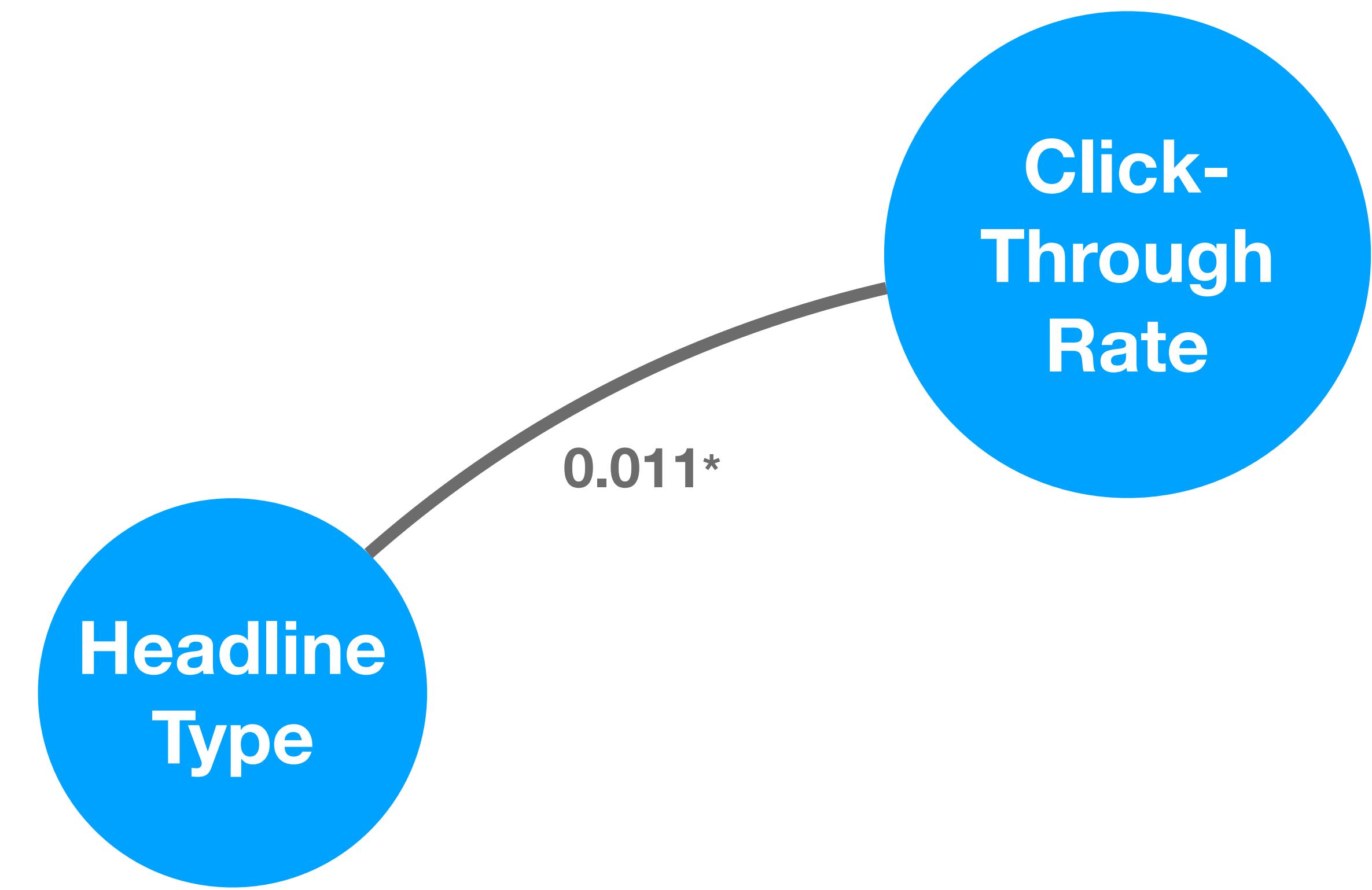
group	click.through.rate
Group	0.008643488
Stereotype	0.020039605

Reviewing Historical Data

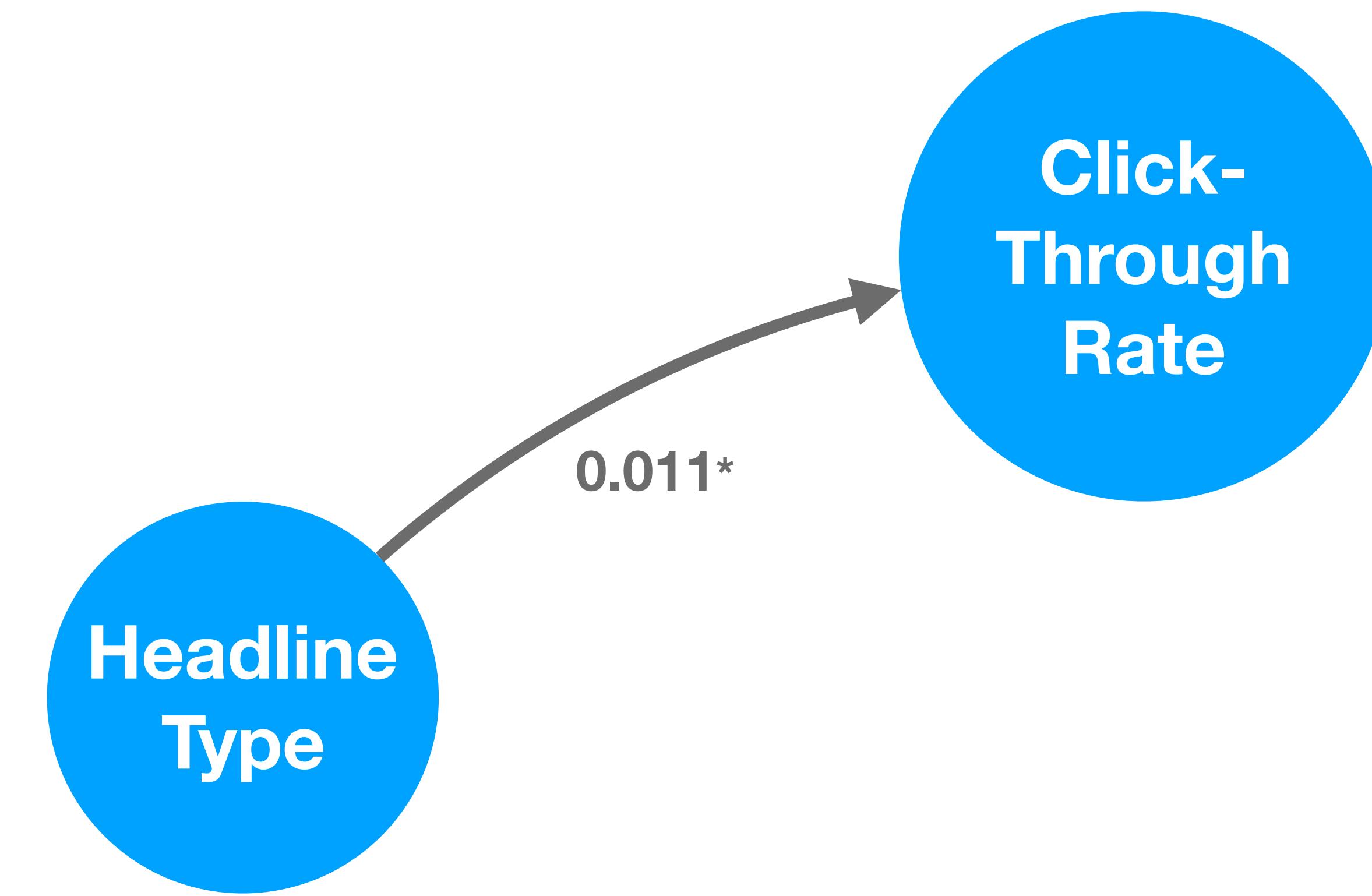
Welch Two Sample t-test

```
data: headlines$click.through.rate by headlines$group
t = -2.7046, df = 6.0143, p-value = 0.03528
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.021700620 -0.001091616
sample estimates:
mean in group Group mean in group Stereotype
0.008643488          0.020039605
```

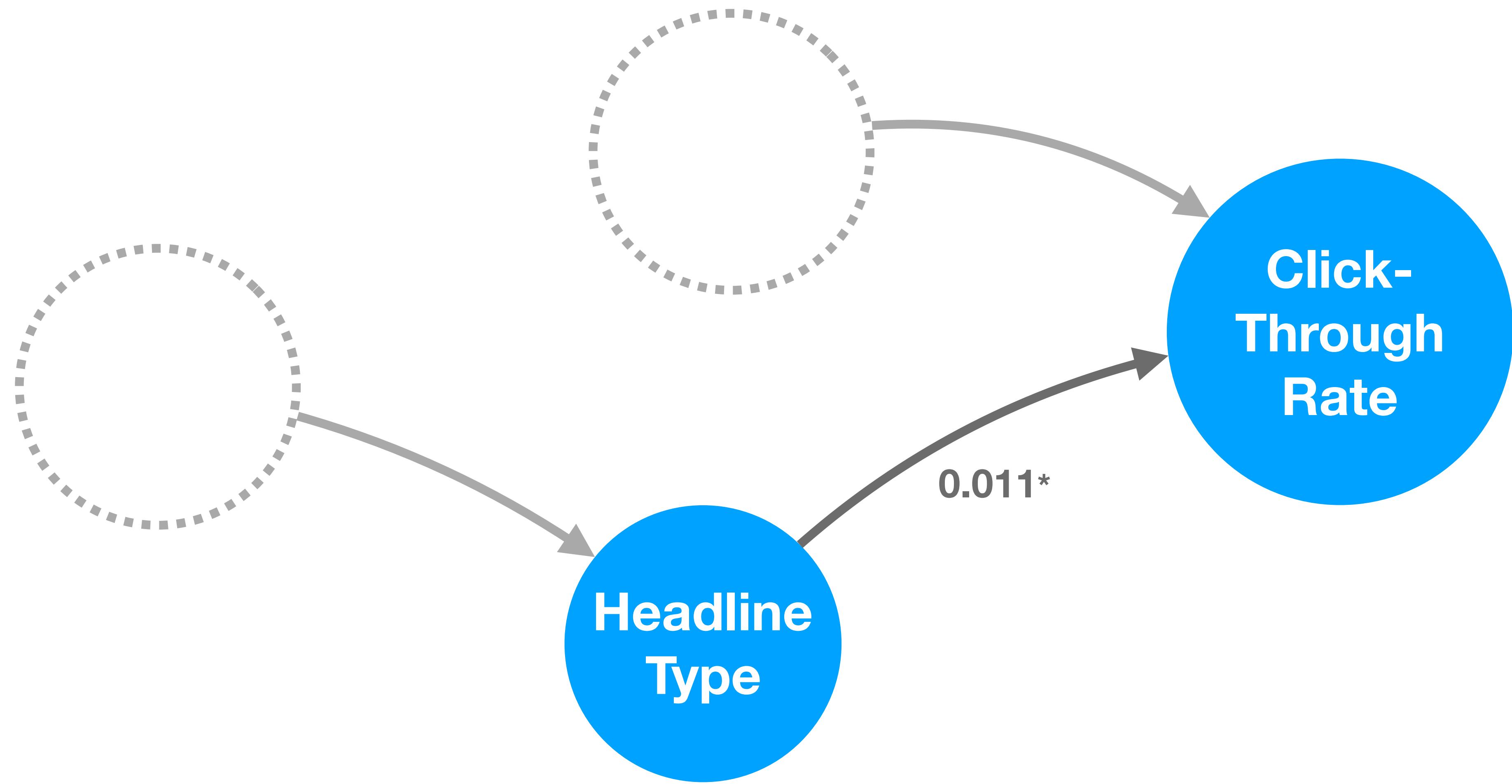
Path Diagram



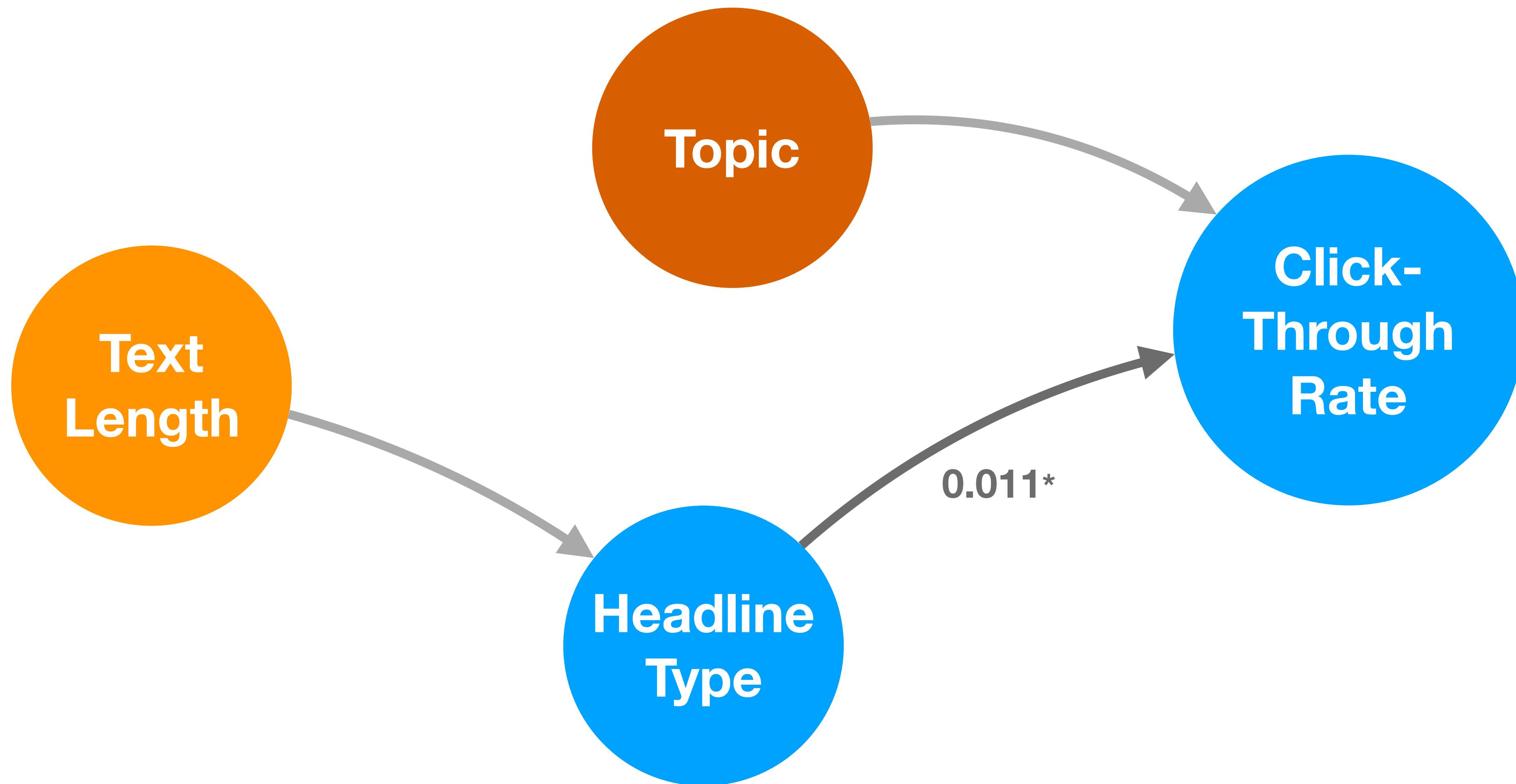
What do we mean by Correlation != Causation?



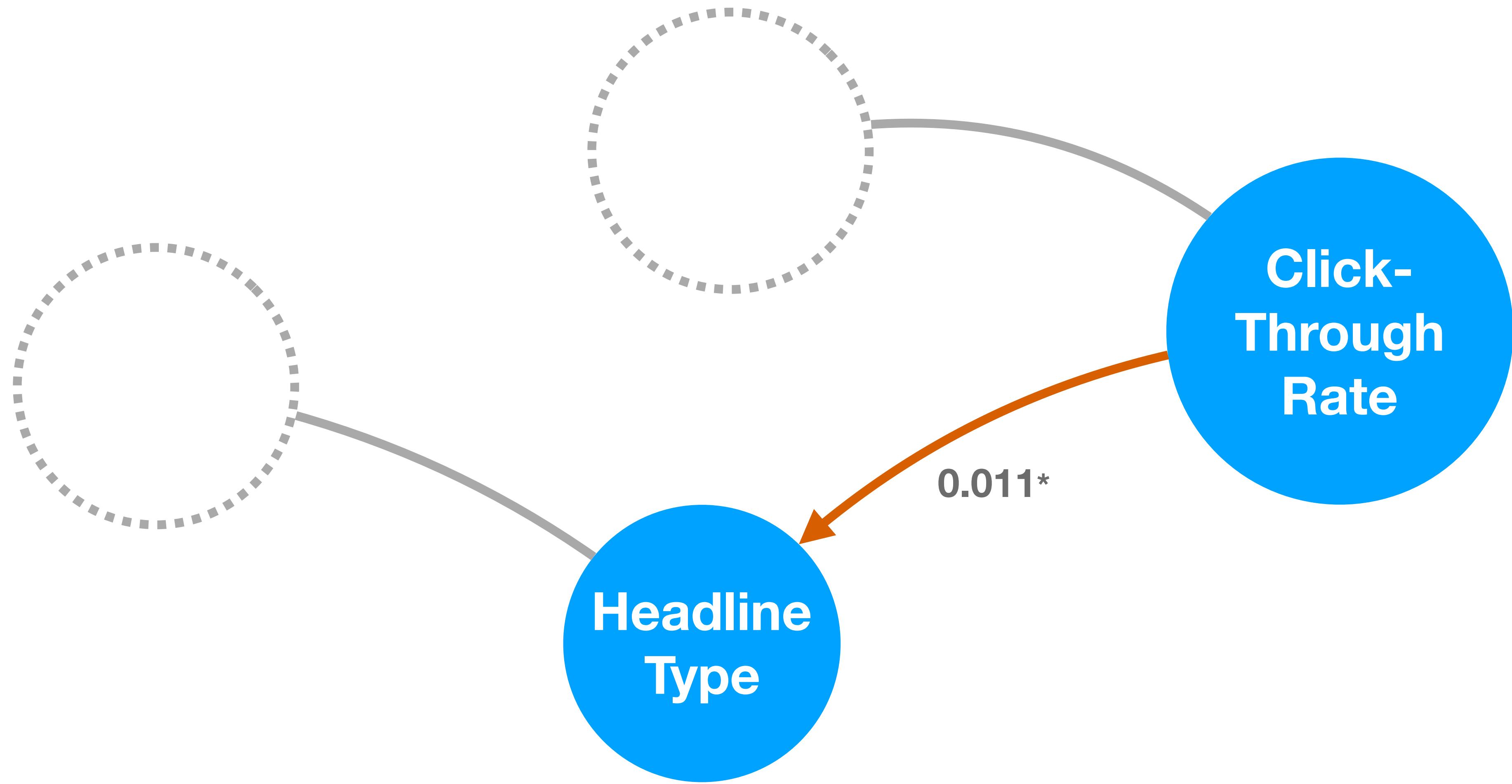
Unobserved Confounders



Confounders (?) Topic & Text Length



Direction of Causality



Why Experiment?

- Causal explanation and 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

worthy.com/ | Go OCT DEC JAN
5 2011 01 2012 2013 2014 About

UPWORTHY SHARE BETTER

Hi!

We're Upworthy.

We're a social media outfit with a mission: **To make meaningful stories go viral.**

Here's *AllThingsD*'s take on our first 6 months, and a *New York Times* piece on our first 100 days.

JOIN US, AND NEVER MISS OUR BEST CONTENT:

[Facebook](#) [Email List](#) [Tumblr](#)

Email Follow

Featured Stories

 **Open Your Mouth Now - Your Jaw Is About To Drop**
 **Femi Oke** Nov 19 2012
Introducing the \$20 bike. Seriously!
[Let's see it »](#)

 **Open Your Mouth Now - Your Jaw Is About To Drop**
 This African Teenager Taught Himself...
 **WATCH: Girl Tells Boy She Has HIV. Boy's Reaction...**

.upworthy.com/ | Go FEB APR MAY
ures Oct 2019 01 2013 2014 2015 About

UPWORTHY

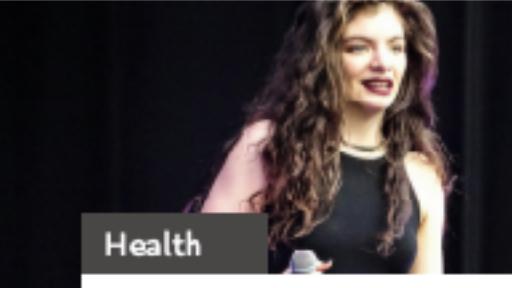
Search Stories

Things that matter. Pass 'em on.

 **A Woman Shares Her Day Through Google Glass. It Seems Lovely And Ordinary — Until The End**

 **Lorde Found A Photoshopped Photo Of Herself And Decided To Tweet An Untouched Real One Next To It**

 **A Guy Sells Something So Shocking On Wall Street That A Crowd Stops And Stares**

 **Community**
 **Entertainment & Culture**
 **Health**

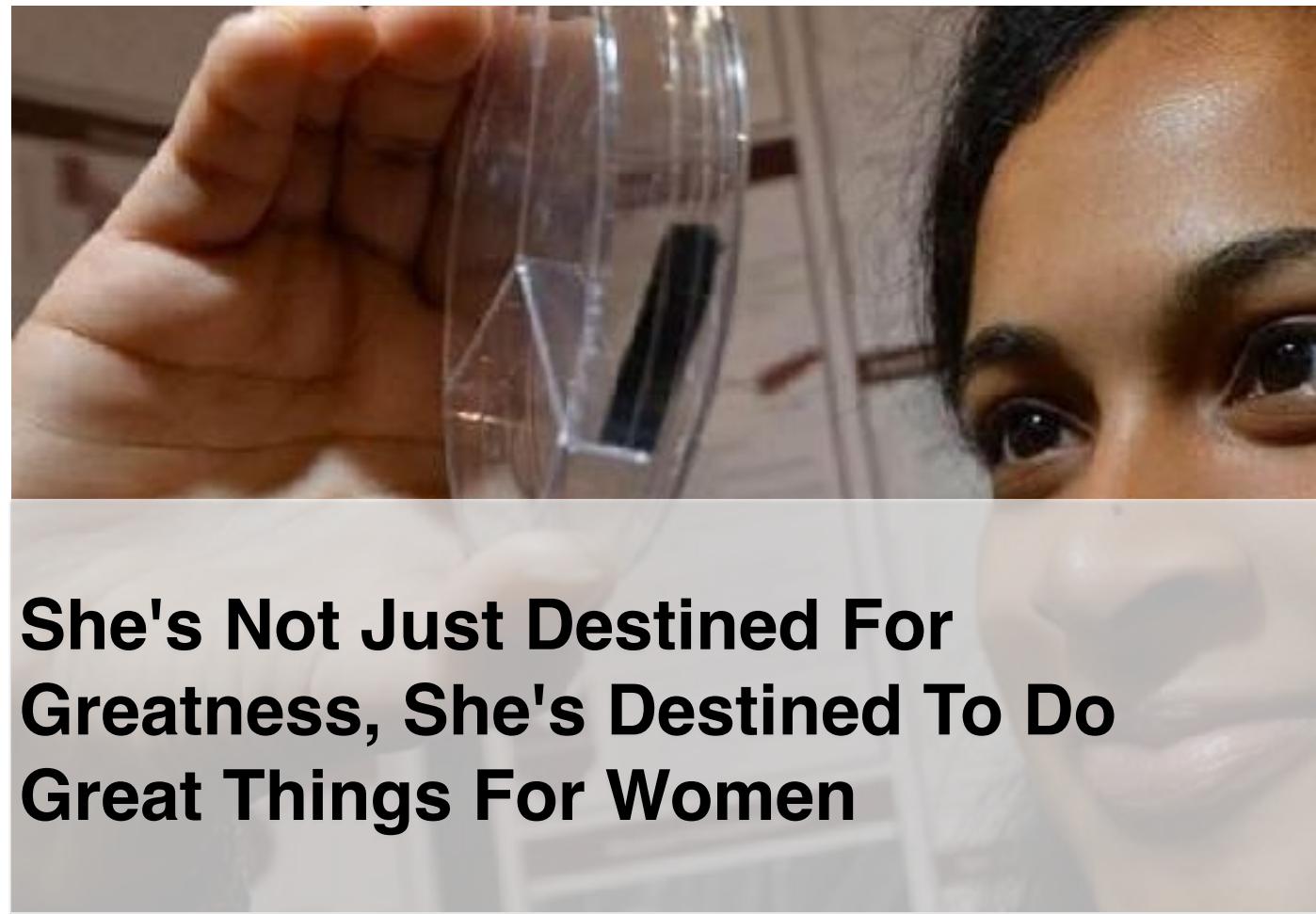
A Supersized Art Project Shows Us Something Tragic That Always Happens But We Rarely Ever See
By Carolyn Silveira

One Group Had A Hypothesis About Sexism. They Made A Bunch Of Pies To Prove It.
By Carolyn Silveira

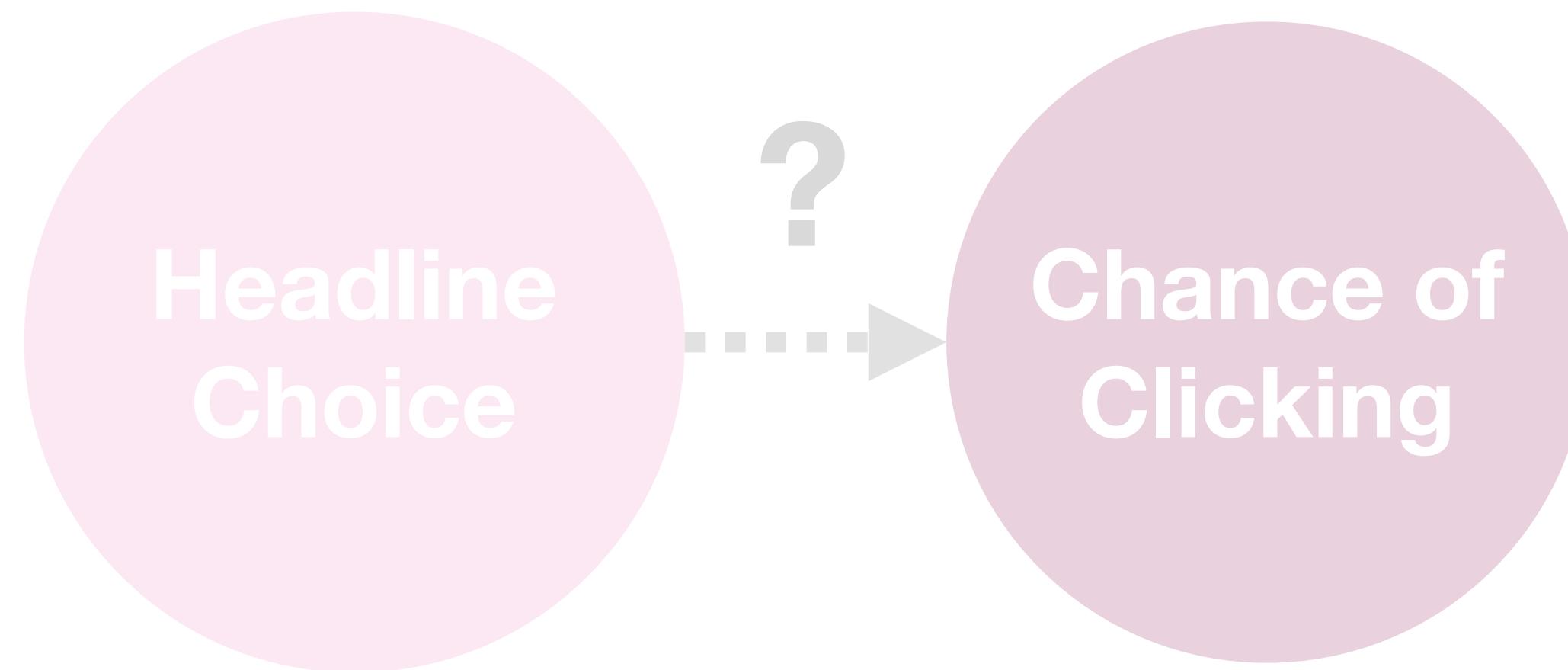
Lorde Found A Photoshopped Photo Of Herself And Decided To Tweet An Untouched Real One Next To It
By Rossalyn Warren

Parts of an Experiment (example)

Upworthy wants to reach **the greatest number of people** with the inspiring story of a teenager who won a major engineering competition



Parts of an Experiment



1. Are people more **likely to click** Headlines on **Group Identity** than **Overcoming Stereotypes** on avg?
2. If so, **by how much** on average?



She's Not Just Destined For Greatness, She's Destined To Do Great Things For Women

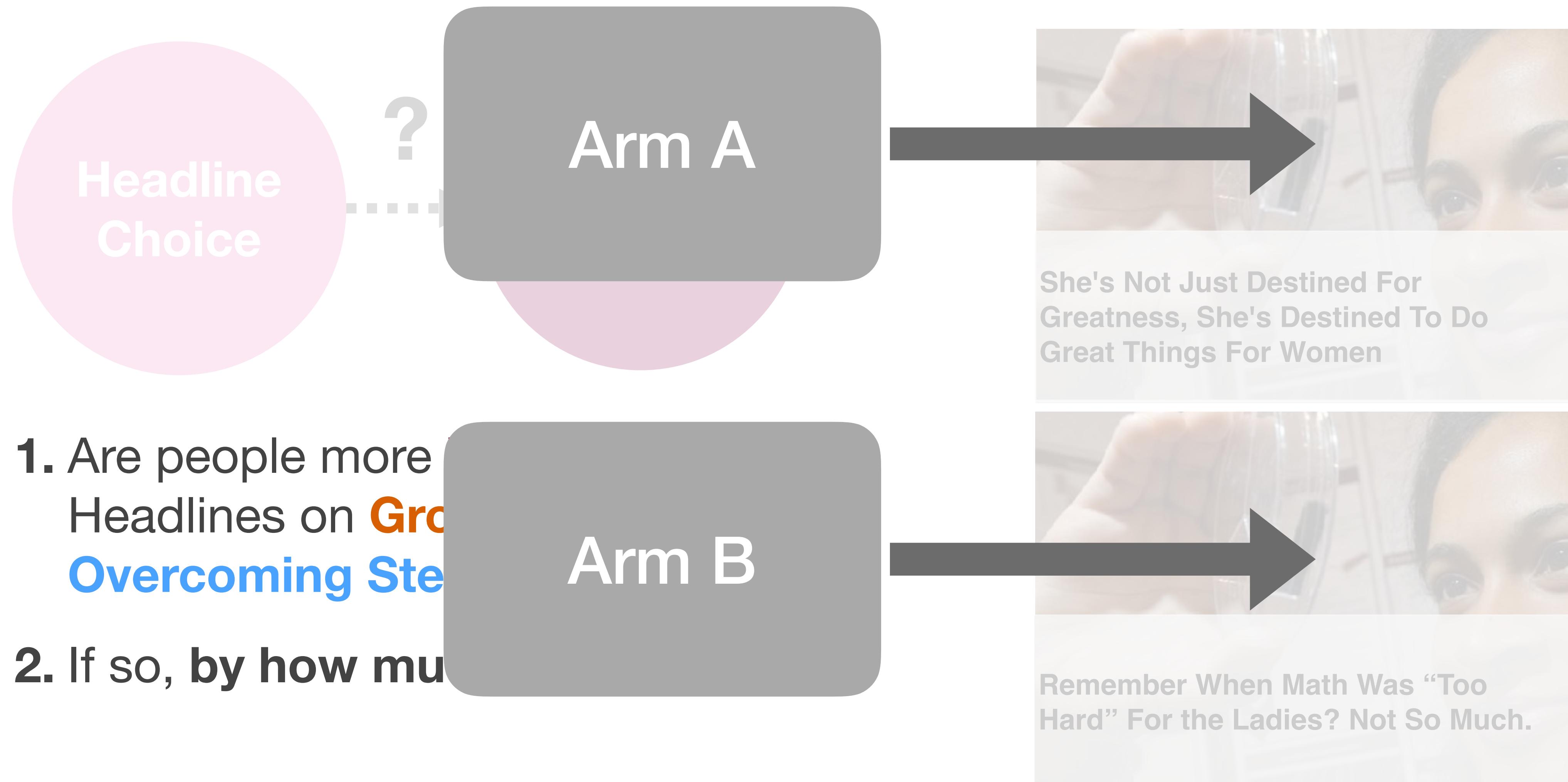
A



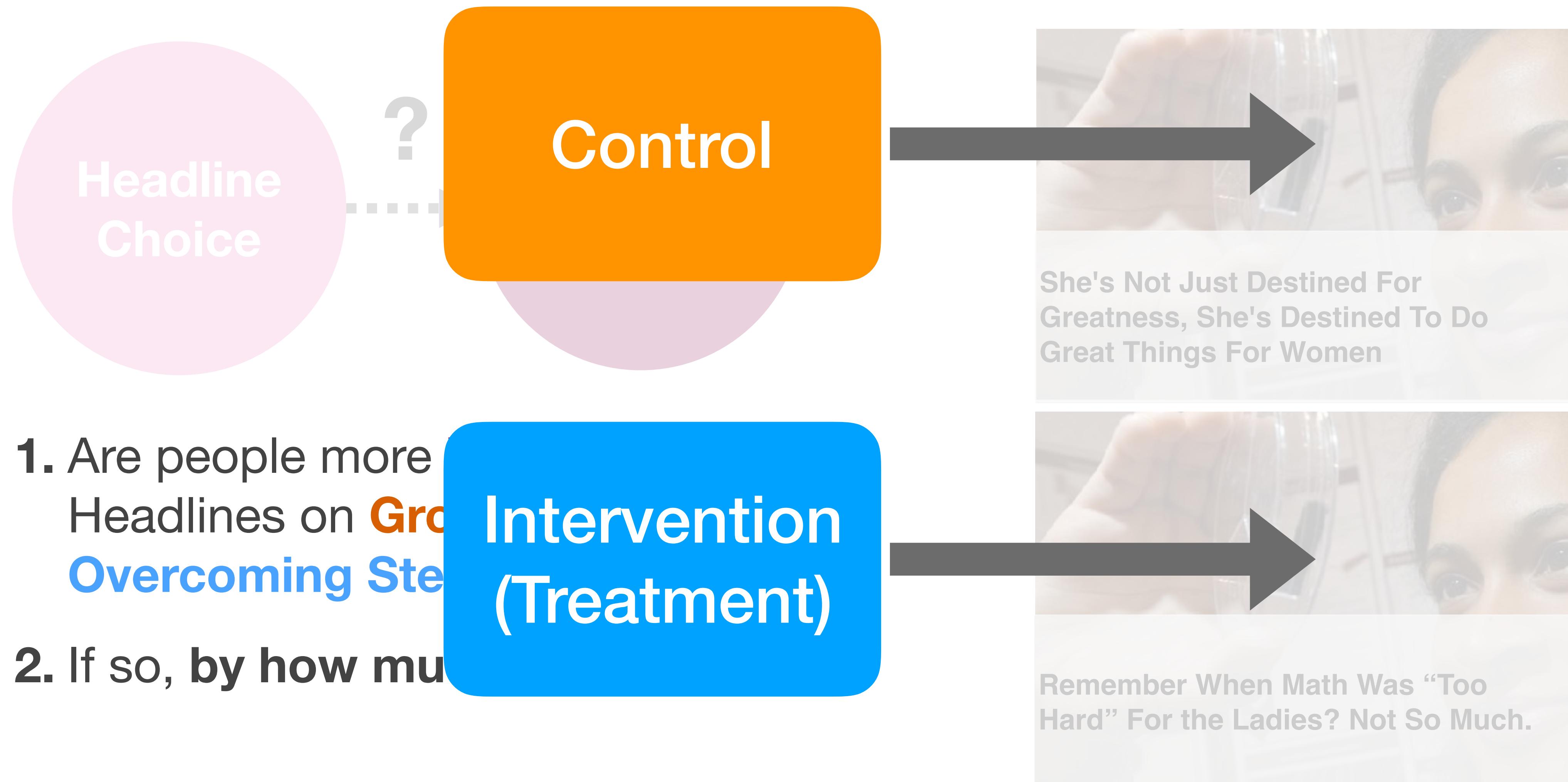
Remember When Math Was “Too Hard” For the Ladies? Not So Much.

B

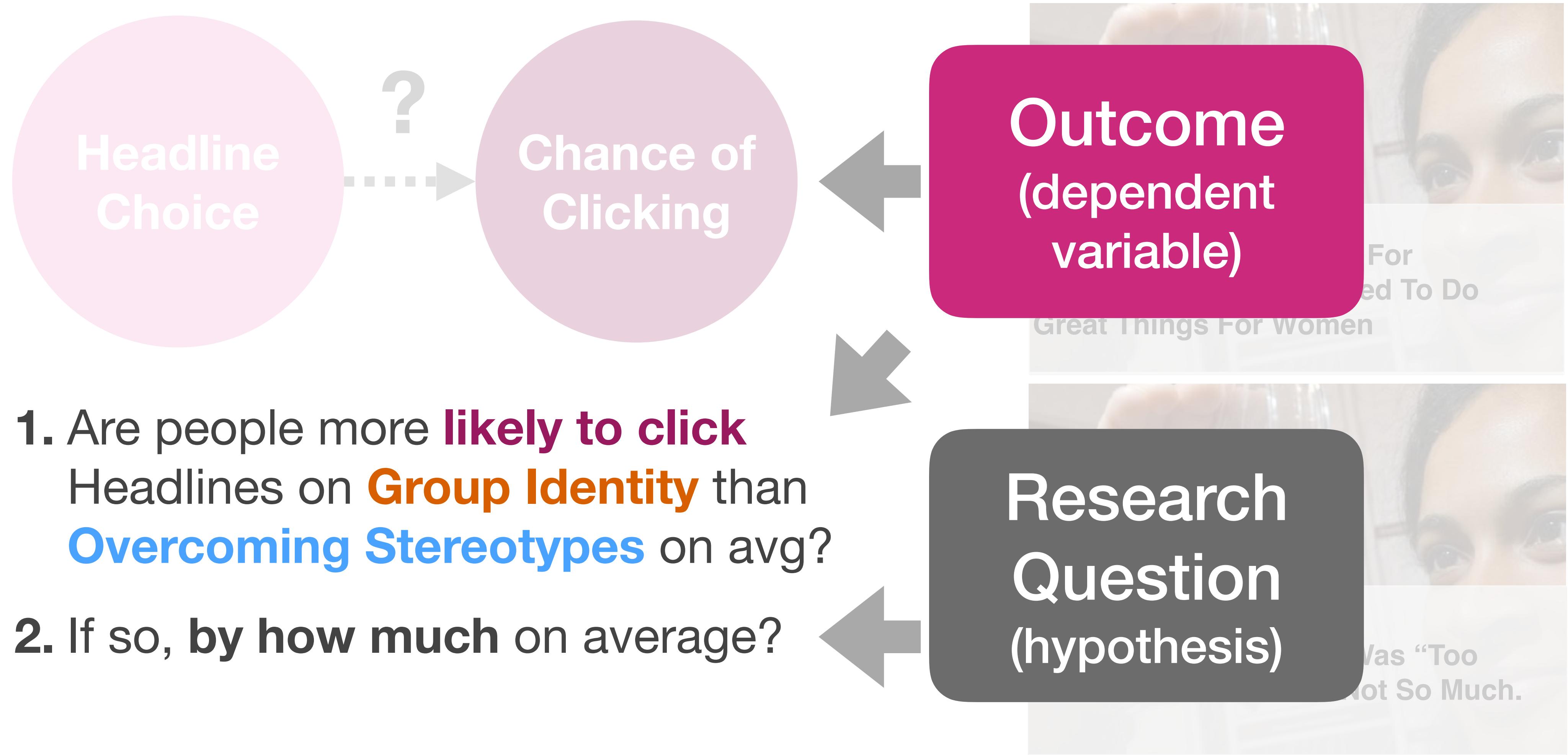
Parts of an Experiment (arms)



Parts of an Experiment (arms)



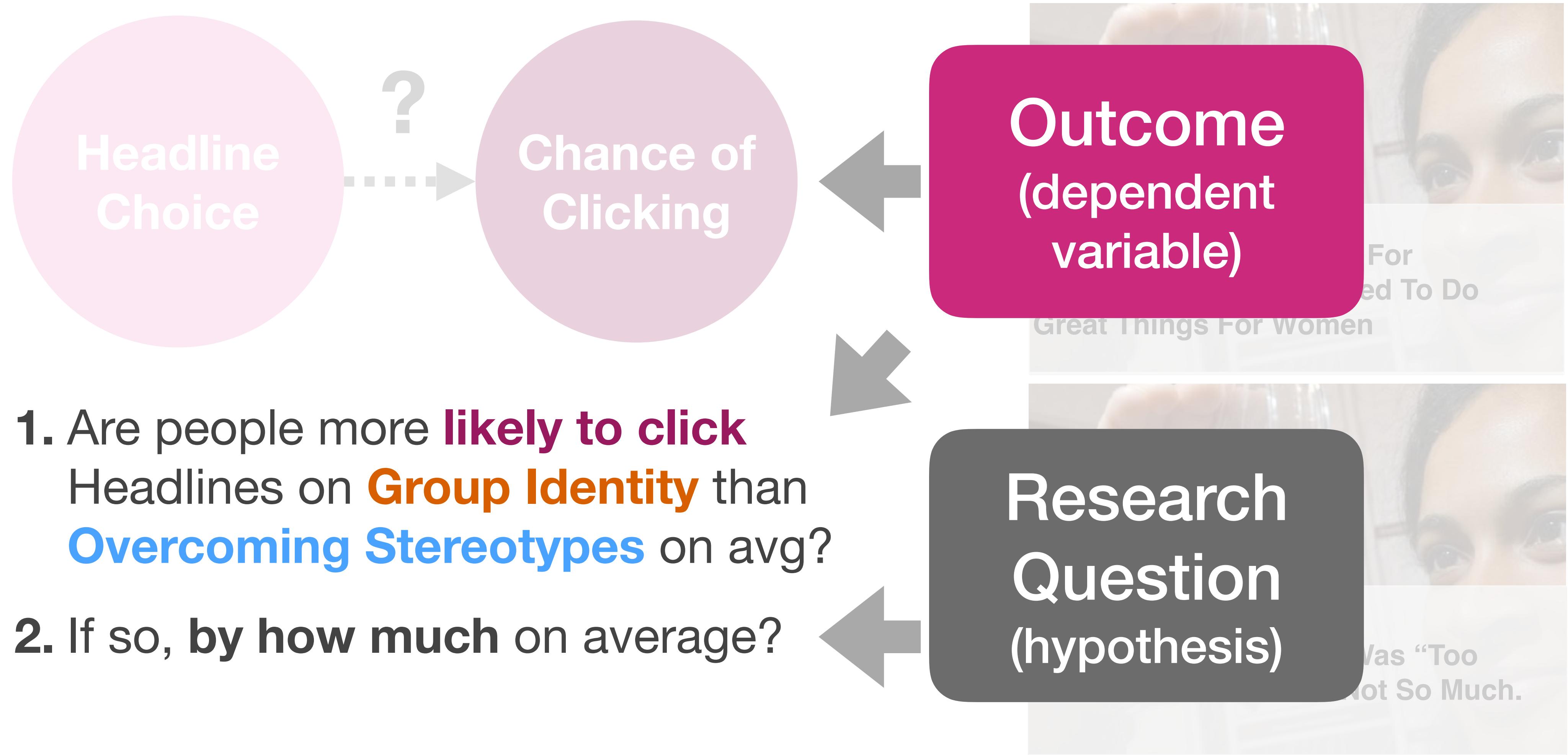
Parts of an Experiment (arms)



A

B

Parts of an Experiment (arms)

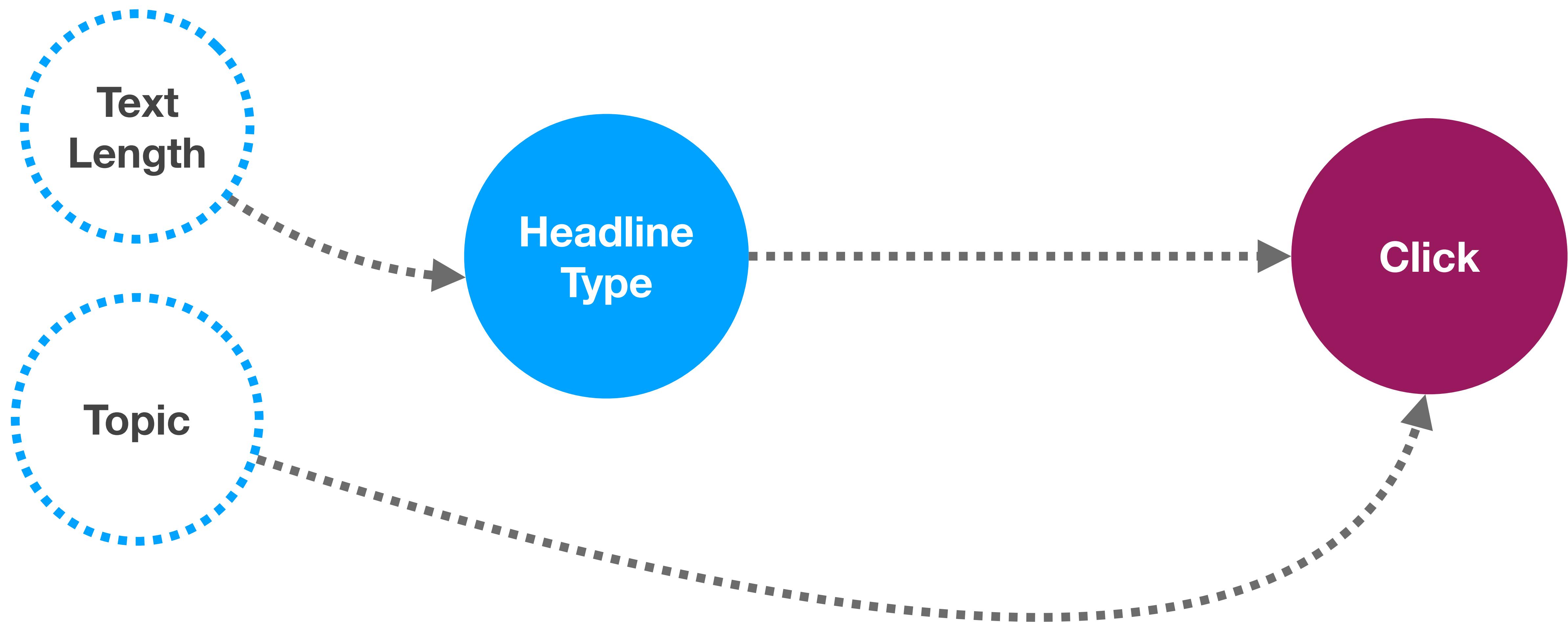


A

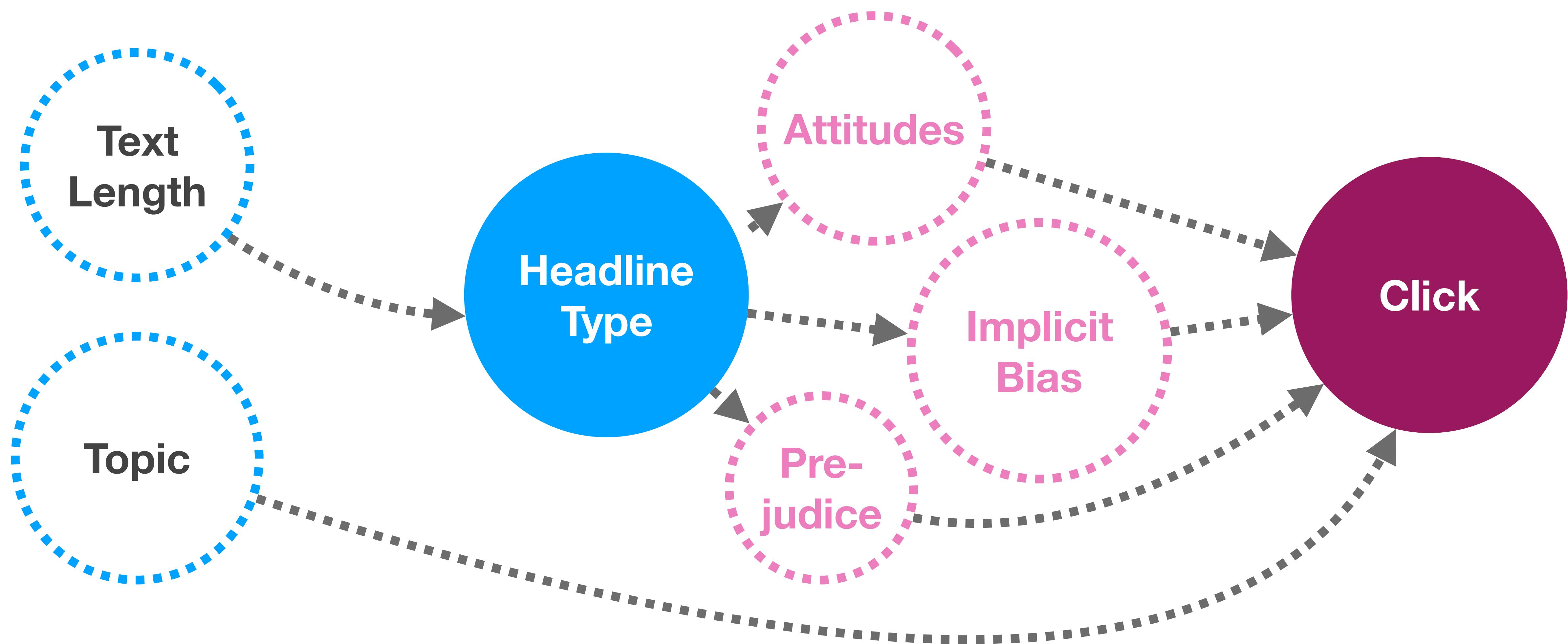
B

Unobserved Confounders

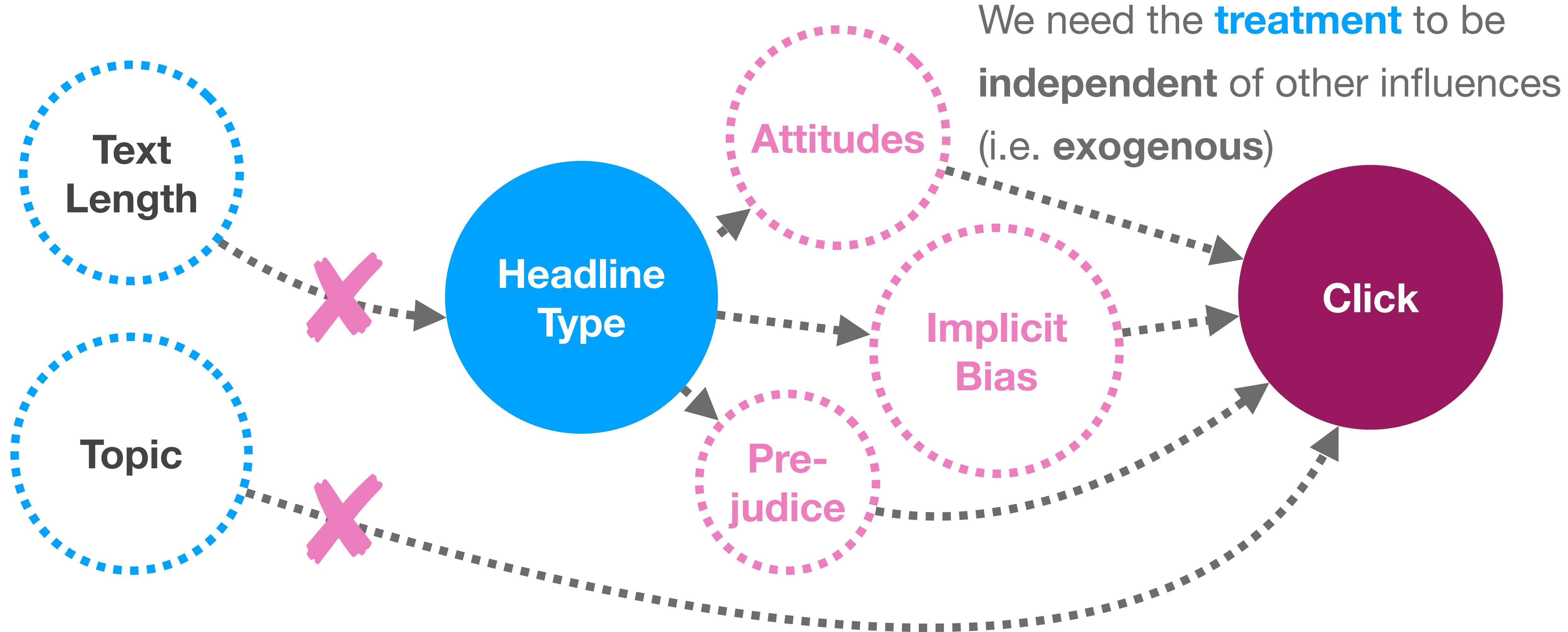
(by the way, this graph is called a path diagram)



Unobserved Confounders



Independence of Treatment



Random Assignment

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

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

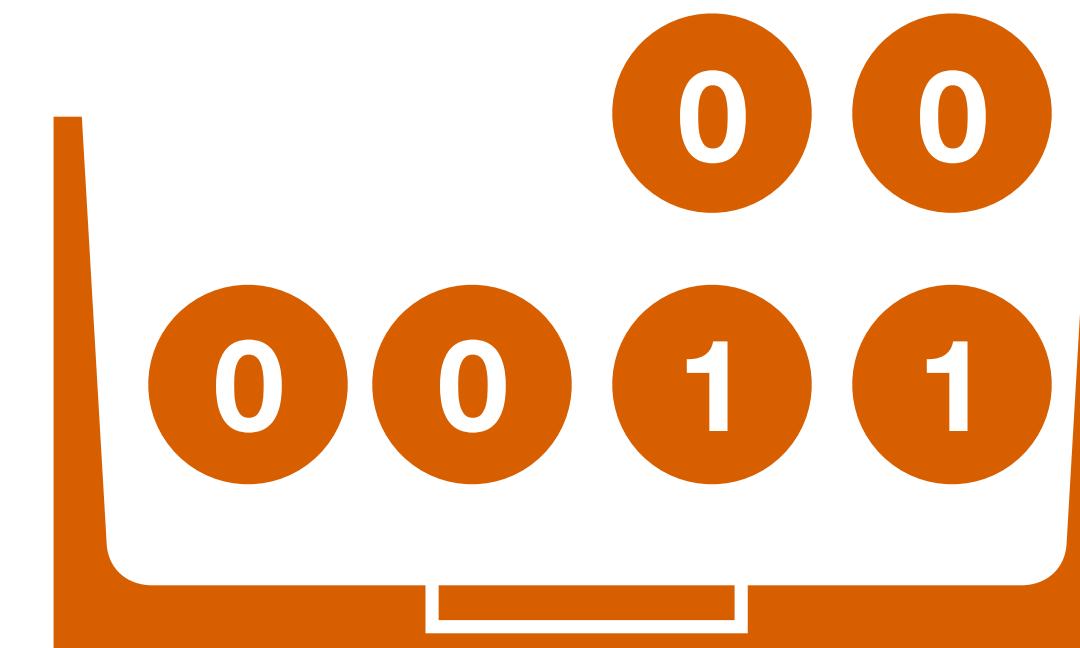
Imagine each headline type has a basic potential “clickability” and that the writer unconsciously allocates colors based on that factor

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

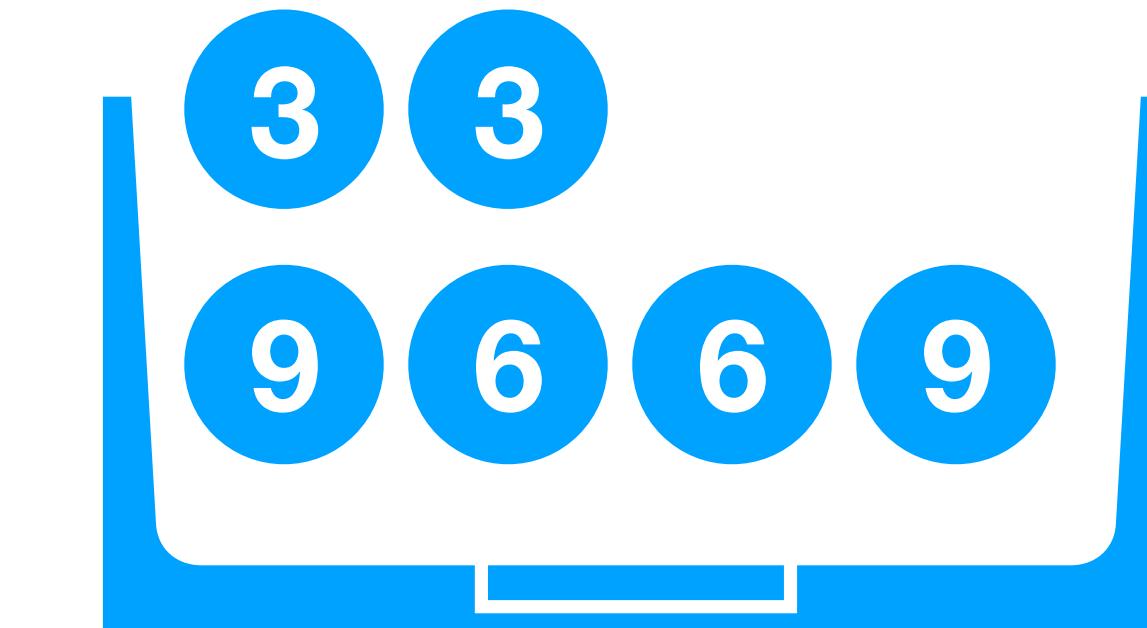


Understanding Randomization via Sampling

Imagine each headline type has a basic potential “clickability” and that the writer unconsciously allocates colors based on that factor



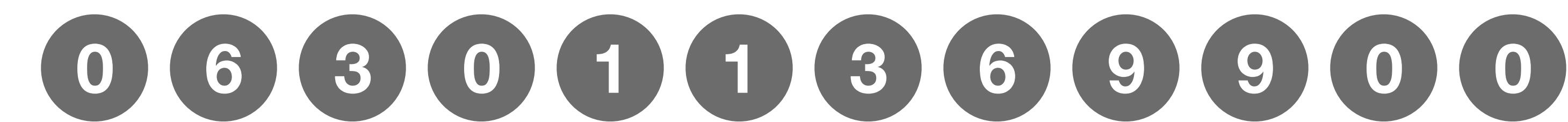
Group
(0 interactions)



Stereotype
(6 interactions)

Understanding Randomization via Sampling

Now imagine that we allocate articles to get a headline type based on a random sample

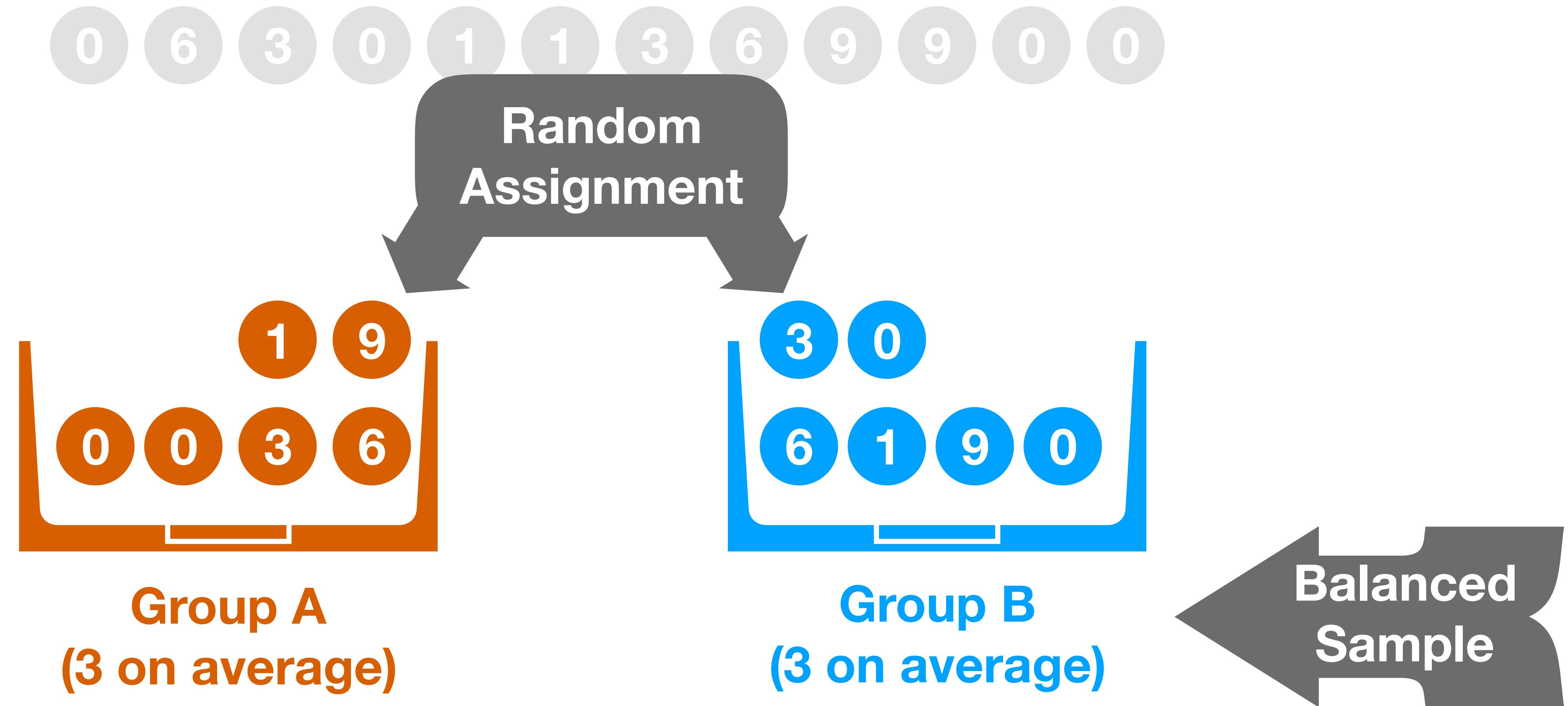


3 interactions
on average



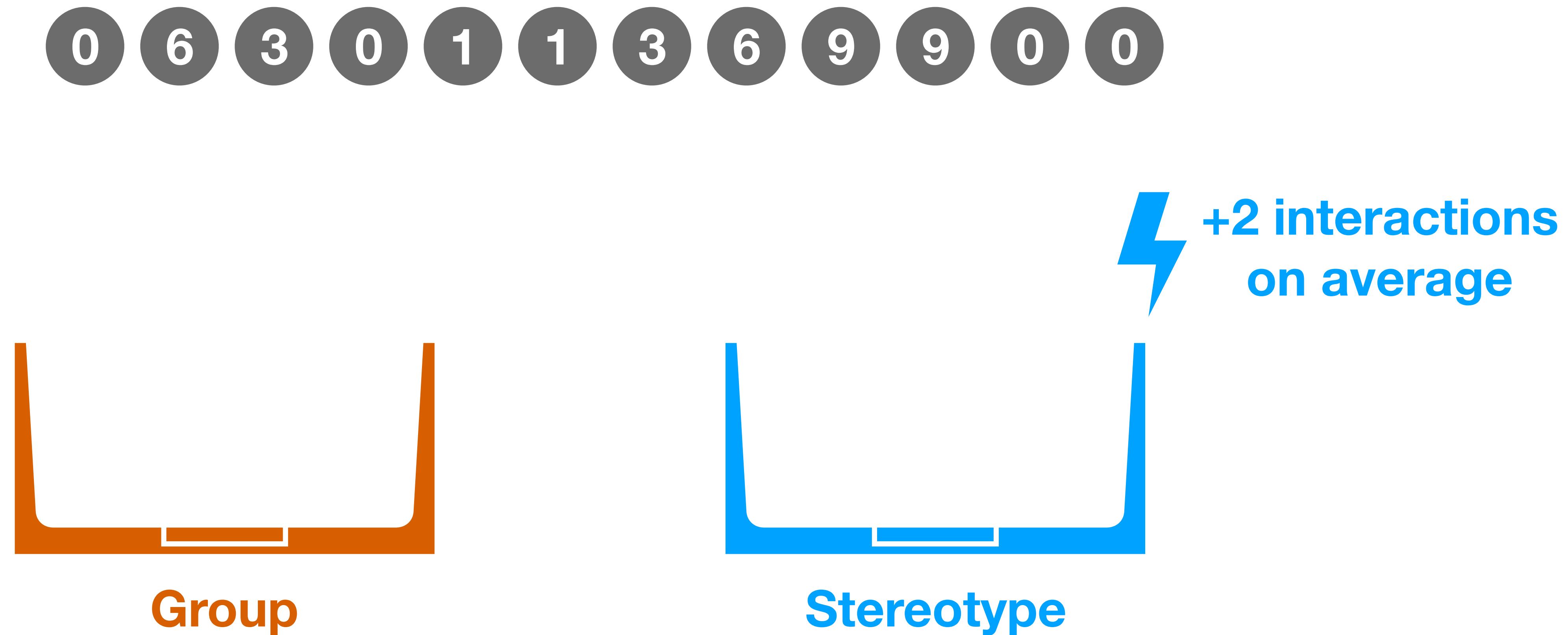
Understanding Randomization via Sampling

Now imagine that we allocate articles to get a headline type based on a random sample



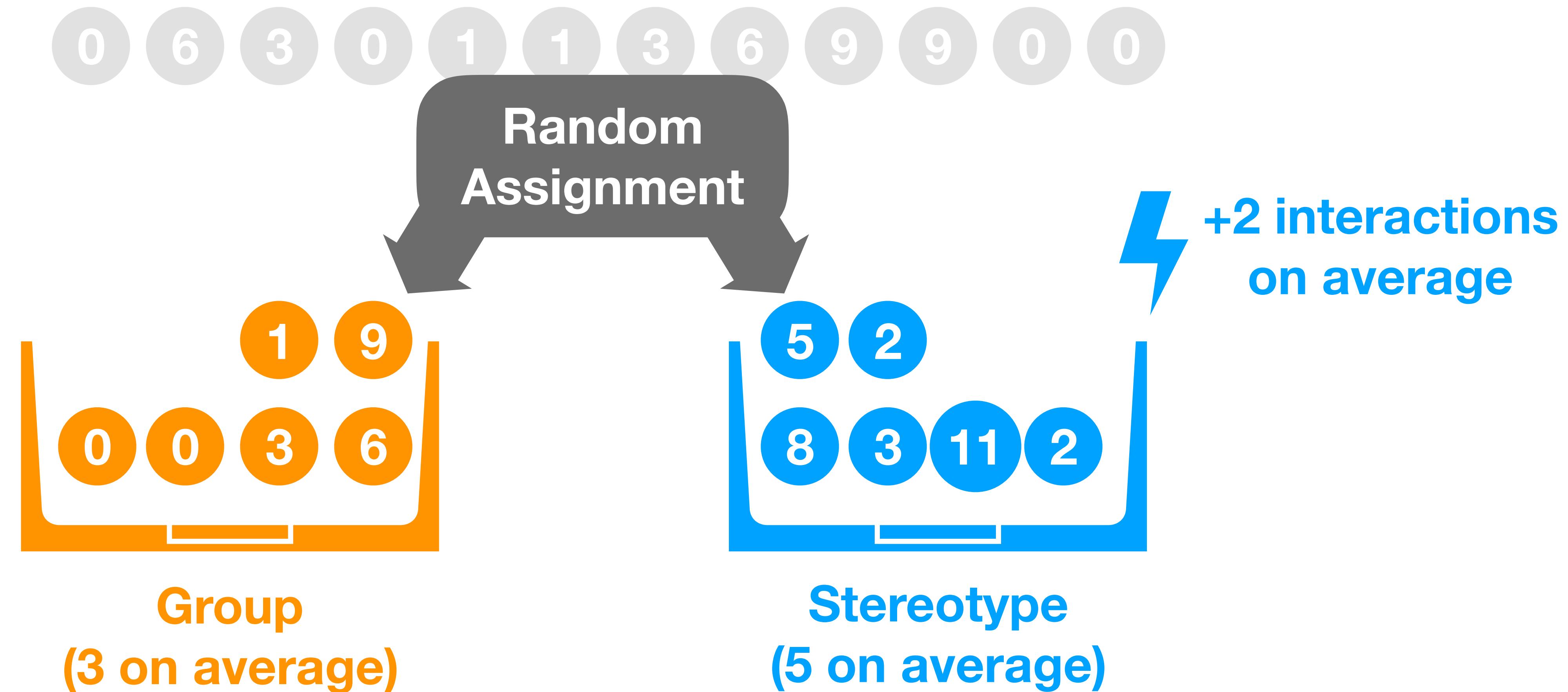
Understanding Randomization via Sampling

Now imagine that an “overcoming stereotype” headline has an effect



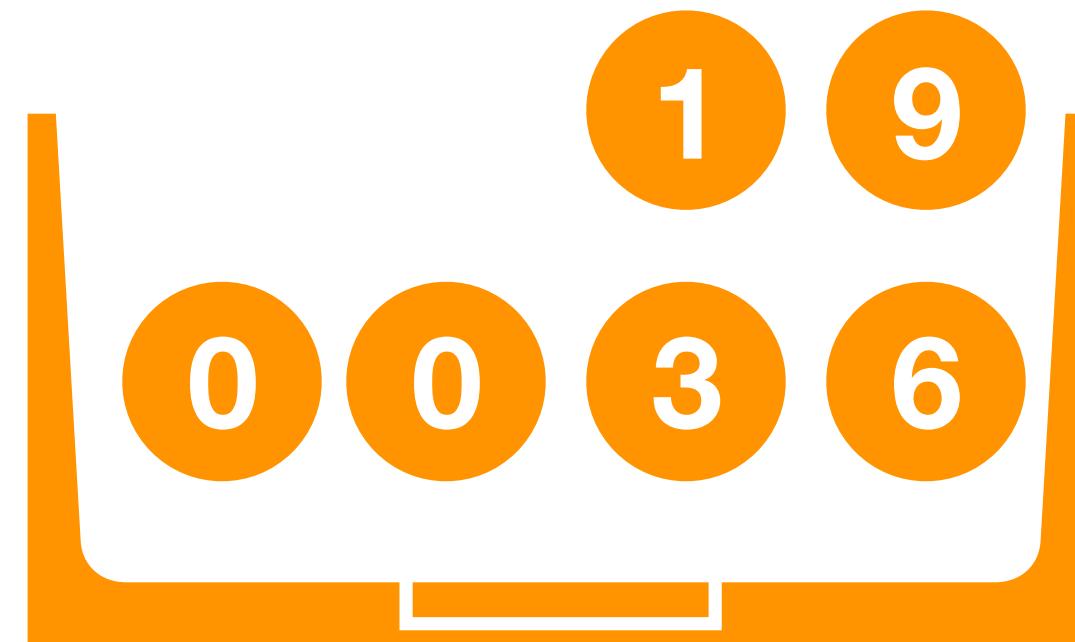
Understanding Randomization via Sampling

Now imagine that an “overcoming stereotype” headline has an effect

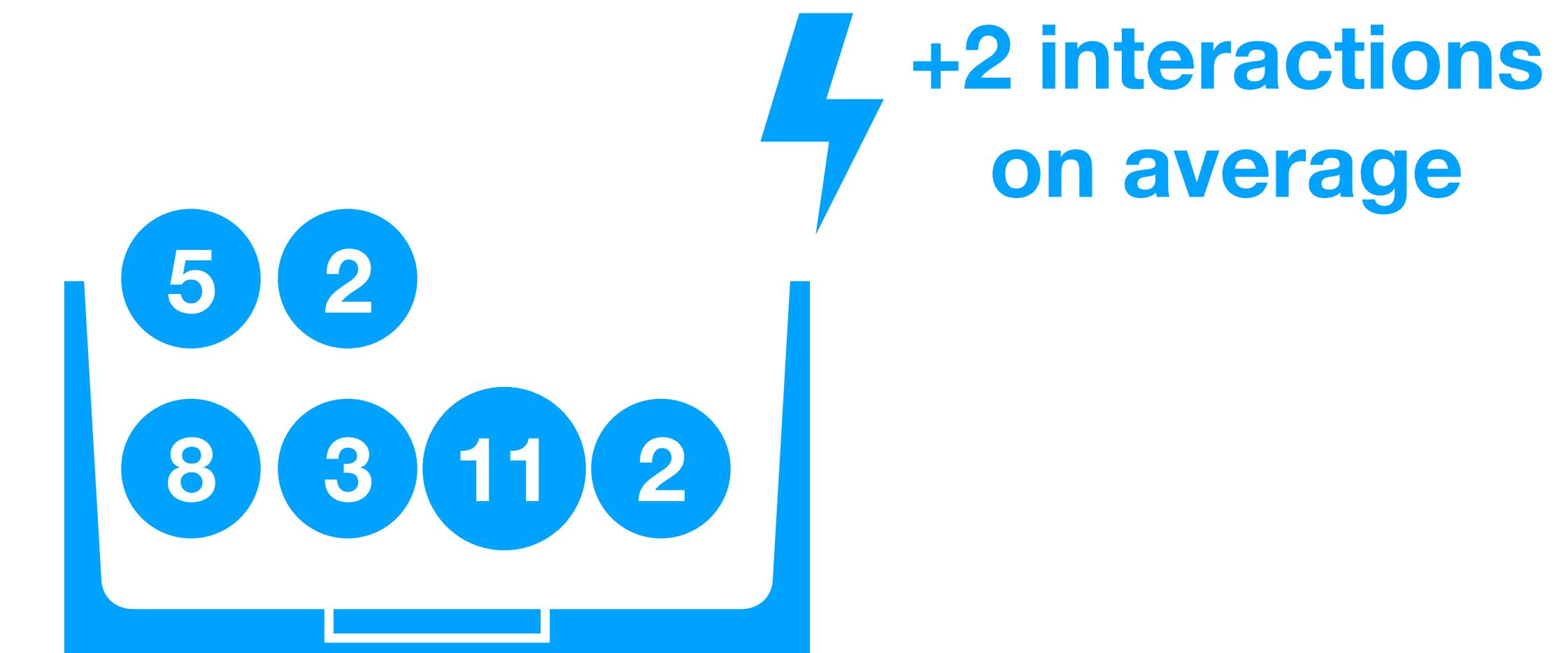


Average Treatment Effect

On average, the difference between the treatment and control



Group
(3 on average)



Stereotype
(5 on average)

+2 interactions
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 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 / representative 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

Tuesday's Assignment

This Week's Assignment

A



She's Not Just Destined For Greatness, She's Destined To Do Great Things For Women

B



This Young Woman Just Took Silicon Valley By Storm And She's Not Stopping There

C



Feminism 101: This Girl Is Going Places And She's Taking Other Girls With Her

D



Remember When Math Was “Too Hard” For the Ladies? Not So Much.

Tuesday's Assignment

For this assignment, create a report for Upworthy that describes what you learned and proposes which headline to use. You should also explain the benefits of causal inference, and argue why field experiments could help the foundation test its headlines and beyond.

Your essay should include:

- a paragraph **describing the experiment design**, including the intervention being tested, the outcome measures being used, and how many participants were included.
- a paragraph **summarizing the findings**. It should summarize the outcome variable, the means for each condition, and include a statement of the effect size.
- a paragraph that **suggests a course of action**, contextualizing the findings in a way that the organization would normally think about, such as the payoff per thousand people who see the headline. Think about whether the result could inform future headline writing. Make sure to reflect on the limitations of the sample, which is drawn from the Upworthy's homepage.
- include a **table of results** and an **illustration of the average treatment effect**. You could (a) show the effect with error bars or (b) show fitted(predicted) values for each condition, with error bars for the treatment (color). If you show fitted values, document details of any covariates(predictors) used to generate the fitted values (such as weekend).
- a paragraph that builds on this finding in the attempt to **convince Upworthy to do more testing** with headlines and in the organization.

Using CoCalc

- Receive up-to-date assignment code & data
- Run analyses on CoCalc Server
- Collaborate over Code & Data with Partners
- Submit Assignments that Work
- Backup & Archive records
- Keep code & data after you graduate (if you update your email address)

Using CoCalc

The screenshot shows the CoCalc web interface. At the top, there is a navigation bar with icons for Projects, J. Nathan Matias - con, CoCalc, Help, Account, a bell, and a search icon. Below the navigation bar, the word "Projects" is displayed in large, bold, dark gray letters. A search bar with the placeholder "Search for projects..." and a clear button is present. A green button labeled "+ Create New Project..." is visible. A pink hand-drawn circle highlights the first project in the list, which is titled "J. Nathan Matias - comm4940_2020". This project has a "No description" note and was created "less than a minute ago". To the right of this project, the text "COMM 49490 Folder" is prominently displayed in large orange letters. Below the highlighted project, another project is listed: "Max Klein (4 hours ago), Comm 4940" with a status of "Running". At the bottom of the project list, there is a blue button labeled "Load any older projects...".

Projects

Deleted

Search for projects...

+ Create New Project...

J. Nathan Matias - comm4940_2020

No description

less than a minute ago

Max Klein (4 hours ago), Comm 4940

Running

Welcome to CoCalc!

7 hours ago

Stopped

Load any older projects...

Viewing Folders

The screenshot shows the CoCalc interface for viewing files. The top navigation bar includes 'Projects' (with a yellow sun icon), the current workspace 'J. Nathan Matias - con' (with a pencil icon), 'CoCalc' (info icon), 'Help' (help icon), 'Account' (user icon), a bell icon for notifications, a Wi-Fi icon, and a full-screen icon.

The main menu bar has tabs for 'Files' (selected), 'New', 'Log', 'Find', 'Settings', 'headline-analysis-example.R.ipynb' (the current file), 'README.md', and a 'Share' button.

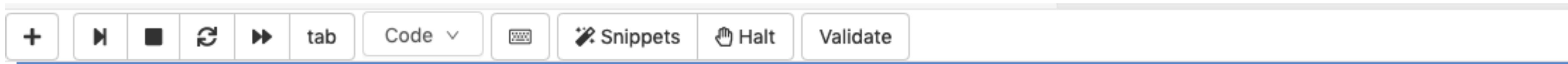
Below the menu is a search bar labeled 'Search or create file' with a clear button, a 'New' button with a dropdown arrow, and a breadcrumb navigation bar showing the path: 'design-...rimentments / assignments / 1-headl...eriment'.

On the left, there's a checkbox labeled 'Check All'. On the right, there are buttons for 'Library', 'Upload', a refresh icon, a download icon, a cloud icon, and 'Backups'.

A message indicates '7 items — Click on the checkbox to the left of a file to copy, move, delete, download, etc.'

Type	Name	Date Modified	Size	Cloud
<input type="checkbox"/>	DUE_DATE.txt	4 hours ago	47 bytes	
<input type="checkbox"/>	headline-analysis-example.R.ipynb	4 hours ago	5.3 KB	
<input type="checkbox"/>	headline-analysis-example.R.ipynb~	4 hours ago	31.6 KB	
<input type="checkbox"/>	DUE_DATE.txt~	4 hours ago	47 bytes	
<input type="checkbox"/>	README.md	4 hours ago	3.9 KB	
<input type="checkbox"/>	headline-experiment-heds.csv	4 hours ago	322 bytes	
<input type="checkbox"/>	headline-experiment-impressions.csv	4 hours ago	59.8 KB	

Editing & Running Jupyter Notebooks



Headline Analysis Example (assignment one)

J. Nathan Matias, COMM 4940 (Jan 2020)

This code includes an analysis example for the [Week 2 Headline Experiment](#).

DATASET DESCRIPTION

Each row in *headline-experiment-impressions.csv* represents a single "impression" - a browser that viewed a given headline:

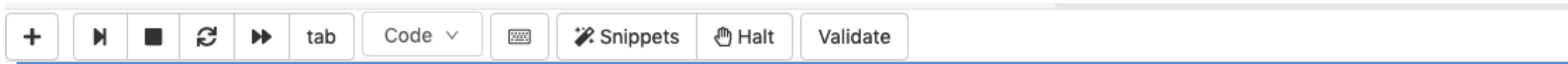
- **hed**: which headline number was used (the actual headline text is in *headline-experiment-heds.csv*)
- **click**: whether that participant clicked on the article to read further

```
In [ ]: ## LOAD LIBRARIES
library(ggplot2) ## FOR PLOTS
library(gmodels) ## FOR CrossTable

## SET GGPLOT TO USE WIDE BUT NOT TOO TALL PLOTS
options(repr.plot.width=6, repr.plot.height=4)
```

```
In [ ]: # load participants file
participants    <- read.csv("headline-experiment-impressions.csv")
# convert headline to a factor, to simplify analysis later on
```

Editing & Running Jupyter Notebooks



Headline Analysis Example (assignment one)

J. Nathan Matias, COMM 4940 (Jan 2020)

This code includes an analysis example for the [Week 2 Headline Experiment](#).

DATASET DESCRIPTION

Each row in *headline-experiment-impressions.csv* represents a single "impression" - a browser that viewed a given headline:

- **hed**: which headline number was used (the actual headline text is in *headline-experiment-heds.csv*)
- **click**: whether that participant clicked on the article to read further

```
In [ ]: ## LOAD LIBRARIES
library(ggplot2) ## FOR PLOTS
library(gmodels) ## FOR CrossTable

## SET GGPLOT TO USE WIDE BUT NOT TOO TALL PLOTS
options(repr.plot.width=6, repr.plot.height=4)
```

```
In [ ]: # load participants file
participants    <- read.csv("headline-experiment-impressions.csv")
# convert headline to a factor, to simplify analysis later on
```

Editing Markdown Documents

The screenshot shows the CoCalc interface for editing a Markdown document named README.md. The left pane displays the raw Markdown code with line numbers. The right pane shows the rendered HTML output.

Raw Markdown (Left):

```
1 # Headline Experiment(Week 2 Assignment)
2
3 **COMM4940**
4
5 Cornell University Department of Communication
6
7 J. Nathan Matias ([@natematias]
8 (https://twitter.com/natematias))
9 > *Disclaimer: while this assignment includes reference to
10 factual people, events, and experiments, the assignment is
11 a hypothetical scenario.
12
13 ## Background
14
15 One day in 2013, you receive an email from Upworthy, an
organization dedicated to amplifying meaningful and
inspiring stories.  
Upworthy wants to reach the greatest number of people with
```

Rendered View (Right):

Headline Experiment(Week 2 Assignment)

COMM4940

Cornell University Department of Communication

J. Nathan Matias (@natematias)

*Disclaimer: while this assignment includes reference to factual people, events, and experiments, the assignment is a hypothetical scenario.

Background

One day in 2013, you receive an email from Upworthy, an organization dedicated to amplifying meaningful and inspiring stories.

Working with an Assignment

- Navigate to the assignment folder under *design-governance-experiments/assignments*
- Complete the assignment in CoCalc (or upload files to CoCalc)
- The assignment will be auto-submitted at the due date
- If you need longer, reach out to the professor

Working with an Assignment

- Navigate to the assignment folder under *design-governance-experiments/assignments*
- Complete the assignment in CoCalc (or upload files to CoCalc)
- The assignment will be auto-submitted at the due date
- If you need longer, reach out to the professor