

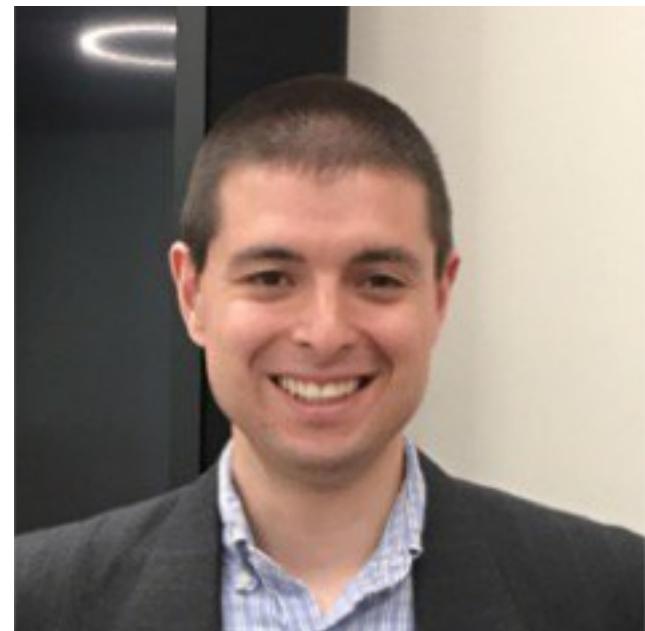
# COMM 4940

# The Design & Governance of

# Field Experiments

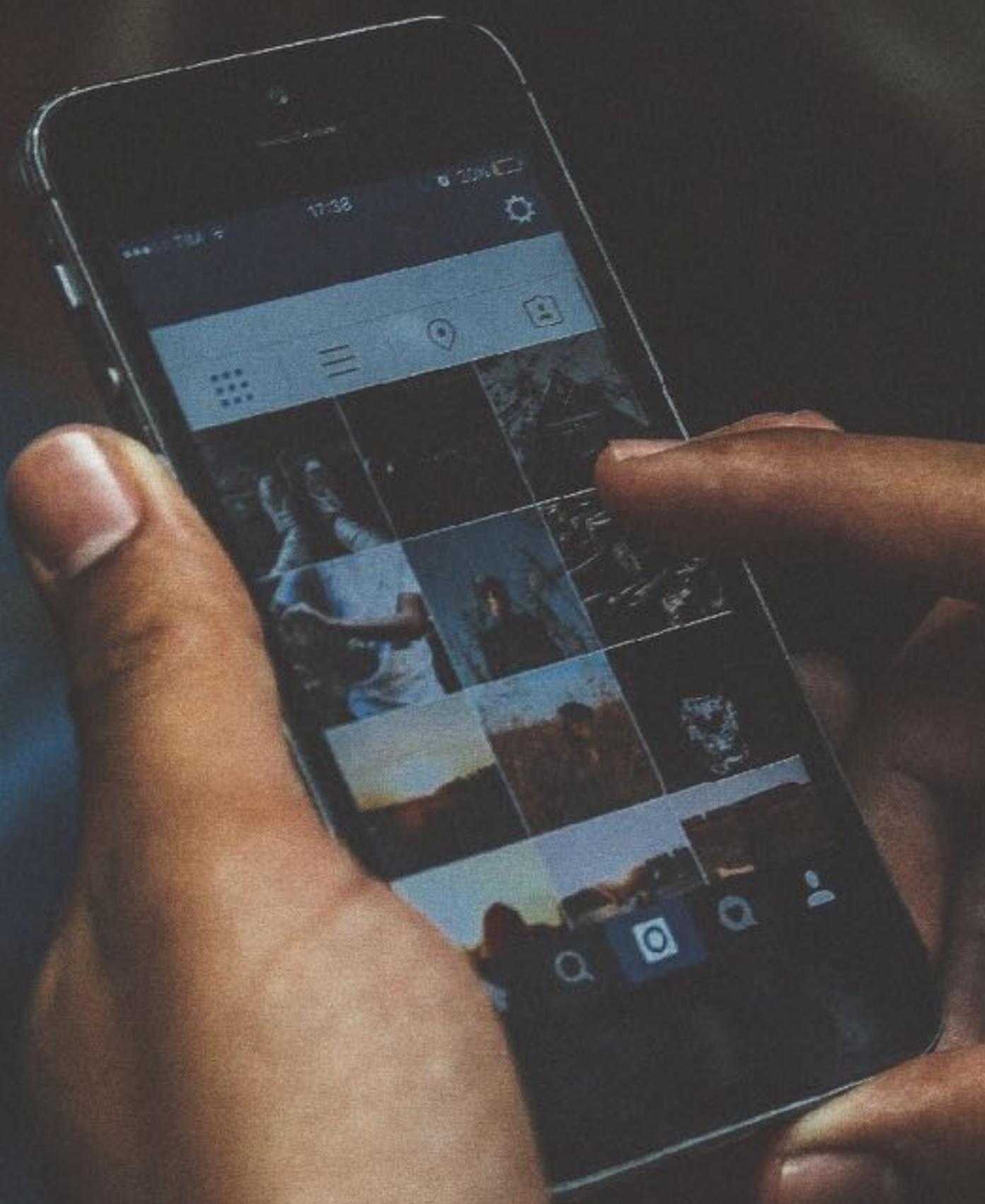
Jan 22, 2020

Kennedy Hall 213



**J. Nathan Matias**  
@natematias  
[citizensandtech.org](http://citizensandtech.org)  
[natematias.com](http://natematias.com)





# Disinformation Is Becoming Unstoppable



A man is silhouetted against a video screen with a Facebook logo on Aug. 14, 2013. Dado Ruvic—Reuters

By **DIPAYAN GHOSH AND BEN SCOTT**

January 24, 2018

## Weekend magazine technology special

### 'Our minds can be hijacked': the tech insiders who fear a smartphone dystopia

Google, Twitter and Facebook workers who helped make technology so addictive are disconnecting themselves from the internet. **Paul Lewis** reports on the Silicon Valley refuseniks alarmed by a race for human attention

by Paul Lewis in San Francisco

The WorldPost • Opinion

### Pierre Omidyar: 6 ways social media has become a direct threat to democracy

By **Pierre Omidyar**

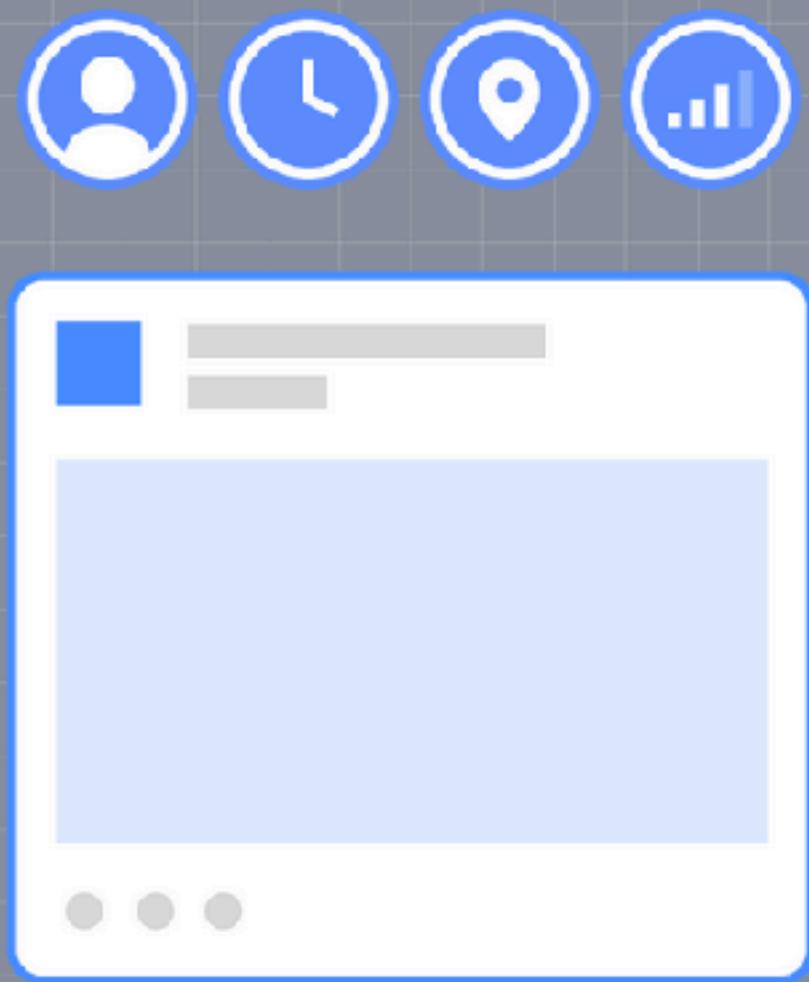
October 9, 2017



People hold smartphones against a backdrop projected with the Twitter logo. (Kacper Pempel/Reuters)

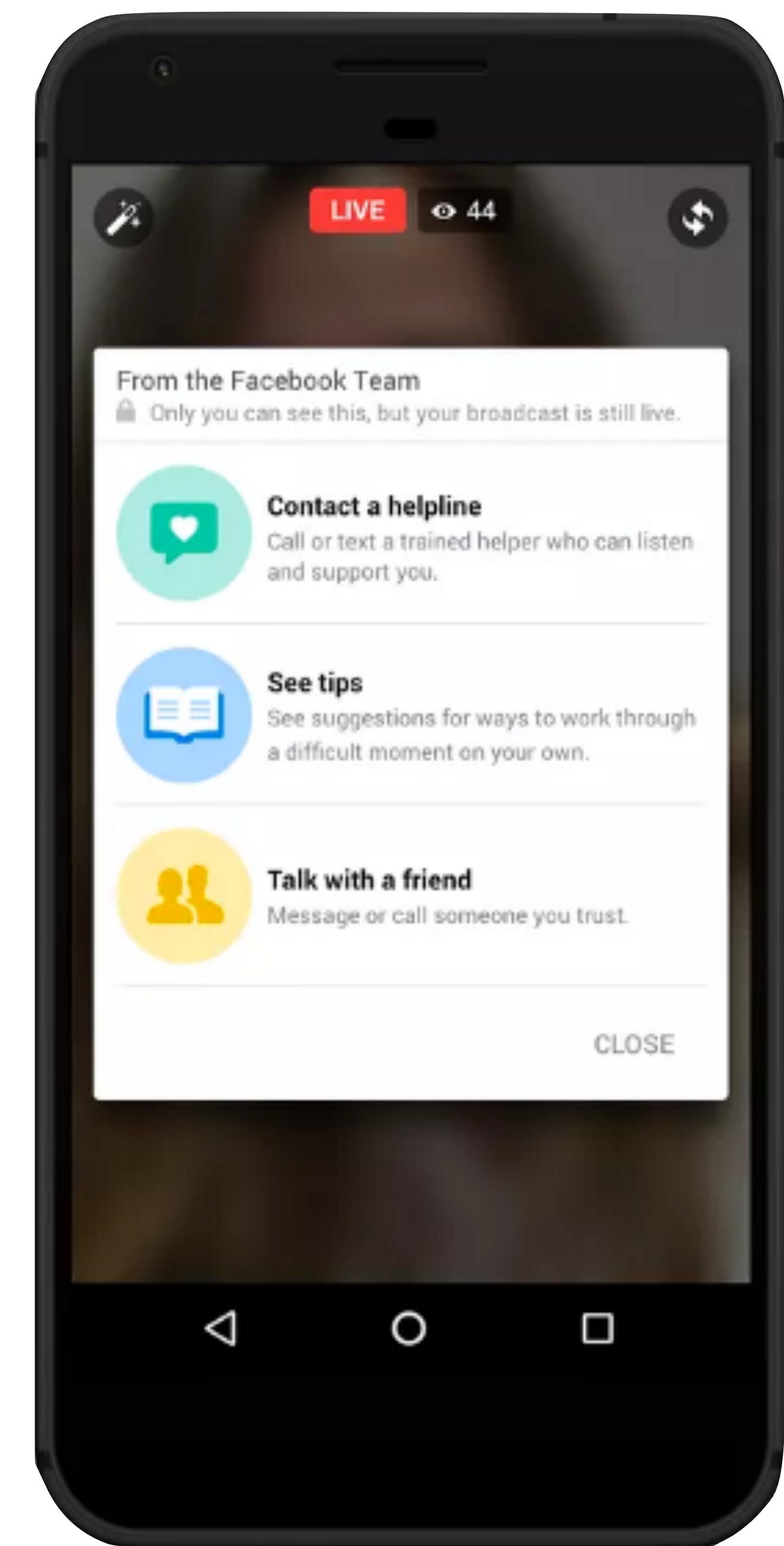
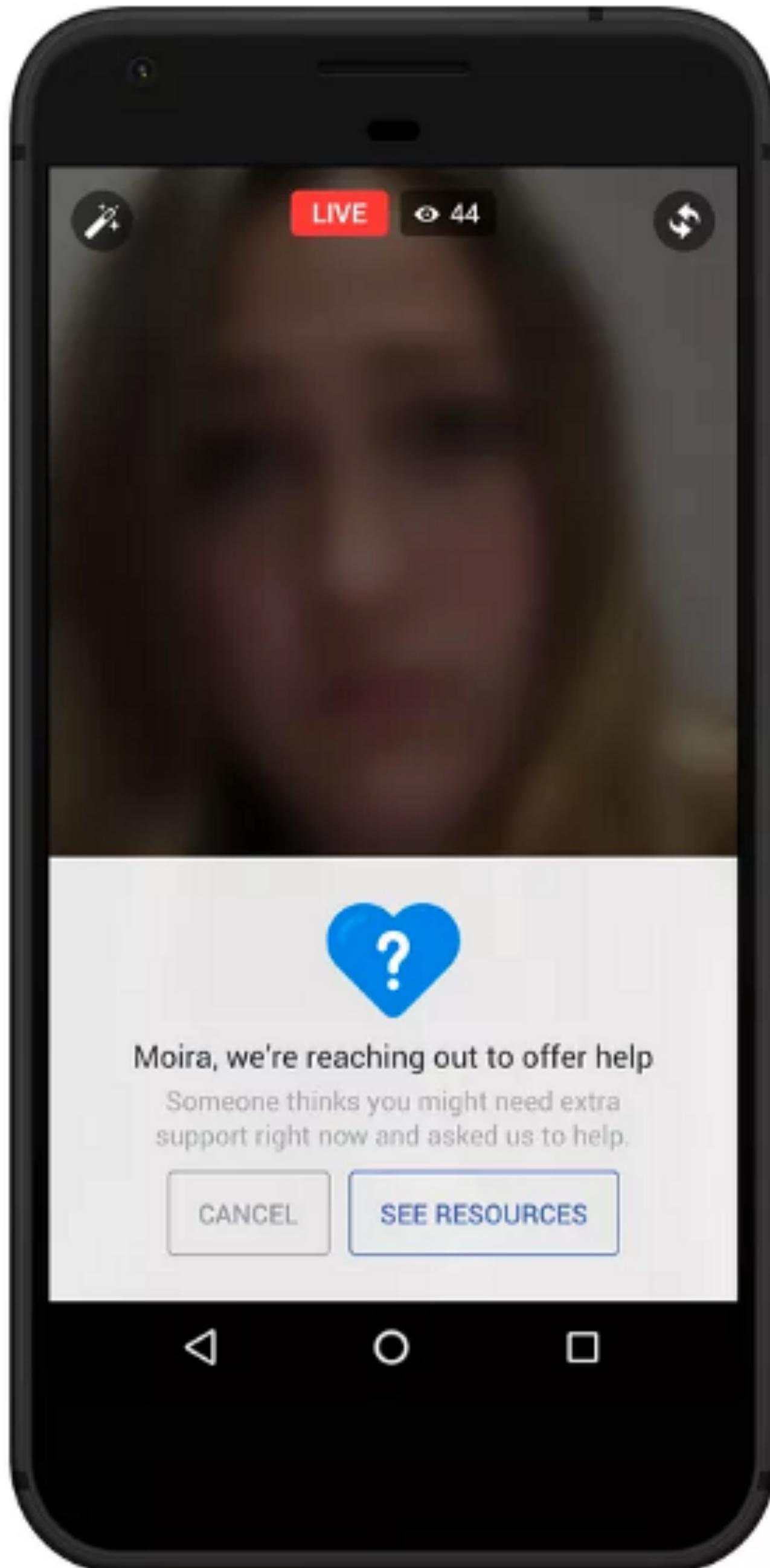
September 10, 2018

# How Facebook AI Helps Suicide Prevention



By Catherine Card, Director of Product Management

Around the world, a death by suicide occurs every 40 seconds and is the second-leading cause of death for 15- to 29-year-olds. In the US, nearly 45,000 people take their life every year.



How can you know if an intervention is making things **better** or **worse**?

## Options

Step Up  
Hide Links  
Zoom In  
Normal Scale  
Zoom Out

## Bookmarks

Import

## Lecture Segments

Listen  
Online

Introduction:  
Innovation in  
Ancient Greece,  
Tragedy in  
Electronic  
Literature

Kinetic Poetry:  
"Faith", by  
Robert Kendall

Hypertext  
Fiction: "We  
Descend", by

Interactive  
Fiction:  
"Galatea" by

Community &  
Convention in  
Electronic  
Literature  
(Quicktime Movie.)

Why Write  
Tragedy?

### Tragedy in Electronic Literature

If electronic literature empowers characters and Tragedy disempowers them, can Electronic Literature present Tragedy?

This is supplementary information to a lecture by J. Nathan Matias at Cambridge University on 5 Nov 2007. Students browsed this page using a WIFI network set up just for the lecture.

### Lecture Slides

The items on the left present the slides combined with lecture audio. Slides without audio may be downloaded from here: [View](#)

## Examples of Electronic Literature

Faith: An  
expanding  
multiverse in

Hypertext Technology and

## Conventions and Traditions in Electronic Literature

Poems that Go  
Poems that Go is

Where are the  
Hypertext Poets?

Communities and Software

## Why bother with Tragedy?

Modern criticism  
of Tragedy

Lowe's Diagram  
of the Poetics

Antigone as Protest  
Tactic

## Using This Document

Each of these boxes is called a "note" and contains text, data fields, and links. Some of these notes contain other notes, which is shown by a square, white-ish button at the top right. Click on a note to read it, or click on the corner button to step inside that note's contents. For more information, click on "How to use this Document".

Acknowledgements

J. Nathan Matias

How to Use  
this  
Document

## More Information

Don't read this until  
after the lecture

I want you to listen,  
and it might be too  
interesting

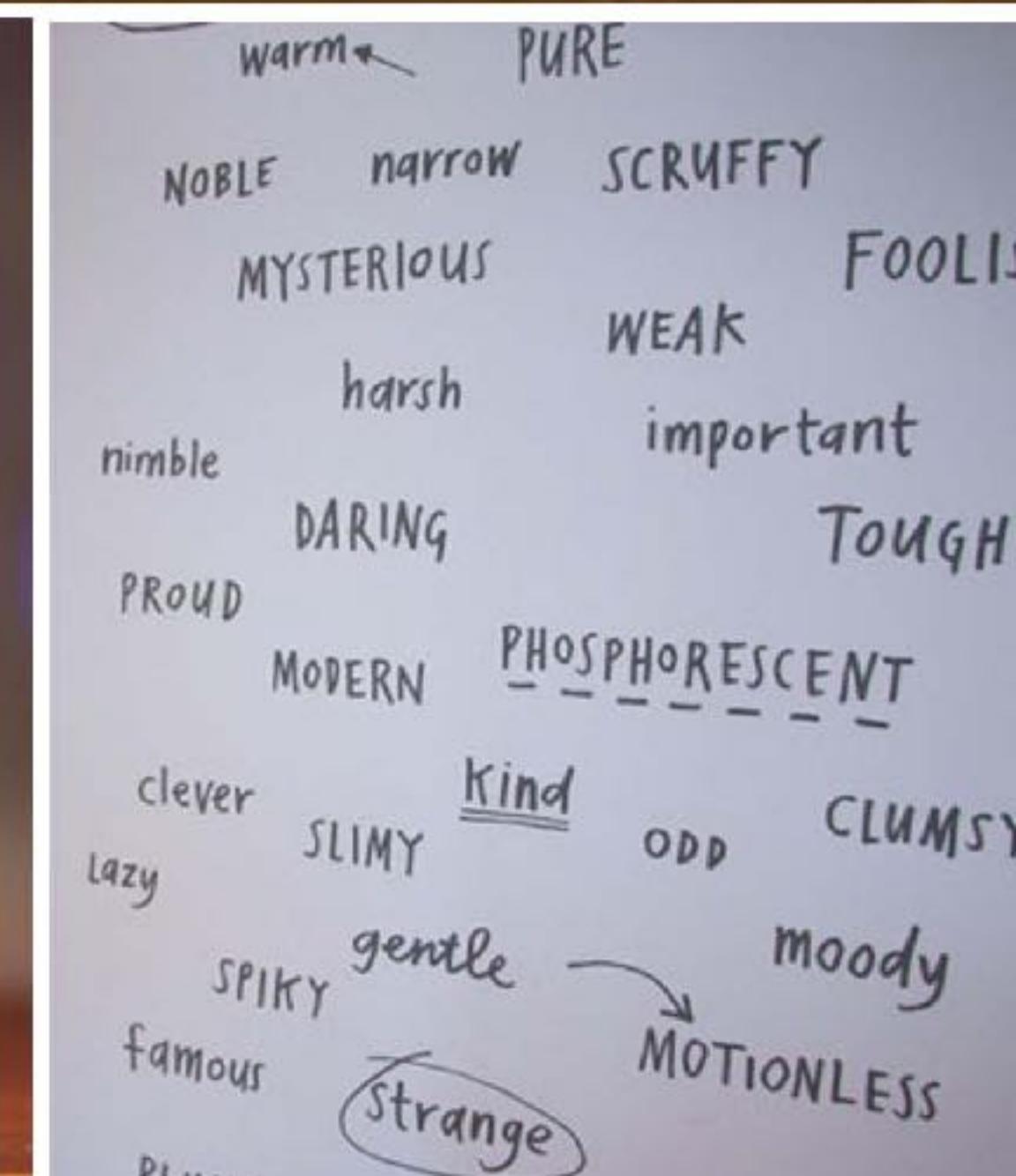
## Murray's Distinctions

Encyclopedic  
Readers tend to  
explore of the available  
information, which

Procedural  
Electronic texts are not  
static. There are many  
ways in which the work

Participatory  
Electronic works invite  
user participation. This  
participation need not

Spatial / Exploratory  
Readers must keep  
track of the structure of  
the story or the



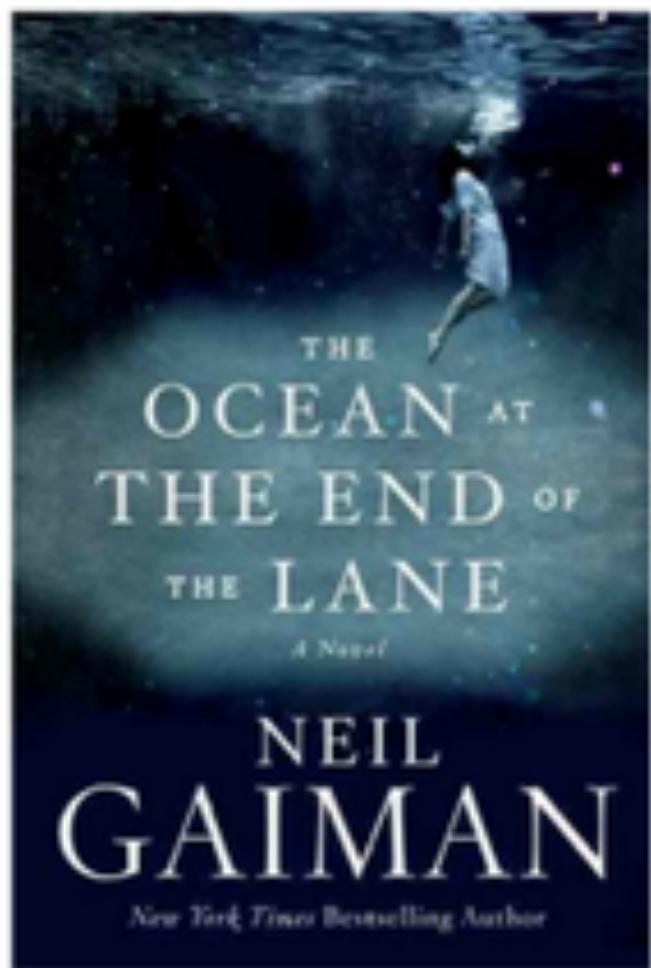


image by Harper Collins

## August Reading Schedule

- |                  |                 |
|------------------|-----------------|
| Week 1: ch 1-4   | <b>#1b140_1</b> |
| Week 2: ch 5-8   | <b>#1b140_2</b> |
| Week 3: ch 9-12  | <b>#1b140_3</b> |
| Week 4: ch 13-15 | <b>#1b140_4</b> |



image © by pinguino k

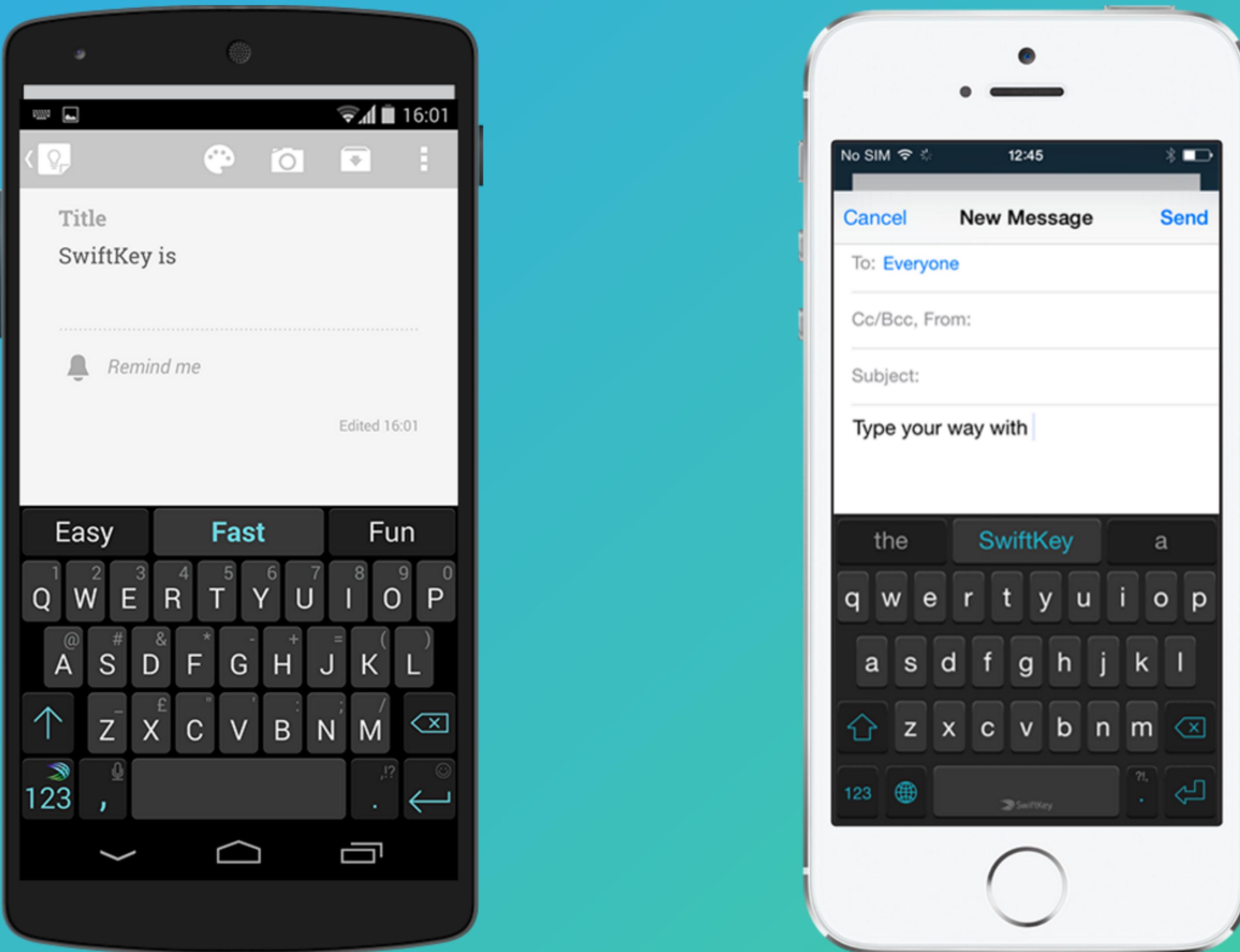
**@1book140**   
Twitter Book Club

have increased their productivity - saving nearly two trillion keystrokes and more than 23,000 years in combined typing time - the equivalent of watching every episode of 'Friends' 19 million times.



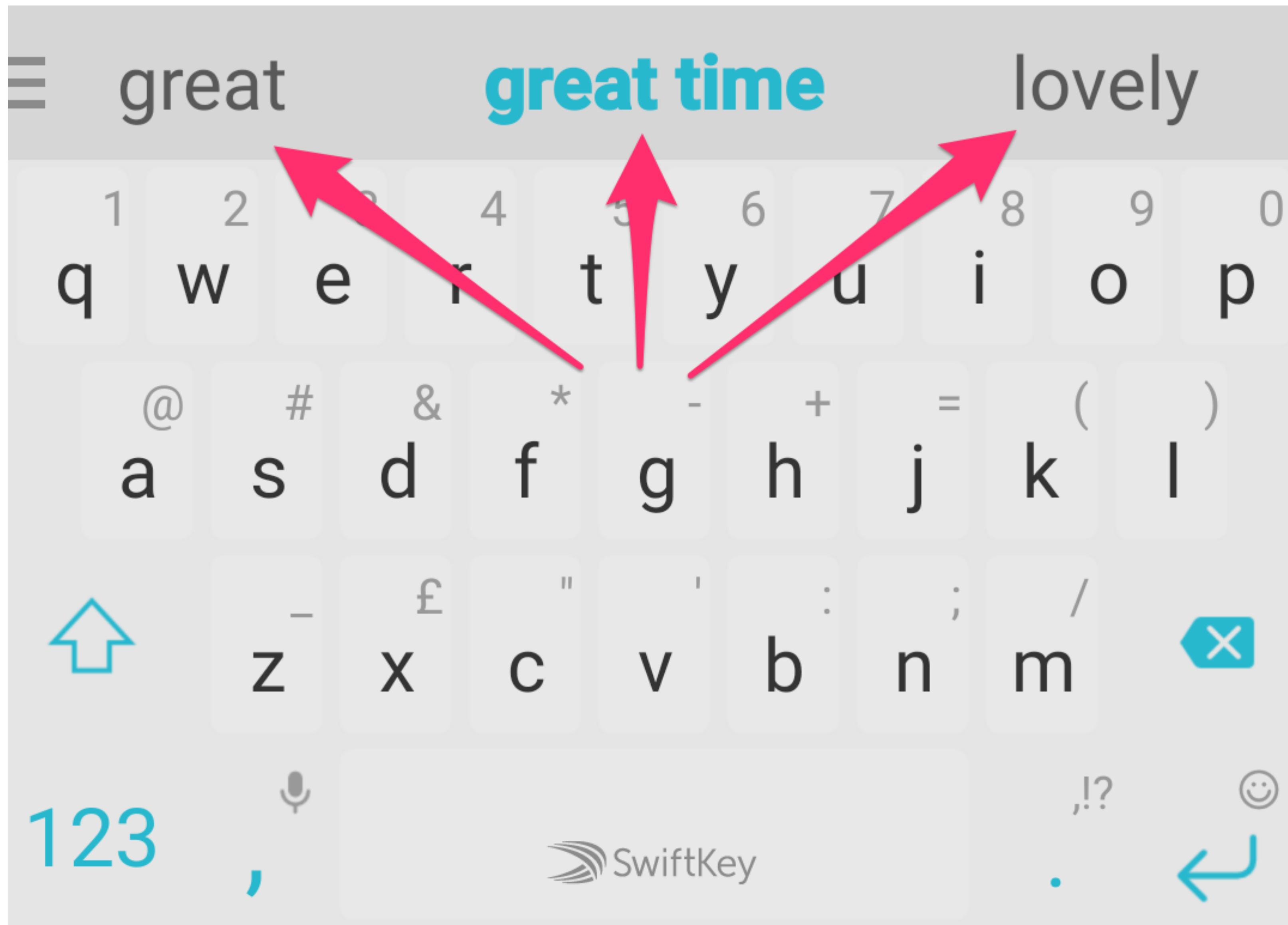
IPHONE & IPAD take

ANDROID





I had such a



How can you know if  
predictive text is doing a  
**good job?**

How can you know if an intervention is making things **better** or **worse**?

# Causal Question:

if people **don't get**  
specific predictions, will  
they be **worse off**?

# Meet Your Neighbor:

- Name
- Year & Major
- An experiment or a question

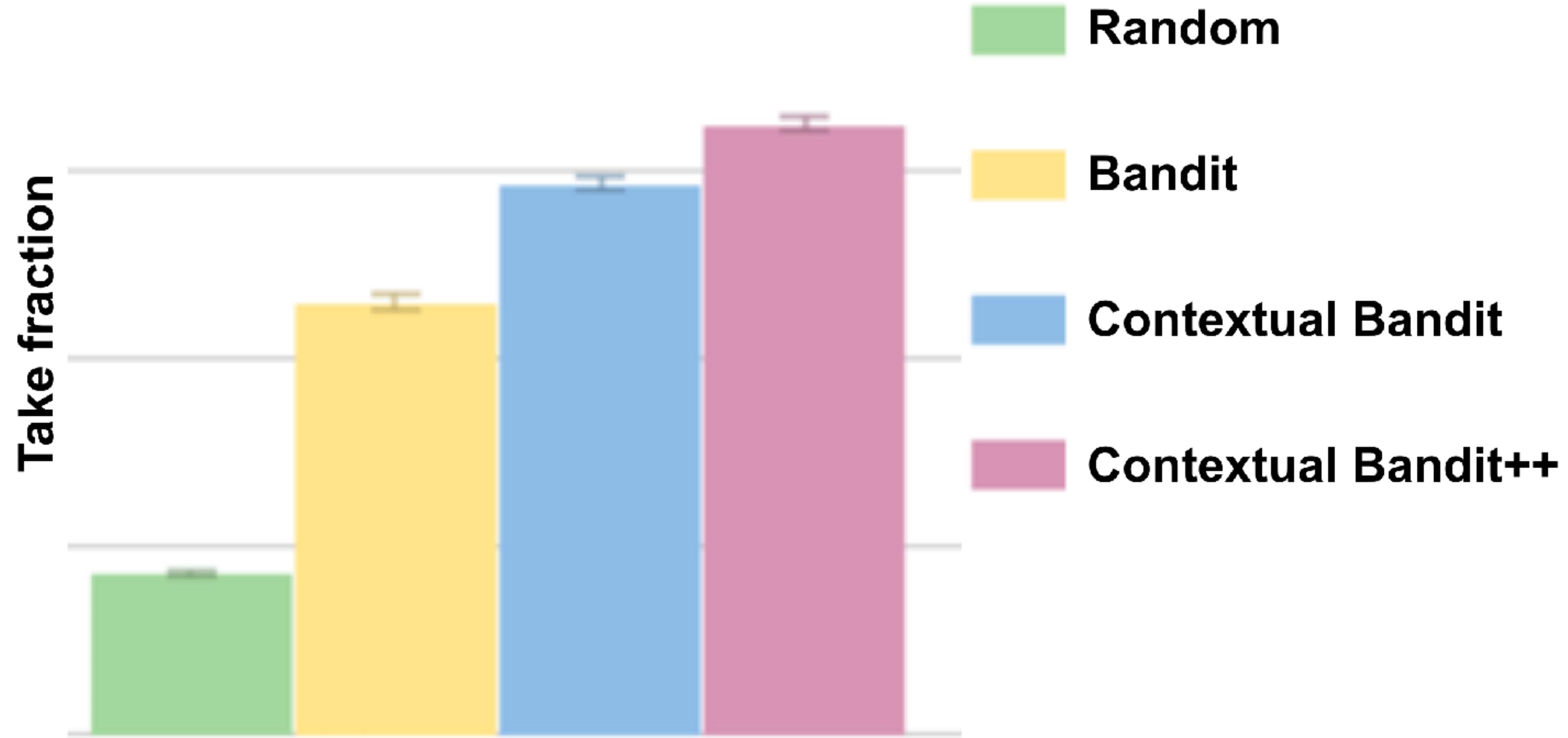
Profile Type	Score Image A	Score Image B
Comedy	5.7	6.3
Romance	7.2	6.5

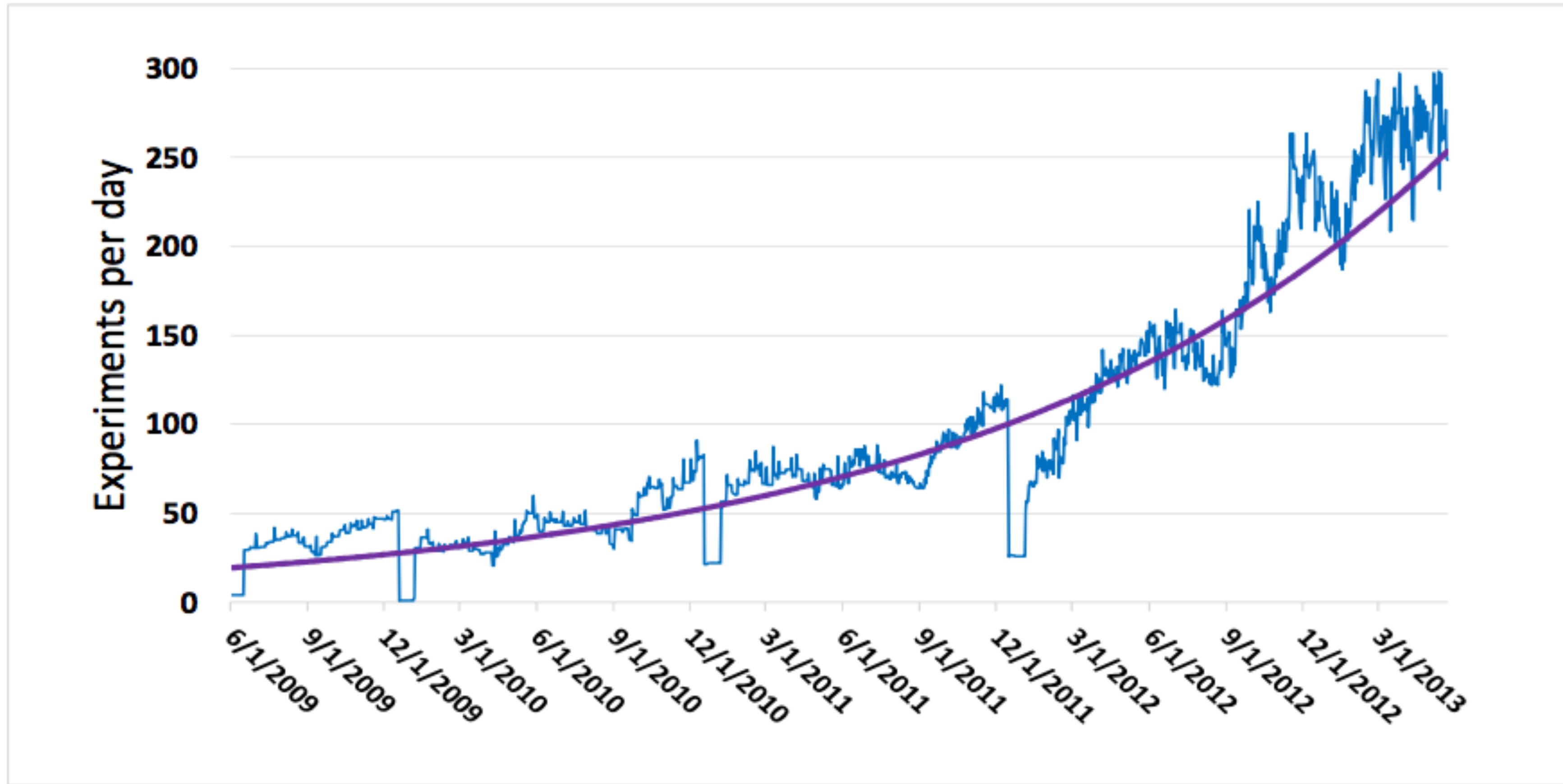


Image A



Image B





## Experiments Per Day on bing.com

Kohavi, R., Deng, A., Frasca, B., Walker, T., Xu, Y., & Pohlmann, N. (2013, August). **Online controlled experiments at large scale**. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1168-1176). ACM.

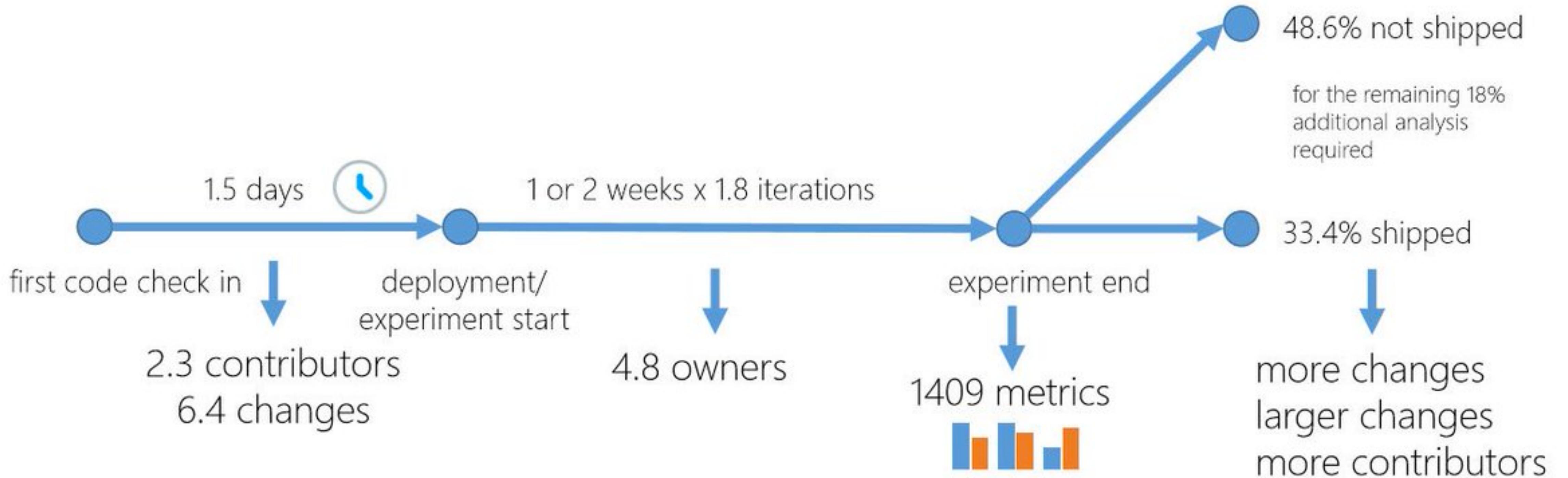


“ the number of concurrent experiments running in ERF has grown from a few dozen (in 2014) to **about 500 concurrent experiments** [May 2017]

Today we compute ~2500 distinct metrics per day and roughly **50k distinct experiment/metric combinations.**



Parks, Jonathan. Scaling Airbnb's Experimentation Platform. Airbnb Engineering & Data Science. May 10, 2017.



Kevic, K., Murphy, B., Williams, L., & Beckmann, J. (2017, May). **Characterizing experimentation in continuous deployment: a case study on bing.** In Proceedings of the 39th International Conference on Software Engineering: Software Engineering in Practice Track (pp. 123-132). IEEE Press.



“

[in] an entire redesign of the Premium Subscription's **payment flow** ...

The experiment showed **an increase of millions of dollars in annualized bookings**, about 30% reduction in refund orders and over 10% lift in free trial orders.

Xu, Y., Chen, N., Fernandez, A., Sinno, O., and Bhasin, A. **From infrastructure to culture: A/b testing challenges in large scale social networks.** In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM (2015), 2227–2236.



**Companies that test  
every business &  
product idea will be  
more successful**

**Companies that test  
every business &  
product idea will be  
better for society**



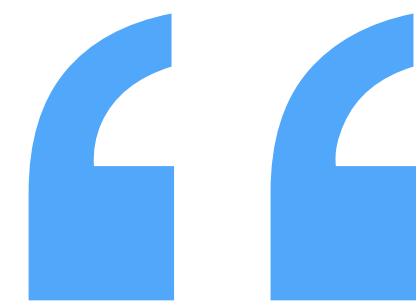
# How Uber Uses Psychological Tricks to Push Its Drivers' Buttons

The company has undertaken an extraordinary experiment in behavioral science to subtly entice an independent work force to maximize its growth.

“ Employing **hundreds of social scientists and data scientists**,

Uber has **experimented** with video game techniques, graphics and noncash rewards of little value that can **prod drivers into working longer and harder** — and sometimes at **hours** and **locations** that are **less lucrative** for them.

Scheiber, N., Huang, J. (2017) **How Uber Uses Psychological Tricks to Push Its Drivers' Buttons**. New York Times Magazine. April 2, 2017.



- A **20 per cent reduction in speeding** over six months
- A **34 per cent increase in acceptances** by students from under-represented schools to top universities.
- An **8% reduction in annual household gas consumption** following installation of smart heating controls
- A **38 per cent reduction in patient referrals** to over-booked hospitals

Behavioral Insights Team. [The Behavioral Insights Team Update Report 2016-2017](#).

THE  
BEHAVIORAL  
INSIGHTS TEAM.



Doleac, J. L., & Stein, L. C. (2013). *The visible hand: Race and online market outcomes*. The Economic Journal, 123(572), F469-F492.

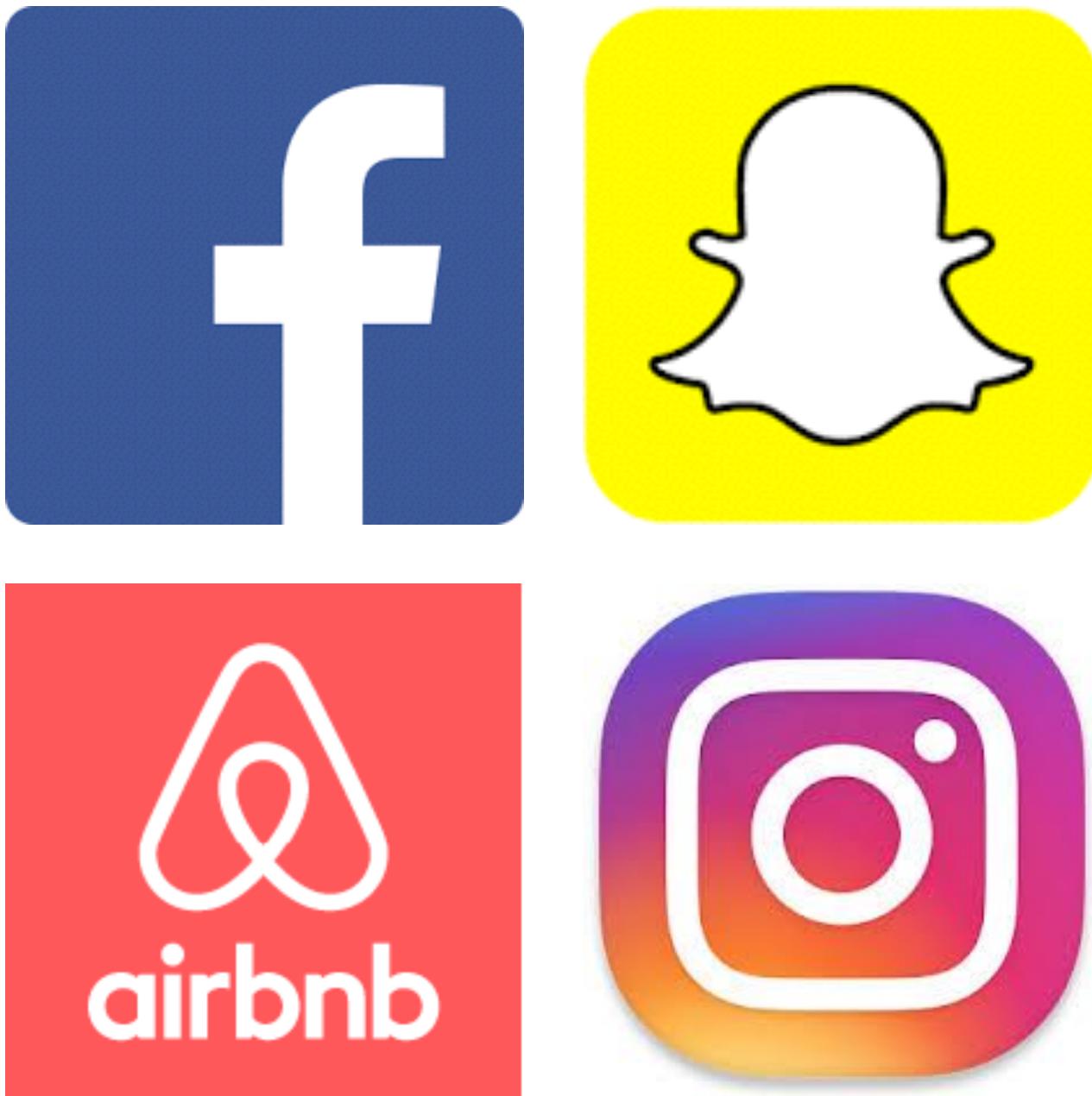
In local online classified advertisements throughout the US... **Black sellers do worse than white sellers** on a variety of market outcome measures:

- They receive 13% **fewer responses**
- 17% **fewer offers**
- Conditional on receiving at least one offer,
  - black sellers also receive 2-4% **lower offers**
  - buyers corresponding with black sellers exhibit **lower trust**:

Doleac, J. L., & Stein, L. C. (2013). *The visible hand: Race and online market outcomes*. The Economic Journal, 123(572), F469-F492.



THE  
BEHAVIORAL  
INSIGHTS TEAM.



Science



**Identify, analyze,  
evaluate** the  
content of field  
experiments in  
social science &  
CS

**Describe &  
Evaluate** the kinds  
of knowledge that  
experiments bring  
to policy & theory

**Design** a novel  
field experiment or  
**synthesize** novel  
results from past  
studies

**Summarize and  
criticize** leading  
views on the ethics  
& governance of  
experiments

**Design & analyze  
research** from the  
perspective of  
rapid A/B testing

# What You Need to Know

- **Statistics**
  - Linear regression (including multivariate)
  - Some experience with R

**February**

**Part I: Understanding Field Experiments**

- Experience designing an experiment
- Experience analyzing experiment results

**March**

**Part II: Planning Your Project**

**April - May**

**Part III: Deploying & Reporting Your Field Experiment**

- May 16:
- Group Results
- Group Analysis
- Personal Essay

# What to Expect Each Week (more or less)

## Lectures/Discussions (laptops often down):

- Professor-led or student-led discussion of an issue

## Workshops (laptops usually needed)

- Working through the craft of experimentation

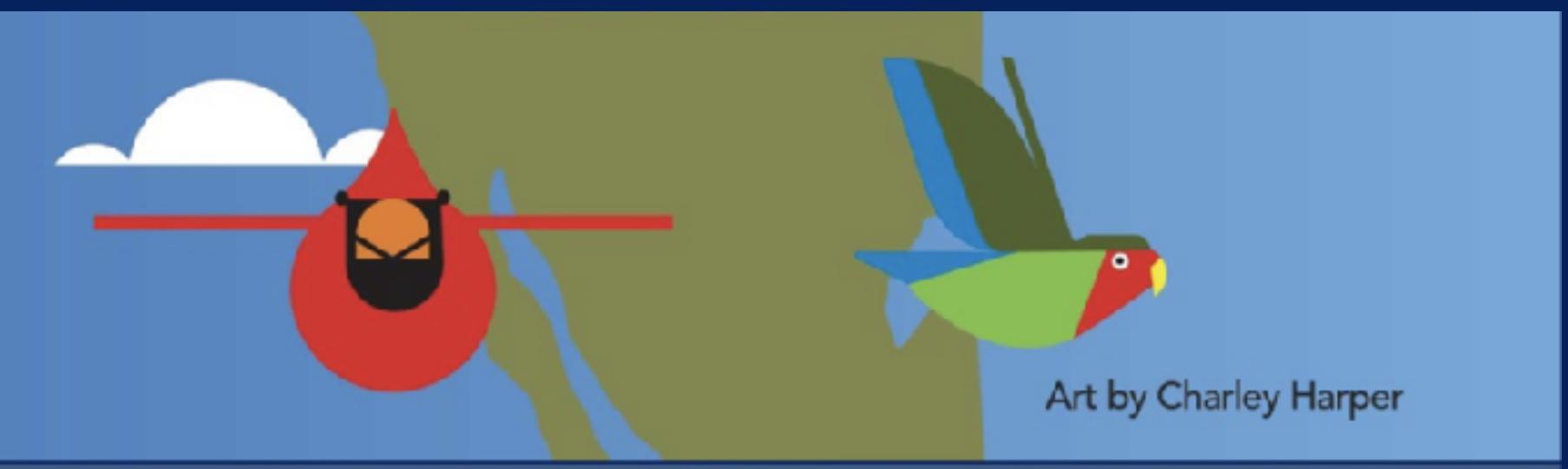
## Homework

- (first part) Analysis/writing assignment (**pairs**)
- (second/third part) Team progress
- Comments & reflection on slack (**individual**)

# Grading Rubric

- **Discussions & Workshops (20%)**
  - Did you attend?
  - Did you add meaningfully to the conversation?
  - Did you present at least once?
- **Weekly Assignments (20%)**
  - Split among Implementation, Analysis, and Writing
- **Midterm Project Proposal (20%)**
- **Final Project (40%)**

The Great Backyard  
Bird Count



Art by Charley Harper

**February 14-17, 2020**

[www.BirdCount.org](http://www.BirdCount.org)





Week 1: Introducing COMM 4940: The Design & Governance of Field Experiments

# Final Project

The final project will involve analyzing results from a field experiment and writing the results for a public audience.

The final can be a project of your own based on the Upworthy Archive, or with permission, you can design a study with a community partner.





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.



“

The most important impacts of digital technology lie not in the capacity of disorganized masses to more easily speak, but the capacity of new civil society organizations to more effectively listen

Karpf, D. (2016). **Analytic activism: Digital listening and the new political strategy.** Oxford University Press.

# Upworthy cofounder Peter Koechley apologizes for the clickbait 'monster' his site unleashed on the internet

Lara O'Reilly Mar 19, 2015, 8:27 AM



Upworthy's cofounder Peter Koechley has an apology to make.

Speaking at The Guardian's Changing Media Summit in London on Thursday, Koechley apologized for what made the three-year-old site [one of the fastest growing of all time](#): its bombastic, sensationalized, clickbait headlines such as "[This Amazing Kid Just Died, What He Left Behind Was Wondtacular](#)" and "[His First 4 Sentences Are Interesting. The 5th Blew My Mind. And Made Me A Little Sick.](#)"



Upworthy's cofounder spoke at The Guardian's Changing Media Summit in London. Business Insider/Lara O'Reilly

Those types of headlines were punished by a change in Facebook's algorithm last year, which sought to clean up the social-media site's News Feed to surface more "high quality" stories, resulting in [huge drops of traffic for Upworthy and similar sites](#).

Koechley said: "We sort of unleashed a monster. Sorry for that. Sorry we kind of broke the internet last year. I'm excited going forward to say goodbye to clickbait."

# The Upworthy Research Archive

Agreement to Study, Publish, and  
Release the Full Archive for Research  
and Educational Purposes



## “Package”

- Headline
- Subhead
- Thumbnail
- Social Media summary
- (sometimes)

The screenshot shows the Upworthy homepage with a large orange "Hi!" button on the left. The main headline reads "We're Upworthy." Below it, a subhead states: "We're a social media outfit with a mission: To make meaningful stories go viral." A text block follows: "Here's *AllThingsD*'s take on our first 6 months, and a *New York Times* piece on our first 100 days." A "JOIN US, AND NEVER MISS OUR BEST CONTENT:" section includes links for Facebook, Email List, and Tumblr.

**Featured Stories:**

- Open Your Mouth Now - Your Jaw Is About To Drop** by Femi Oke (Nov 19 2012). Includes a photo of a person on a bike and a video thumbnail.
- This African Teenager Taught Himself...** by Carolyn Silveira.
- WATCH: Girl Tells Boy She Has HIV. Boy's Reaction...** by Carolyn Silveira.

The screenshot shows the Upworthy homepage from April 2014. The main headline is "UPWORTHY". A search bar and navigation menu are at the top. A sidebar on the right says "Things that matter. Pass 'em on." An article about Lorde is circled in orange. Below are three main stories: "A Woman Shares Her Day Through Google Glass. It Seems Lovely And Ordinary — Until The End", "A Guy Sells Something So Shocking On Wall Street That A Crowd Stops And Stares", and "Lorde Found A Photoshopped Photo Of Herself And Decided To Tweet An Untouched Real One Next To It".

**Community**

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

**Entertainment & Culture**

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

**Health**

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

## Outcome information:

- Impressions
- Clicks
- “Winner”

**Some Jerk Corporation Patented Your Genes.  
Literally. I'm Not Joking.**

**Your Body Was Just Patented By A  
Corporation. You. Your Specific Body. I'm  
Not Joking.**

**Creepy Corporation Does Something So  
Cartoonishly Evil That We Have To Make A Dr.  
Evil Reference**

**The One Where A Corporation Thinks They Can  
Control What You Do With Your Own Body**

**32,487**

**Valid Deployed Tests**

**25,566**

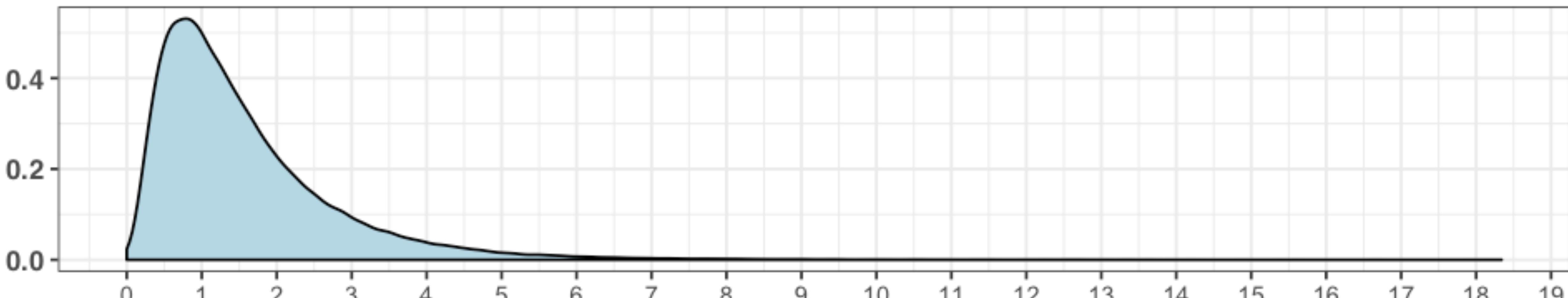
**Tests That Compare  
Headlines**

**9,158**

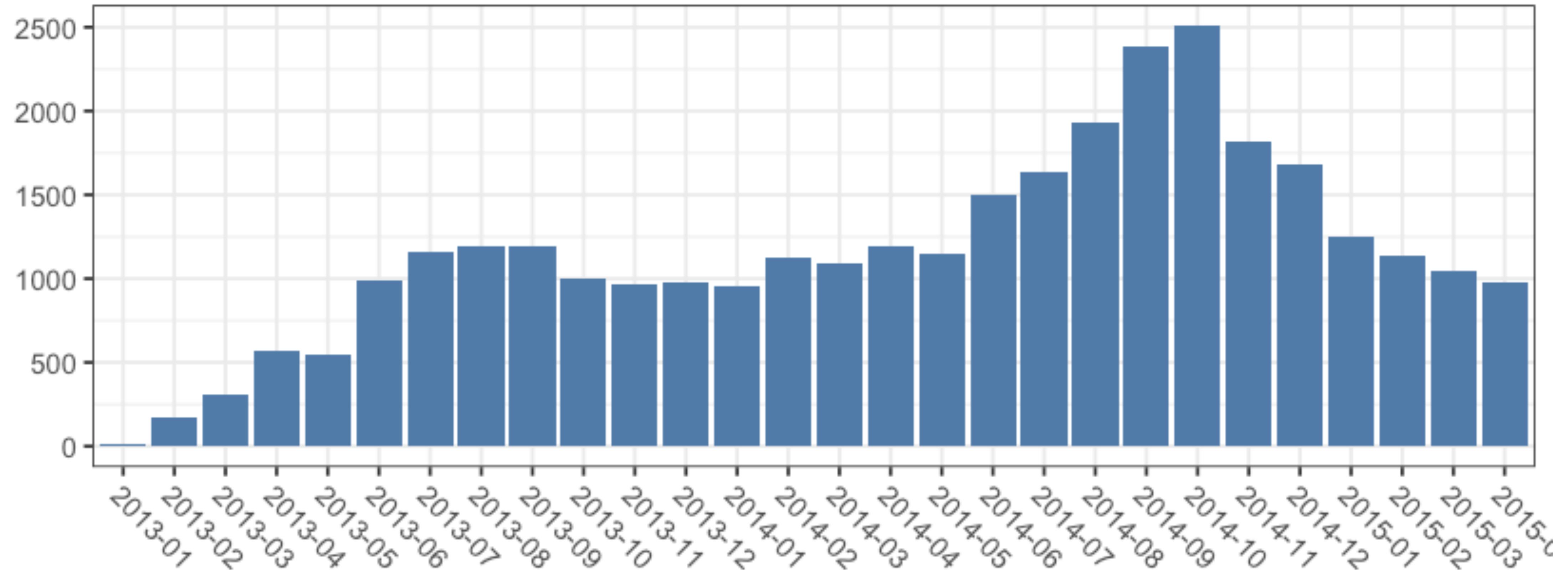
**Recorded Decisions**

**150,817**

**Total Tested Packages**



Clickthrough rate on Upworthy.com across 150,817 packages from 2013-01-23 to 2015-11-02.  
Data analysis by J. Nathan Matias and Marianne Aubin Le Quere



A/B tests per month conducted by Upworthy between 2013-01-24 and 2015-04-30.  
Analysis by J. Nathan Matias and Marianne Aubin Le Quere



Analytics

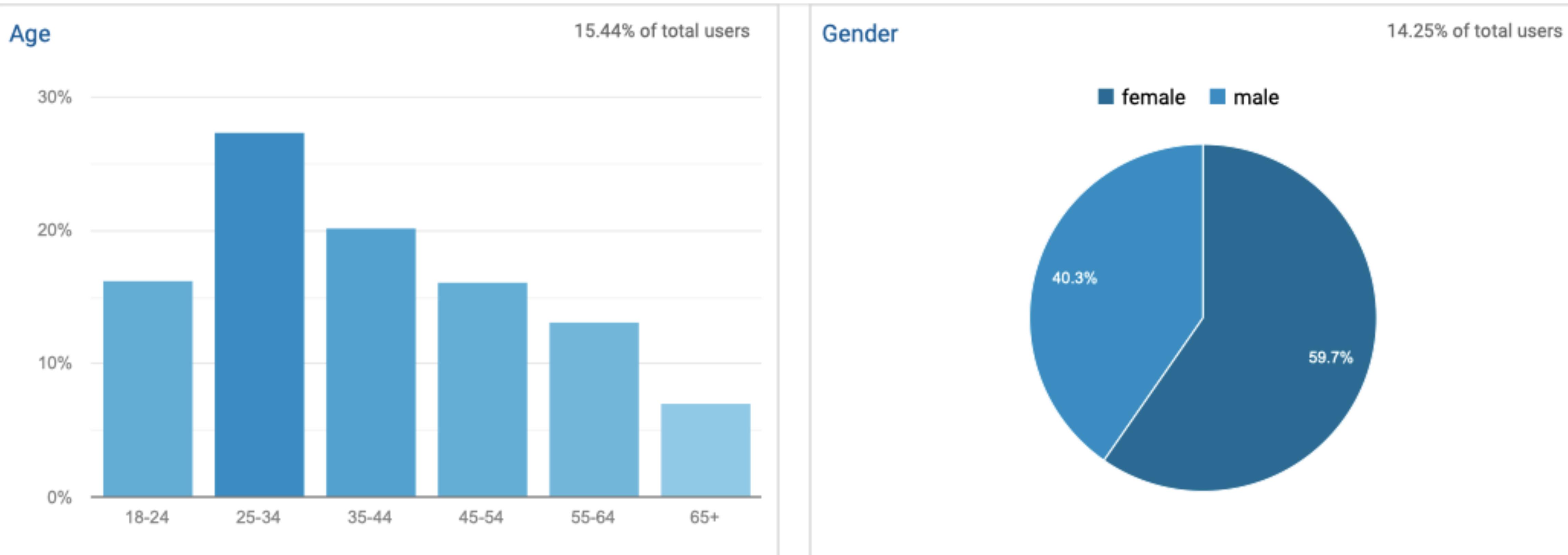
Upworthy  
Upworthy[Go to report](#)

## Demographics: Overview

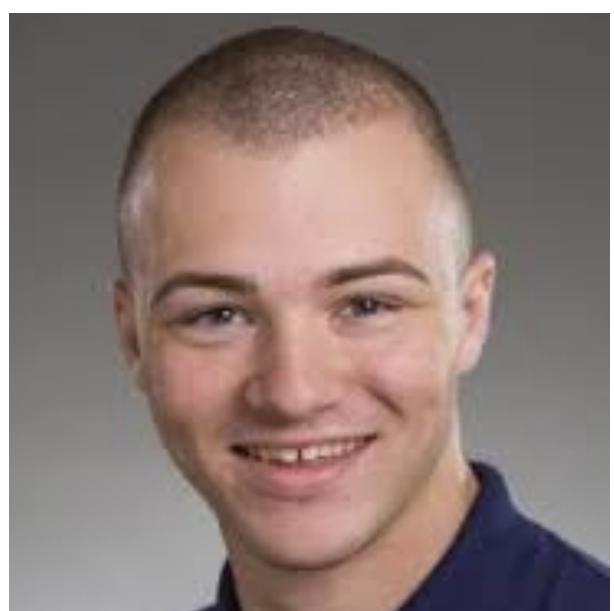
All Users  
100.00% Users

Jan 1, 2013 - Dec 31, 2015

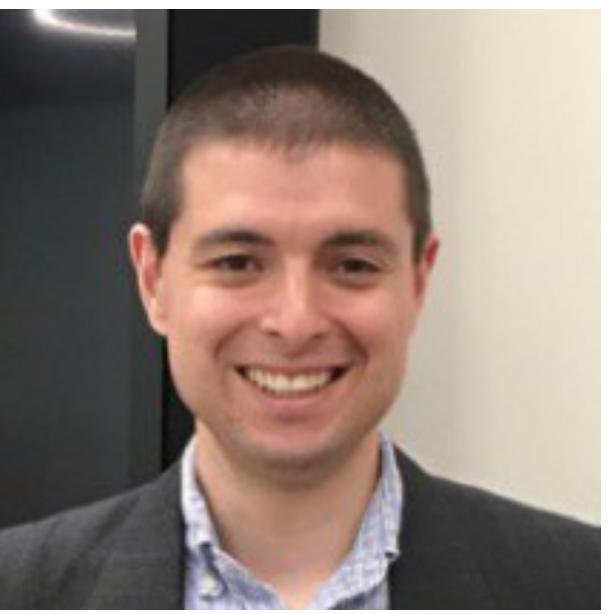
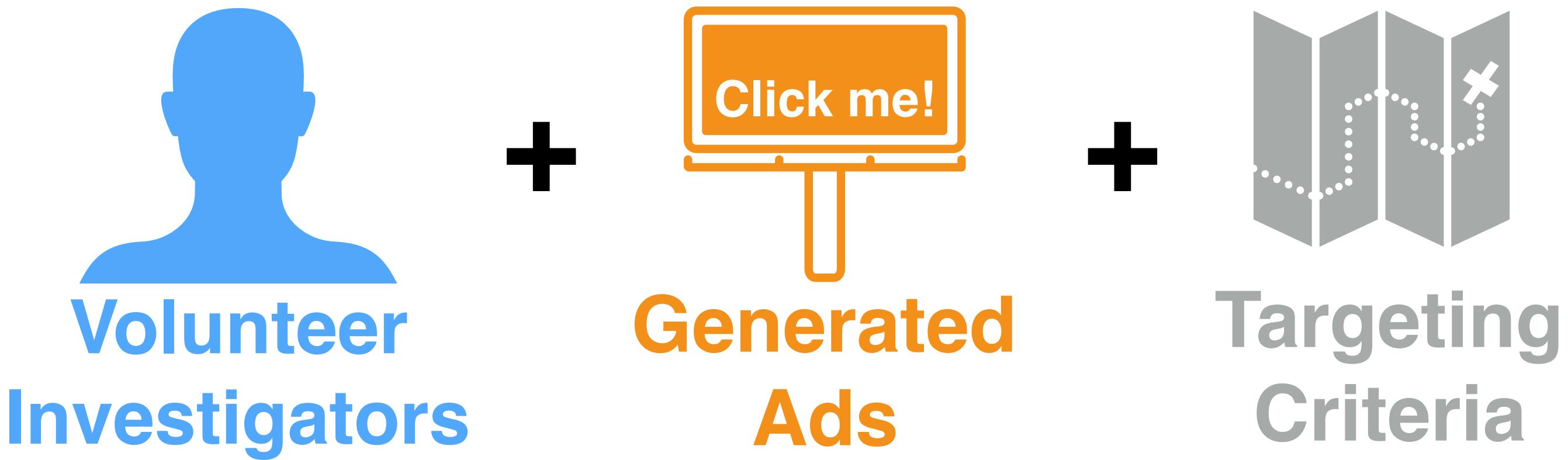
Key Metric:



# Auditing How Tech Companies Enforce Political Ad Policies



Austin Hounsel

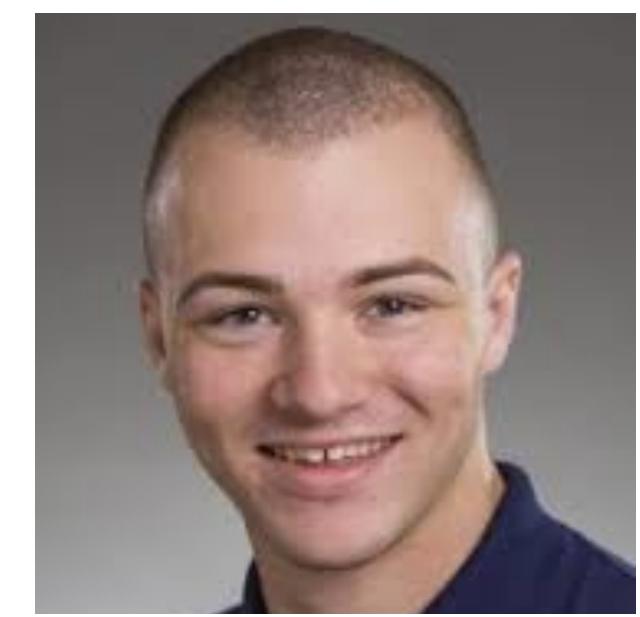


J. Nathan Matias

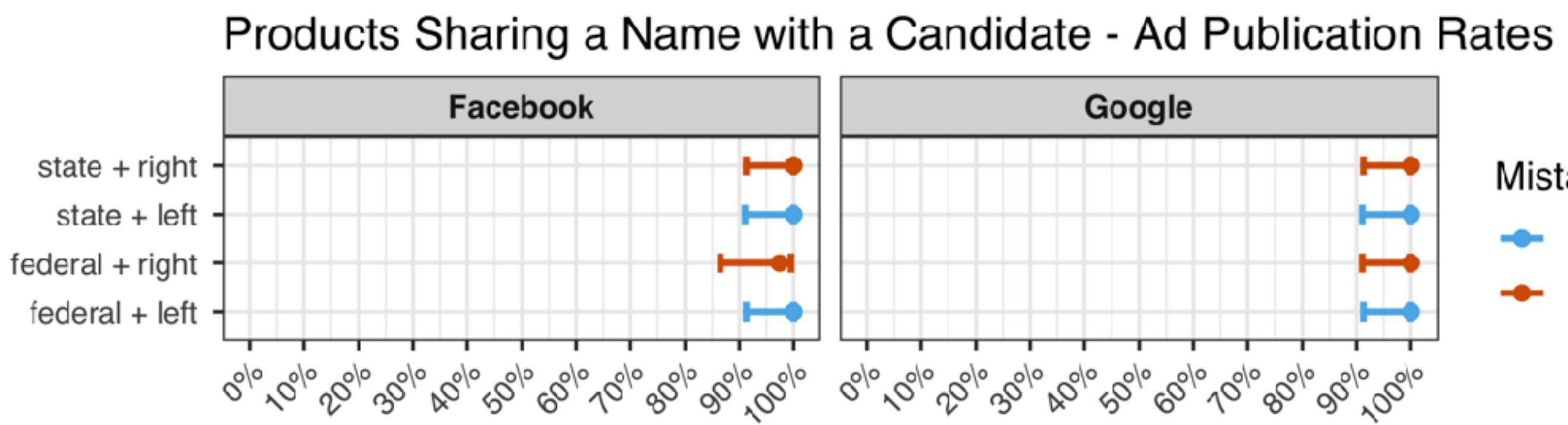
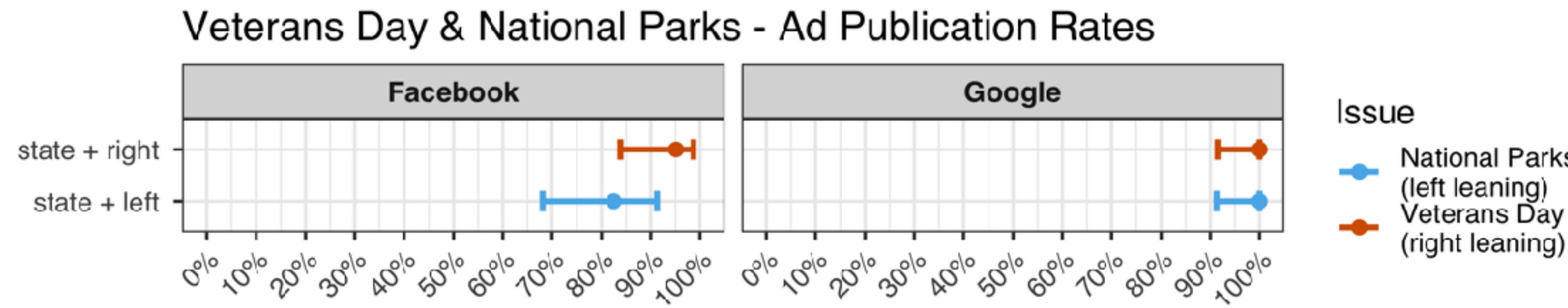


Nick Feamster

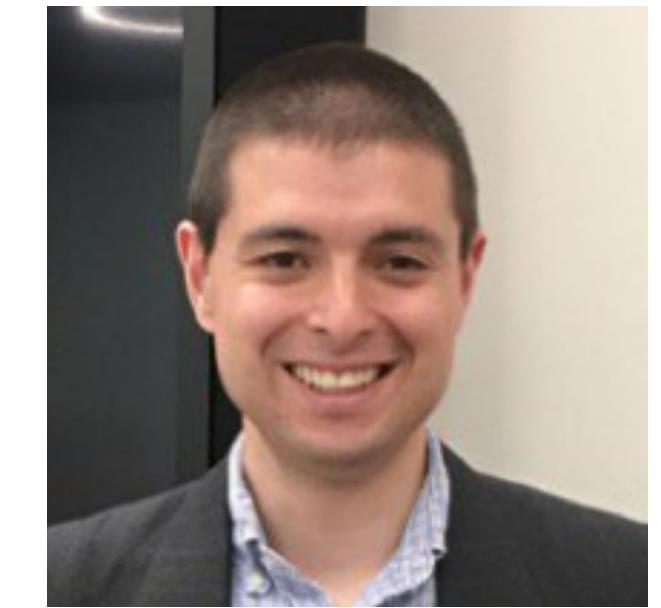
# Auditing How Tech Companies Enforce Political Ad Policies



**Austin Hounsell**



Estimated chance of publication for a given ad combination(election, leaning). Ad placement attempts: 477 ads placed by 7 people from 2018-09-17 to 2018-10-10. Product ads are music albums that share a word with a candidate name. Veterans Day & National Park ads are about events and places that could be mistaken by platform policy enforcers as election-related ads of national importance. 95% confidence intervals use the Wilson method. Code & data at: <https://github.com/citp/mistaken-ad-enforcement>  
Data and analysis by J. Nathan Matias & Austin Hounsell of Princeton University, with Melissa Hopkins, Ben Wermuth, Jason Griffey, Chris Peterson, Scott Hale, and Nick Feamster



**J. Nathan Matias**



**Nick Feamster**

# The Gray Phone Challenge



**Zenobia Chan**



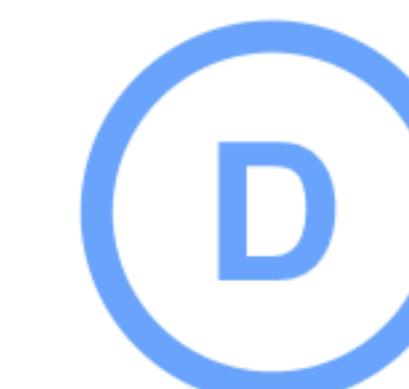
**J. Nathan Matias**

# Gray Phone Challenge Daily Routine

Checking-in just takes a minute before bed or in the morning.  
Time-tracking resets at midnight.



Zenobia Chan

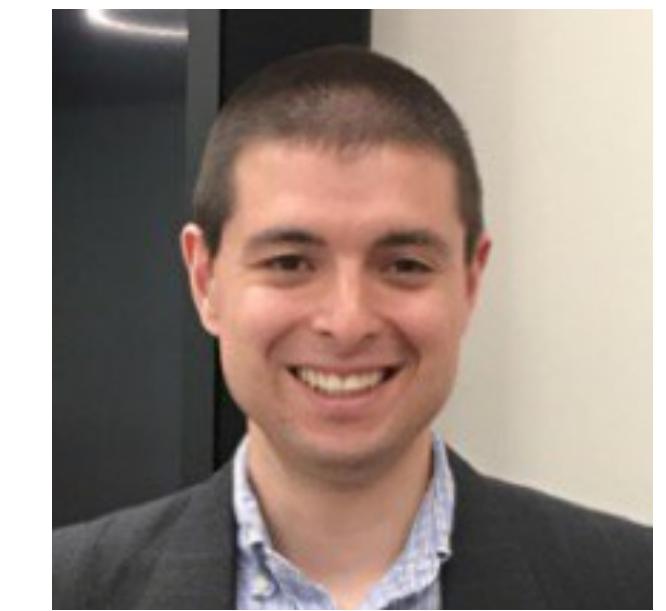


**Receive  
Daily Message**

**Adjust  
Color**

**Confirm  
Daily Status**

**Stick with  
this all day :-)**



J. Nathan Matias

# Promoting Inclusion and Participation in an Online Gender-Related Discussion Community

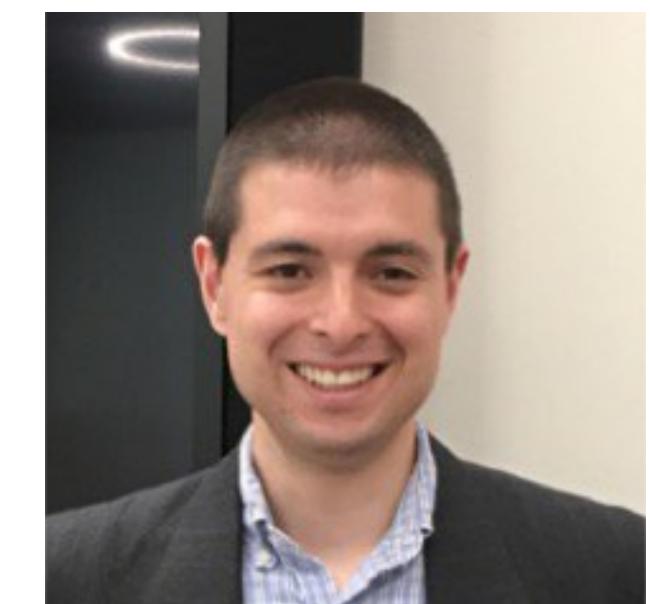
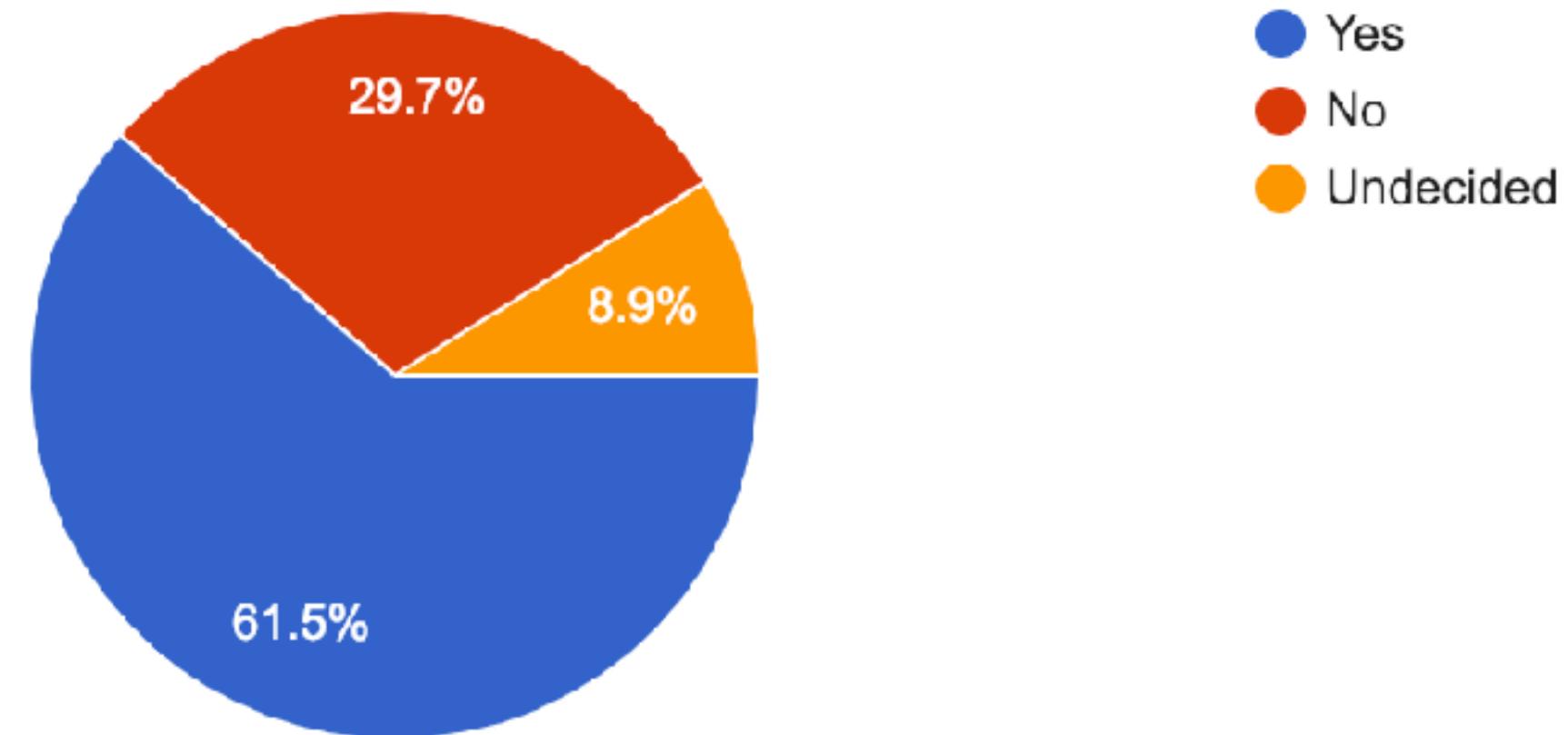


Tyler Simko



Emily Hedlund

Did you identify as a feminist at the time you made your first comment?



J. Nathan Matias

# Office Hours

**Monday, 10am - 12:30am (15 & 30 minutes)**

**Wednesday, 10am - 12:30am (15 & 30 minutes)**

**Mann 486**

**<https://calendly.com/natematias/officehours>**

# IRB Training

**Essential.** If you do not have certification for Human Subjects Research, please take this course ASAP, as your participation in the class will depend on this. It may take ~3-4 hours.

**“Social and Behavioral Researchers”**

**[https://researchservices.cornell.edu/  
resources/irb-training](https://researchservices.cornell.edu/resources/irb-training)**

# Next Steps

- Complete the Survey at:
  - <https://forms.gle/12dsK3K4XsCZNq178>
- Join Slack
- Read:
  - Linden 2006
  - Shadish, Cook, and Campbell
- If necessary, complete Cornell IRB training by Tuesday Feb 12