

A Vocal Minority:

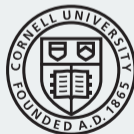
Assessing the Representativeness of Tweeters and Tweets

July 17, 2020

Adam Hughes, Stefan McCabe, Will Hobbs, Emma Remy, Sono Shah, David Lazer



Pew Research Center



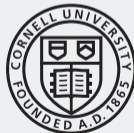
A Vocal Minority: Assessing the Representativeness of Tweeters and Tweets

July 17, 2020

Adam Hughes, **Stefan McCabe**, Will Hobbs, Emma Remy, Sono Shah, **David Lazer**



Pew Research Center



A Vocal Minority:

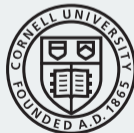
Assessing the Representativeness of Tweeters and Tweets

July 17, 2020

Adam Hughes, Stefan McCabe, Will Hobbs, **Emma Remy**, **Sono Shah**, David Lazer



Pew Research Center



A Vocal Minority:

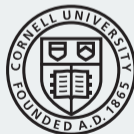
Assessing the Representativeness of Tweeters and Tweets

July 17, 2020

Adam Hughes, Stefan McCabe, **Will Hobbs**, Emma Remy, Sono Shah, David Lazer



Pew Research Center





Kevin Munger ✓

@kmmunger

How badly are scholars of social media starved for descriptive information? There are 1,700 citations to this 3,500 respondent survey by Pew in 2014.

Social media update 2014

M Duggan, [NB Ellison](#), [C Lampe](#)... - **Pew** ..., 2015 - foothillspresbytery ...

([HTTP://WWW.PEWINTERNET.ORG/AUTHOR/MMADDEN/](http://WWW.PEWINTERNET.ORG/AUTHOR/MMADDEN/))(http://www.pewinternet.org/2015/01/09/socialmediaupdate2014/pi_20150109_socialmedia_01/) In a new survey conducted in September 2014, the **Pew** Research Center finds that Facebook remains by far ...

☆ [🔗](#) Cited by 1704 [Related articles](#) [View as HTML](#) [🔗](#)

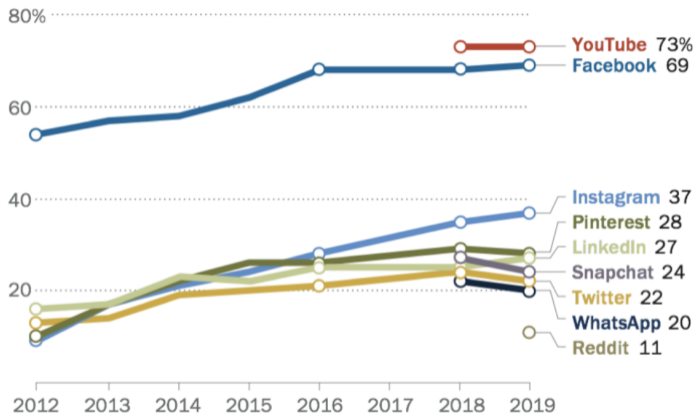
3:49 PM · Jul 25, 2019 · [Twitter Web App](#)

18 Retweets 112 Likes



Facebook, YouTube continue to be the most widely used online platforms among U.S. adults

% of U.S. adults who say they ever use the following online platforms or messaging apps online or on their cellphone



Note: Pre-2018 telephone poll data is not available for YouTube, Snapchat and WhatsApp.

Comparable trend data is not available for Reddit.

Source: Survey conducted Jan. 8-Feb. 7, 2019.

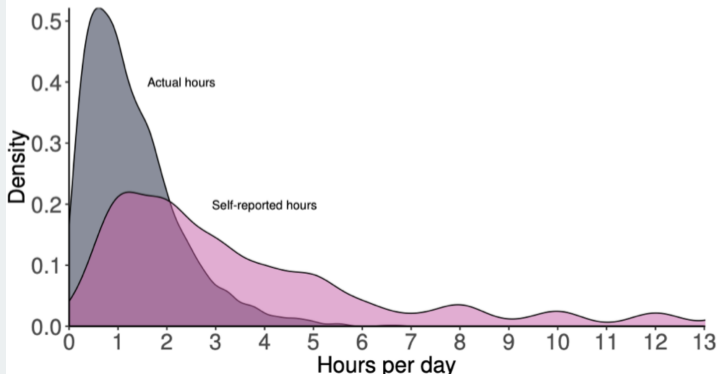
How Well Do People Report Time Spent on Facebook? An Evaluation of Established Survey Questions with Recommendations

Sindhu Kiranmai Ernala
Georgia Tech
sernala3@gatech.edu

Moira Burke
Facebook
mburke@fb.com

Alex Leavitt
Facebook
alexleavitt@fb.com

Nicole B. Ellison
University of Michigan
enicole@umich.edu



Research Question

Do different methods of **linking** individuals with their social media behavior show consistent results?

Research Question

Do different methods of **linking** individuals with their social media behavior show consistent results?

Approaches:

- Self-report
- Record linkage



Four data sources:

- Survey 1: KnowledgePanel
- Survey 2: American Trends Panel
- RDD phone poll
- Voter file + Twitter Decahose

- November 2018–December 2018
- N = 3,649 users say they use Twitter
- 76% provided valid Twitter handles
- Weighted to US population on Twitter based on RDD poll

American Trends Panel

- October 2019–November 2019
- N = 2,561 users say they use Twitter
- 53% provided valid Twitter handles

- Benchmark: N=1,502 users who report using Twitter

- Population of registered American voters from late 2017
- Identities deterministically matched with Twitter profiles on name and self-reported location
- N = 1.5 million
- Voter file sample contains 5% of KP panelists and 6% of ATP panelists

Kenny Joseph
 @kenny_joseph

postdoc @NUnetsci . former @SCSatCMU, @datascifellows, Wolverine, Buffalonian. Yes, I'm still a Bills fan.

Boston, MA

Brennan Klein
 @jkbren FOLLOWS YOU

thinking, thinking about thinking, thinking about thinking about thinking. phd student @NUnetsi

Nir Grinberg
 @grinbergnir FOLLOWS YOU

Phd student @cornell (@cornell_tech), interested in Machine Learning, Statistics, NLP and Social Media, not necessarily in that order.

New York, NY

No match, not in voter records

Voter registrations

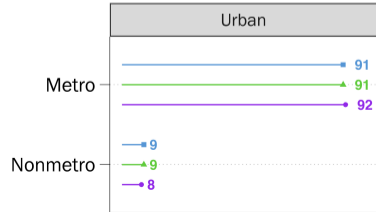
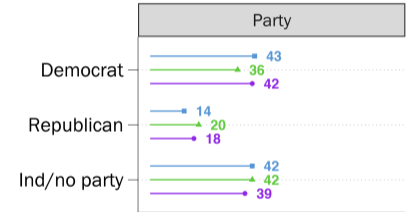
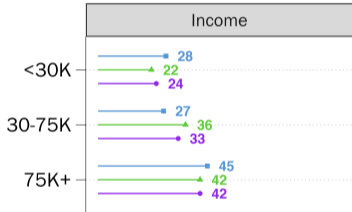
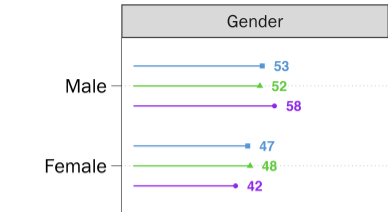
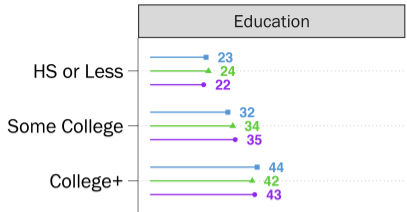
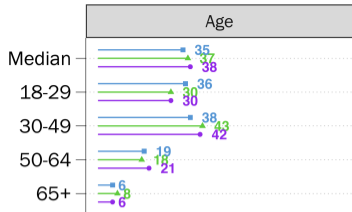
Name	State	Party	Age	Match to Twitter?	...
Kenny Joseph	MA	D	52	Yes!	...
Nir Grinberg	NY	D	19	No, not unique	...
Nir Grinberg	NY	R	25	No, not unique	...
Amelia Tzray	CA	No, not on Twitter	...



Fake news on Twitter during the 2016 U.S. presidential election

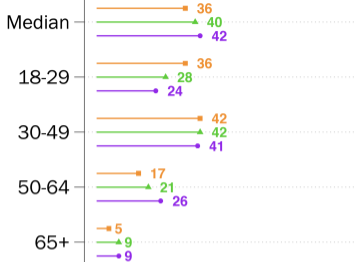
**Nir Grinberg^{1,2*}, Kenneth Joseph^{3*}, Lisa Friedland^{1*},
Briony Swire-Thompson^{1,2}, David Lazer^{1,2†}**

The spread of fake news on social media became a public concern in the United States after the 2016 presidential election. We examined exposure to and sharing of fake news by registered voters on Twitter and found that engagement with fake news sources was extremely concentrated. Only 1% of individuals accounted for 80% of fake news source exposures, and 0.1% accounted for nearly 80% of fake news sources shared. Individuals most likely to engage with fake news sources were conservative leaning, older, and highly engaged with political news. A cluster of fake news sources shared overlapping audiences on the extreme right, but for people across the political spectrum, most political news exposure still came from mainstream media outlets.

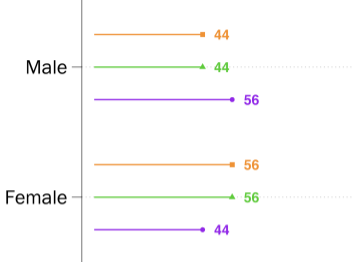


● American Trends Panel
▲ Knowledge Panel
■ RDD Phone Poll

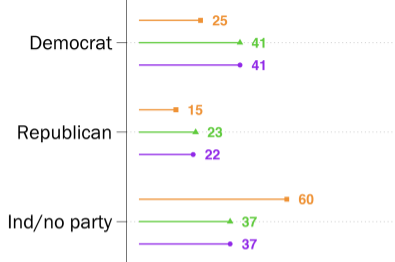
Age



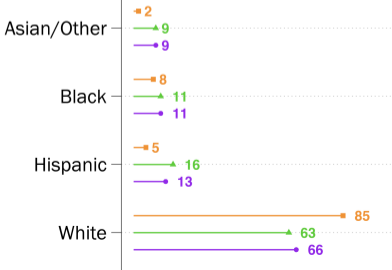
Gender



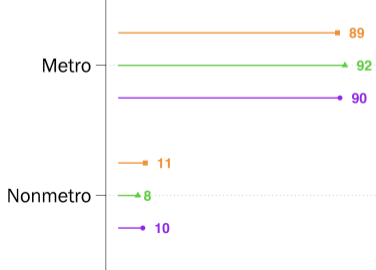
Party



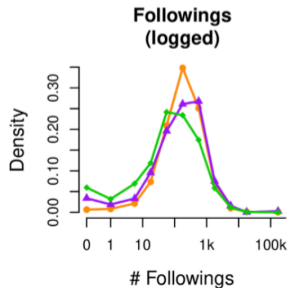
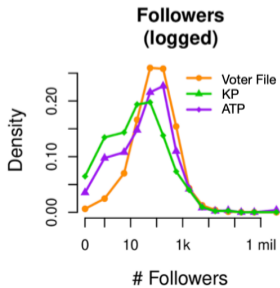
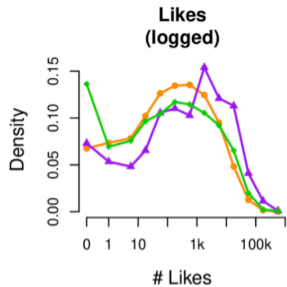
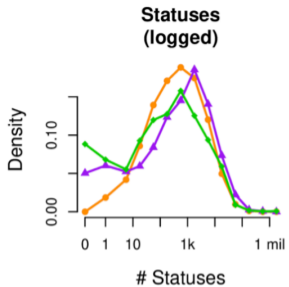
Race/Ethnicity



Urban



- American Trends Panel
- KnowledgePanel
- Voter File



Behavioral Measures

For tweets between October 1 and November 30, 2019:

Behavioral Measures

For tweets between October 1 and November 30, 2019:

Total tweets:

- KP: 118K tweets from 1100 users
- ATP: 96K tweets from 800 users
- VF: 46M tweets from 600K users

Behavioral Measures

For tweets between October 1 and November 30, 2019:

Total tweets:

- KP: 118K tweets from 1100 users
- ATP: 96K tweets from 800 users
- VF: 46M tweets from 600K users

Gini:

- KP: 0.87
- ATP: 0.84
- VF: 0.81

Behavioral Measures

For tweets between October 1 and November 30, 2019:

Total tweets:

- KP: 118K tweets from 1100 users
- ATP: 96K tweets from 800 users
- VF: 46M tweets from 600K users

Gini:

- KP: 0.87
- ATP: 0.84
- VF: 0.81

Number of users with any tweet mentioning “impeach”:

- KP: 200
- ATP: 170
- VF: 88K

- User-first approach to collecting data
- Surveys essential for benchmarking
- Bias-variance tradeoff

Thanks!

sdmccabe.github.io

mccabe.s@northeastern.edu