

# **IN4MATX 153: CSCW**

**Class 19:**  
**Misinformation**

Professor Daniel Epstein  
TA Dennis Wang  
Reader Weijie Du

# Announcements

- Final exam will be posted right after class today
  - Due 11:59pm Wednesday June 12
- A3 grading is underway, probably will be released Monday
- No readings for Wednesday, June 5
  - Sorry one of today's readings was messed up!
- No office hours next week
  - Email the class list if you have questions about the exam

# Today's goals

By the end of today, you should be able to...

- Differentiate misinformation from disinformation
- Reason through theories around why misinformation spread is more prominent in the digital era
- Articulate the role of varied stakeholders such as conventional media, bystanders, and fact checkers in advancing and minimizing spread
- Explain technological design strategies for minimizing the spread of misinformation on social media

**Social media is at the center of the conversation about misinformation**



# Misinformation

Research shows the emergent video platform  
conspiracy theories and other ha

# The New York Times

## THE INTERPRETER

# *'Belonging Is Stronger Than Facts': The Age of Misinformation*

Social and psychological forces are combining to make the sharing and believing of misinformation an endemic problem with no easy solution.



SUBSCRIBE

≡ CNN BUSINESS ▾

Markets →

DOW	33,426.63	0.33% ▼
S&P 500	4,191.98	0.14% ▼
NASDAQ	12,657.90	0.24% ▼

TikTok's search engine delivers misinformation to young user base, report says

# Social media and misinformation

- Social media didn't create misinformation
  - But it lowered the friction associated with sharing
  - It made it easier for groups to find each other and share disinformation
- And even if it didn't start with social media, social media bears responsibility for what happens in those spaces

# Social media and misinformation

- There's reason to be hopeful
  - “*The digitization of information exchange, however, also makes the practices of disinformation detectable, the networks of influence discernable, and suspicious content characterizable.*”



## An Agenda for Disinformation Research

A Computing Community Consortium (CCC) Quadrennial Paper

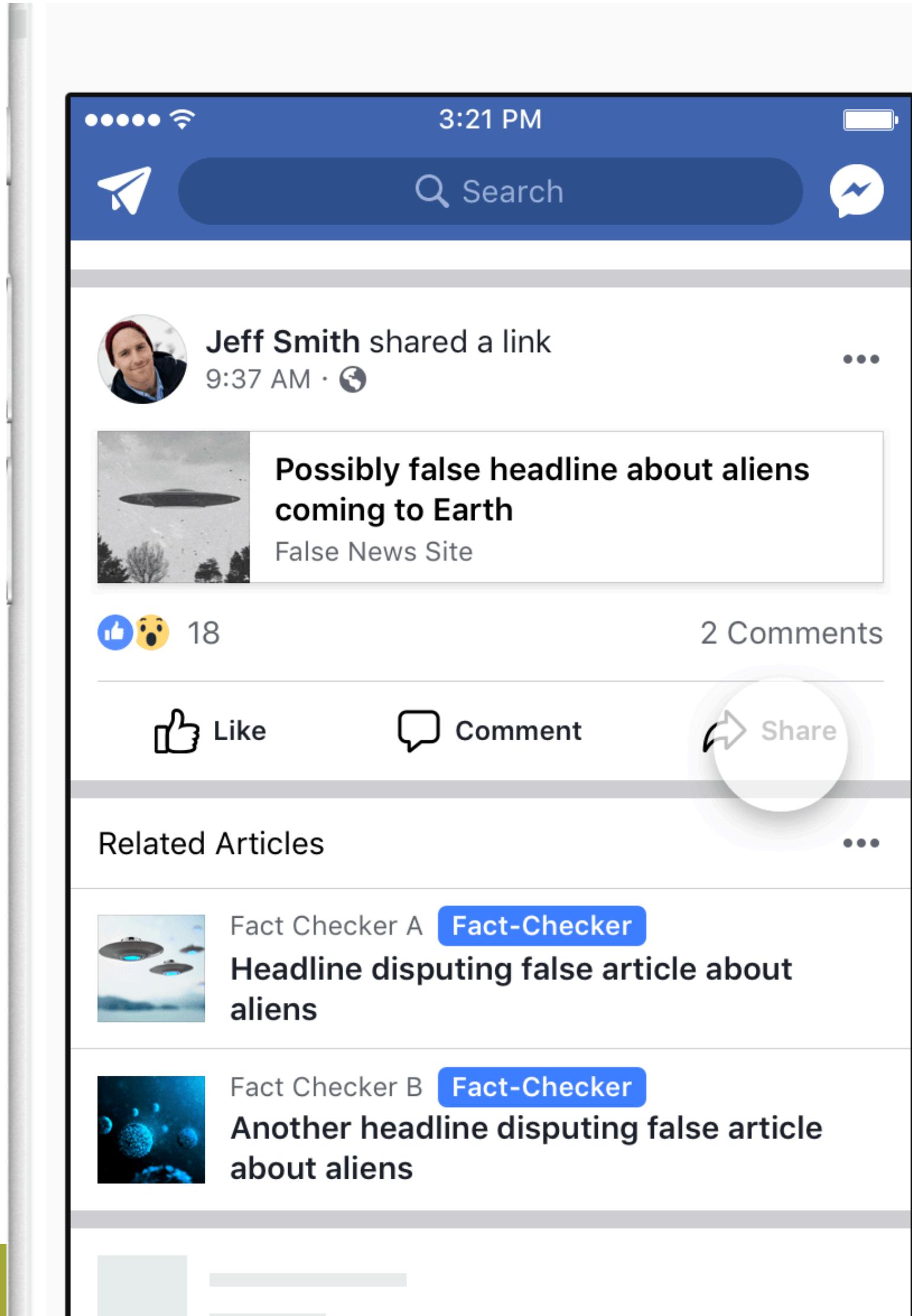
Nadya Bliss (Arizona State University), Elizabeth Bradley (University of Colorado, Boulder), Joshua Garland (Santa Fe Institute), Filippo Menczer (Indiana University), Scott W. Ruston (Arizona State University), Kate Starbird (University of Washington), and Chris Wiggins (Columbia University)

# Outline

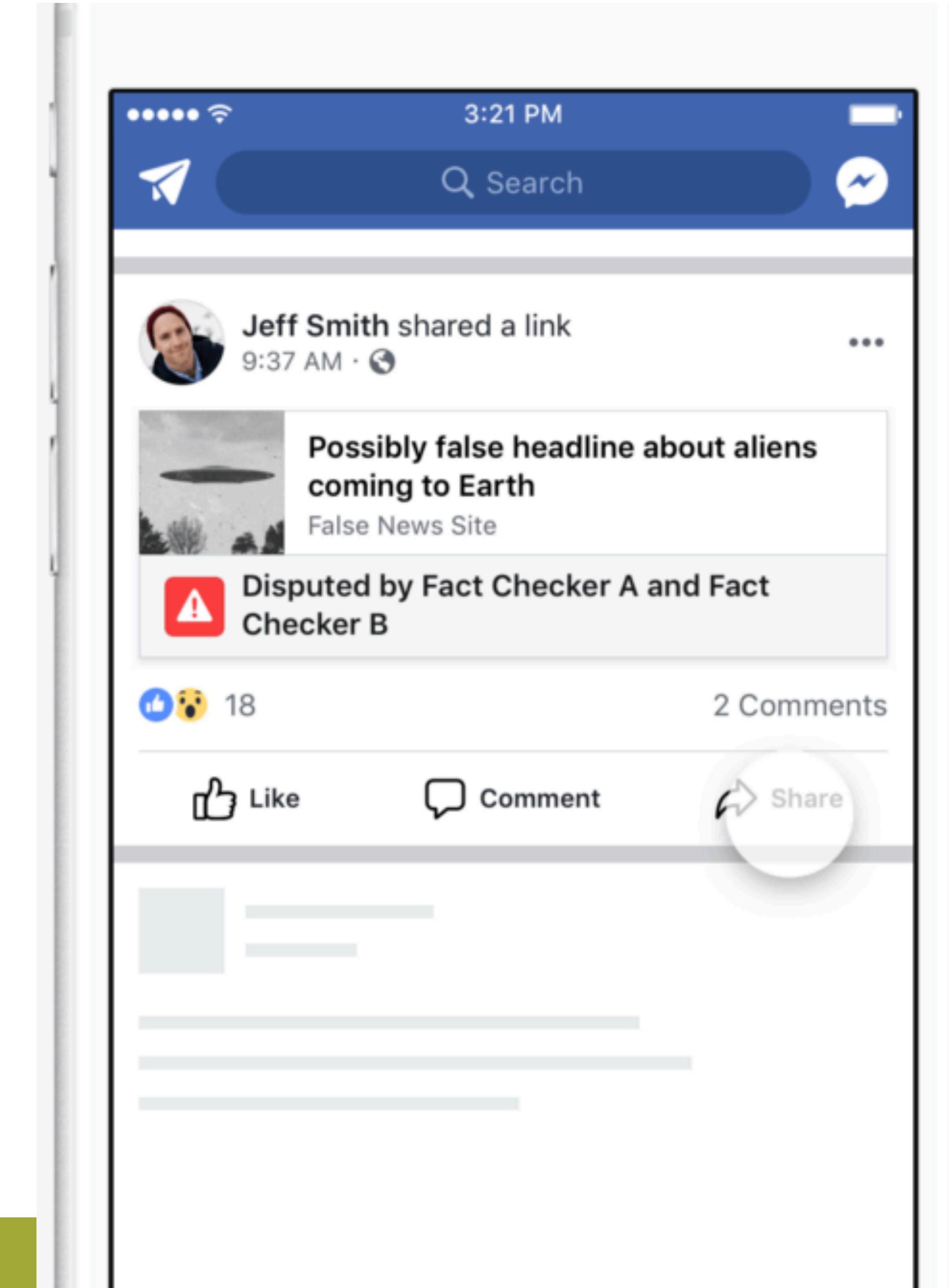
- Misinformation also touches on politics, communication, psychology, and other areas
- We'll focus on how it relates to social media
  - How does misinformation originate and propagate?
  - How can we detect misinformation?
  - How can design help?

# Which of these will reduce misinformation more?

## Related articles



## Fact check

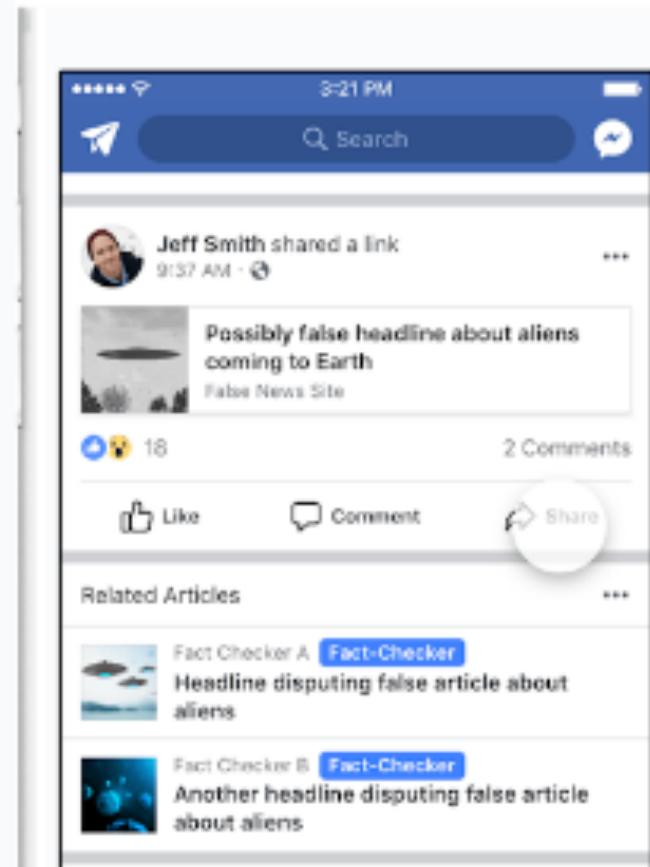


# Which of these will reduce misinformation more?

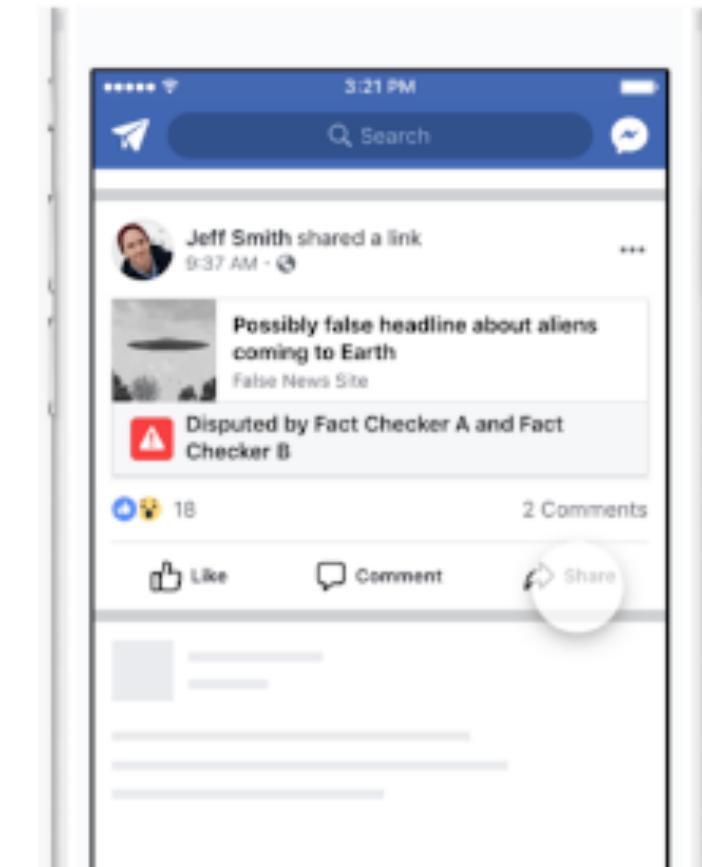
0%

0%

## Related articles



## Fact check



Related articles

Fact check

# Terms

- Misinformation versus disinformation
- Misinformation is **anything false**
  - Might be a rumor, or something not necessarily intentionally false
- Disinformation has the **specific intent is to deceive**
  - Often built around a true or plausible core, wrapped up in a misleading way

# Why now?

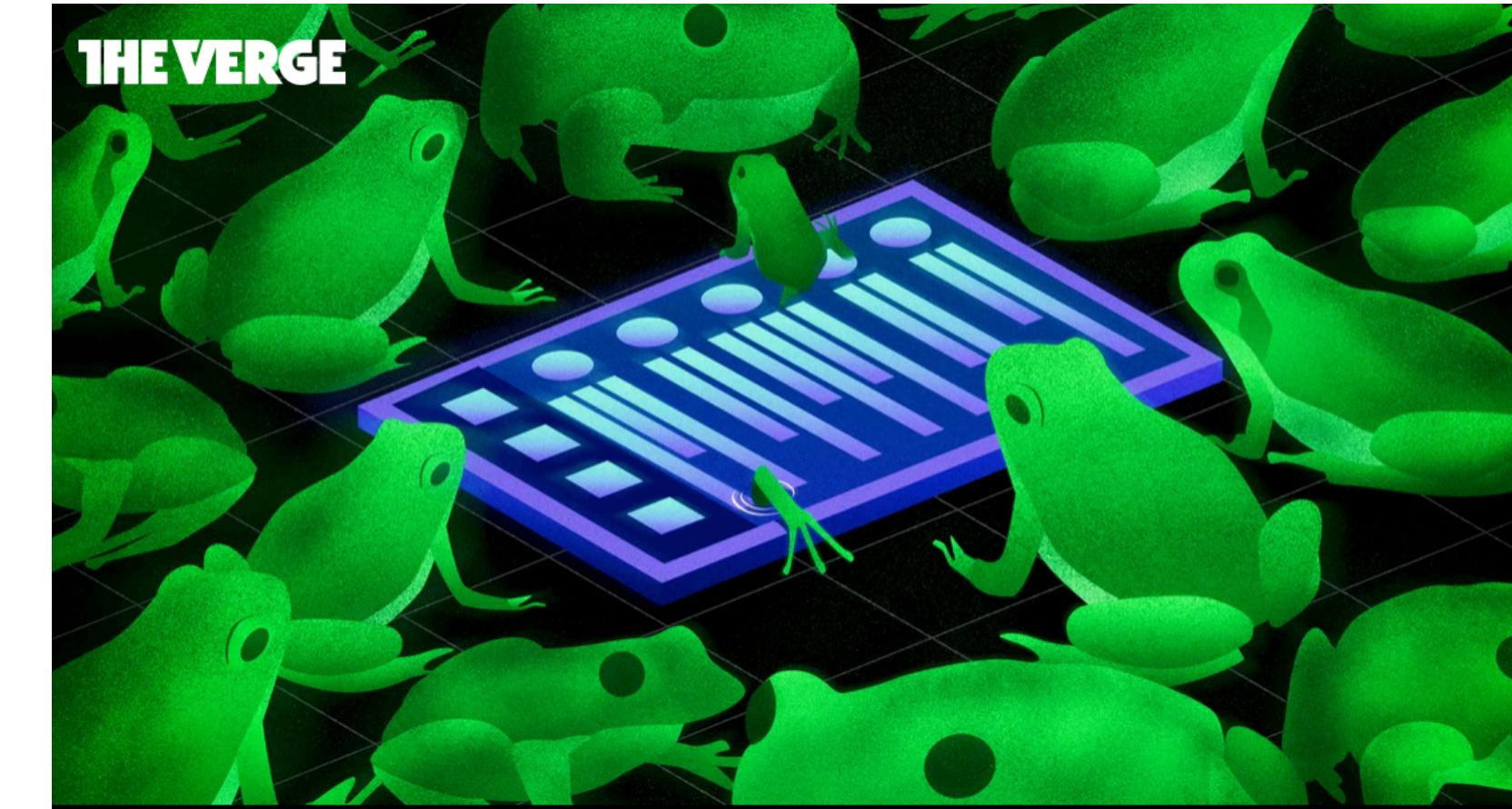
- Common question: what changes with the advent of the internet, compared to offline interaction?
- The effort required to connect groups together has lowered, making it possible for identity-based groups to connect that might otherwise have not
  - Otherwise, there might not have been enough of a critical mass of members
  - Or, it might have been too much effort to coordinate and communicate

# Why now?

- The fact that it's easier to organize is both positive and negative:
  - Positive: social movements that are forced underground, such as LGBTQ military service members, can connect with each other online
  - Negative: hate groups can also connect with each other online

# Why now?

- When groups can convene and push their own narrative, it enables “common knowledge attacks on democracy”
- In other words, it can destabilize democracy by flooding public debate and confusing our shared understandings and expectations, which are required for democracy to function



WEB

## HOW THE BIGGEST DECENTRALIZED SOCIAL NETWORK IS DEALING WITH ITS NAZI PROBLEM

*Mastodon was built to be a kinder, more decentralized version of Twitter — then Gab showed up*

By Adi Robertson | @thedextriarchy | Jul 12, 2019, 2:51pm EDT

Illustration by Alex Castro



SHARE



Over the past few years, Mastodon has become the model for [a friendlier kind of social network](#), promising to keep out the hateful or ugly content that proliferates on larger and more centralized networks.

Journalists hailed it as “[Twitter without Nazis](#)” and for years, it’s generally lived up to that promise. But last week, the social

# Potential theories

- Theory #1: It's trolls
- Disinformation factories such as the Russian Internet Research Agency generate disinformation to harm us
- Coordinated effort, with explicit intent
- **This isn't wrong, but it isn't the theory with the strongest evidence**

Bail, C. A., Guay, B., Maloney, E., Combs, A., Hillygus, D. S., Merhout, F., ... & Volkovsky, A. (2020). Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proceedings of the national academy of sciences*, 117(1), 243-250.

# Potential theories

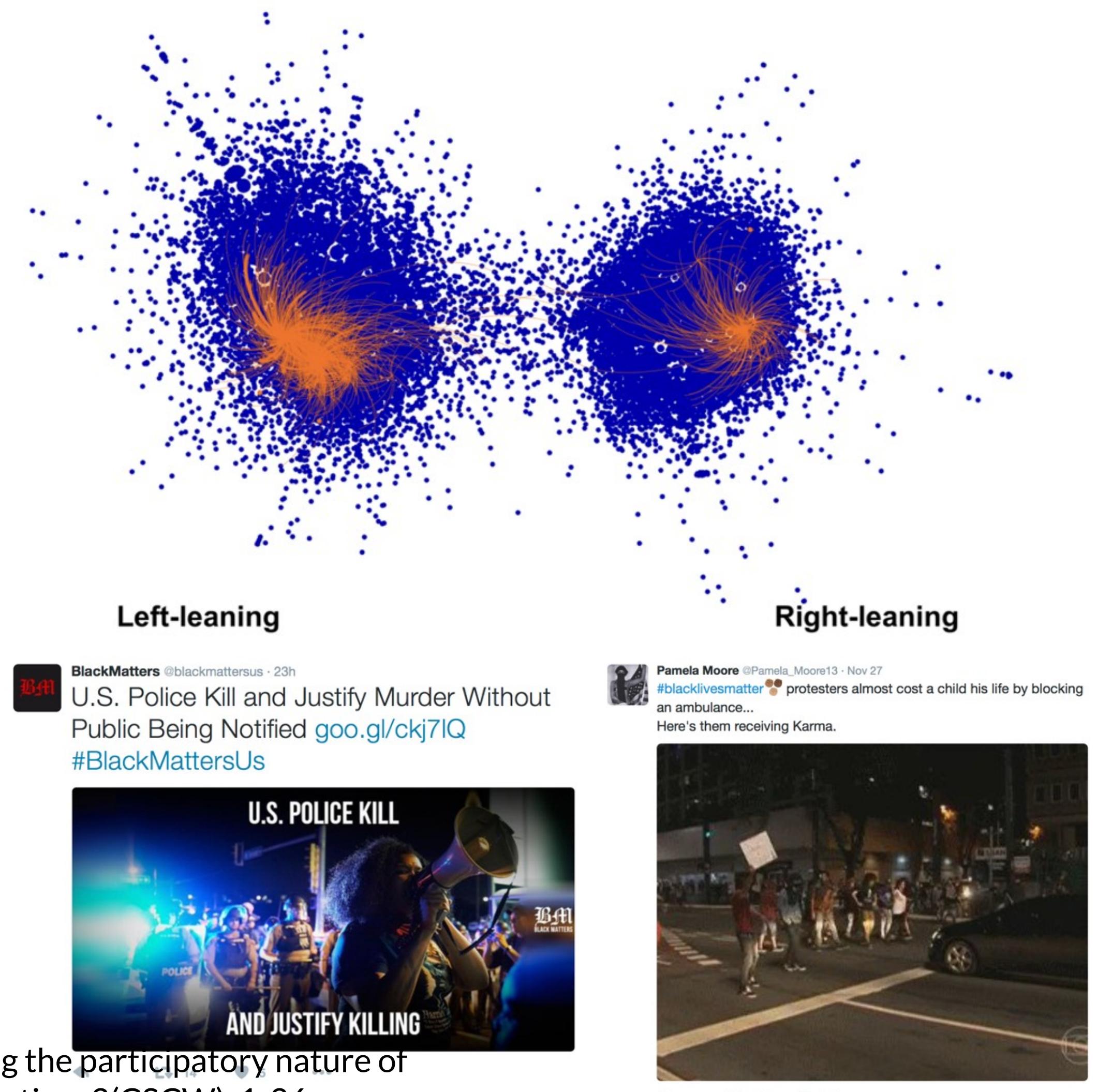
- Theory #2: “Post truth”
- People default to *motivated reasoning*, meaning that we are inclined to believe information that is consistent with our political views, and disinclined to believe information that contradicts our political views
- We are more loyal to political party than loyal to truth
- **This also isn't wrong, but there's still stronger evidence for other theories**

Kahan, D. M. (2017). Misconceptions, misinformation, and the logic of identity-protective cognition.

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in cognitive sciences*, 22(3), 213-224.

# Theory #1: it's trolls

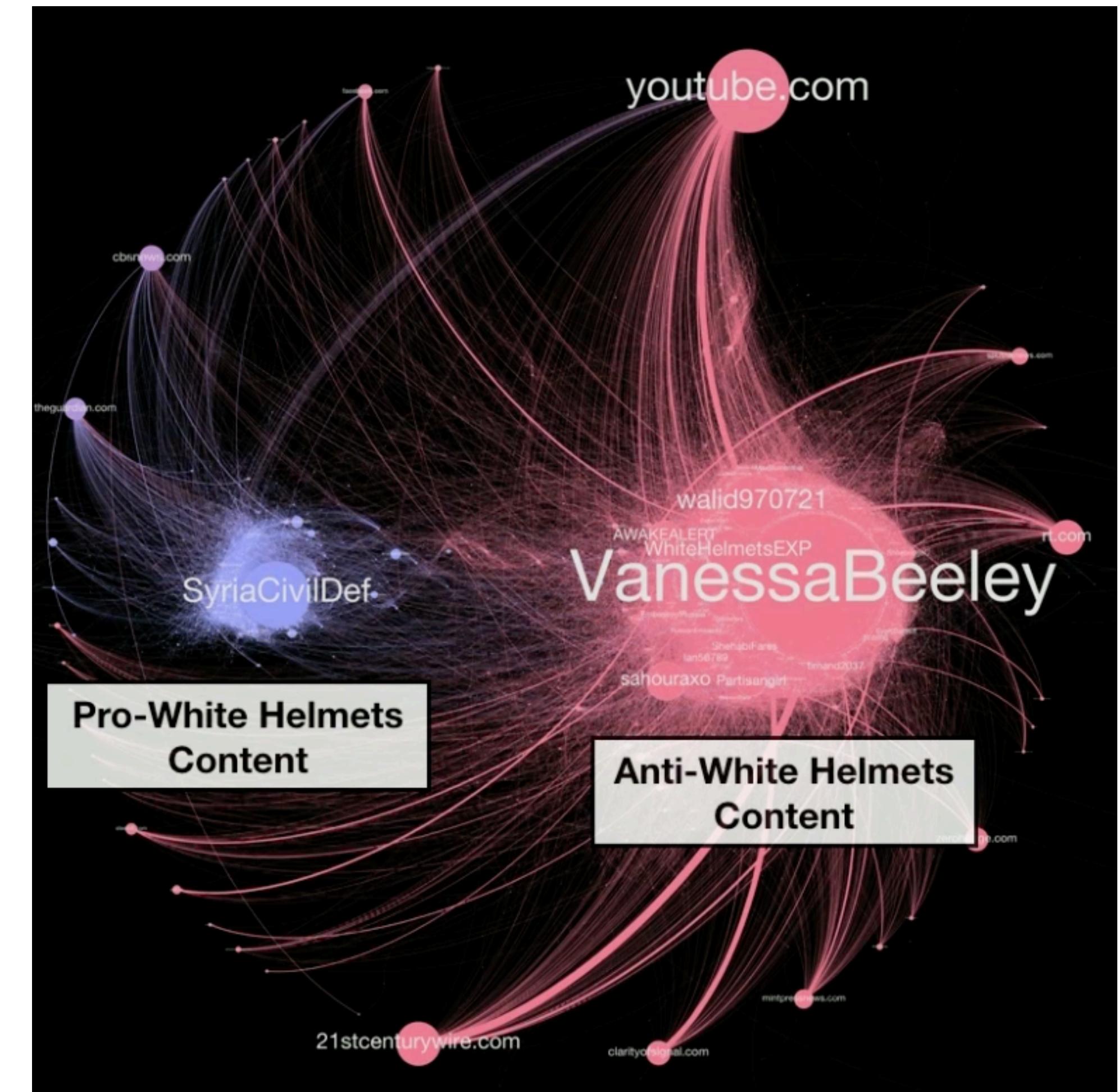
- There definitely are state actors
- Twitter retweet network for Black Lives Matter in 2016
  - Russian IRA (orange) both posed as BLM activists on the left, and infiltrated anti-BLM communities on the right



Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1-26.

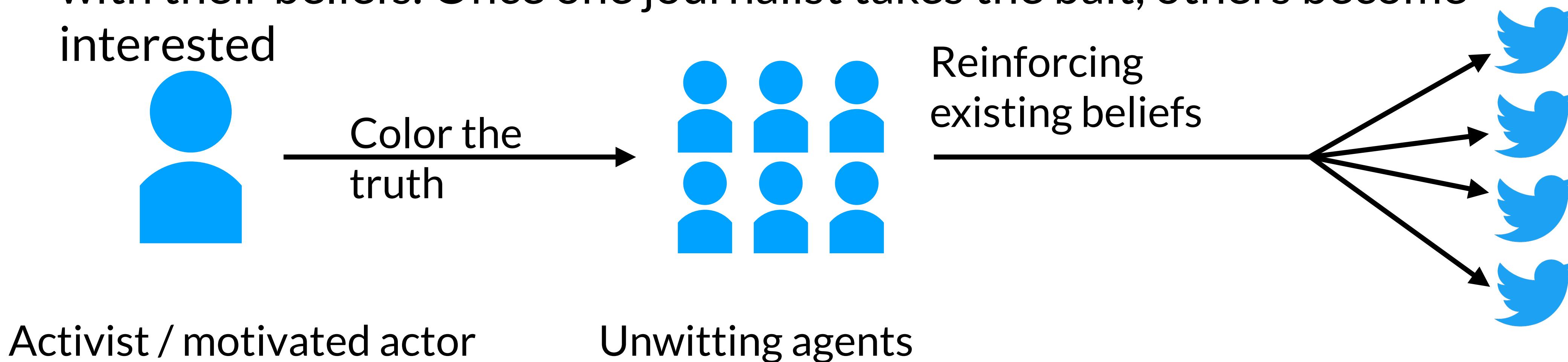
# Theory #1: it's trolls

- Context: Syrian Civil War, the White Helmets are a humanitarian response group
  - Anti-White Helmet accounts (pink) are dominant in volume, delegitimizing the White Helmets' claims
- Not just bots and trolls: lots are journalists aligned with Syrian and Russian government interests, Syrian and Russian government members, and alt. media
- It looks more like *activism* than “just” disinformation



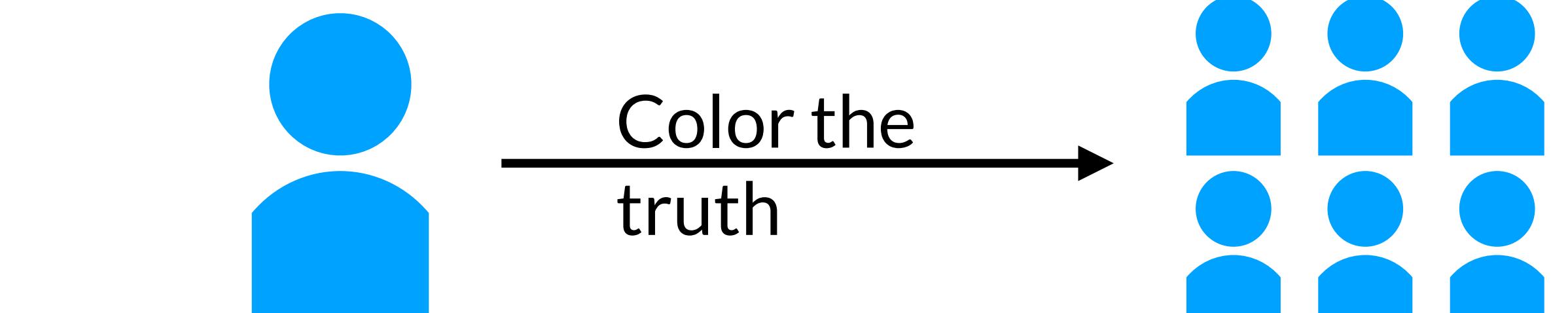
# Theory #1: it's trolls

- Disinformation campaigns often involve unwitting agents who are unaware of their role and whose views and behaviors have been shaped by motivated actors
- Cold War-era Soviet technique: sell journalists on anonymous tips aligned with their beliefs. Once one journalist takes the bait, others become interested



Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-26.  
Bittman, L. (1985). The KGB and Soviet disinformation: an insider's view. (No Title).

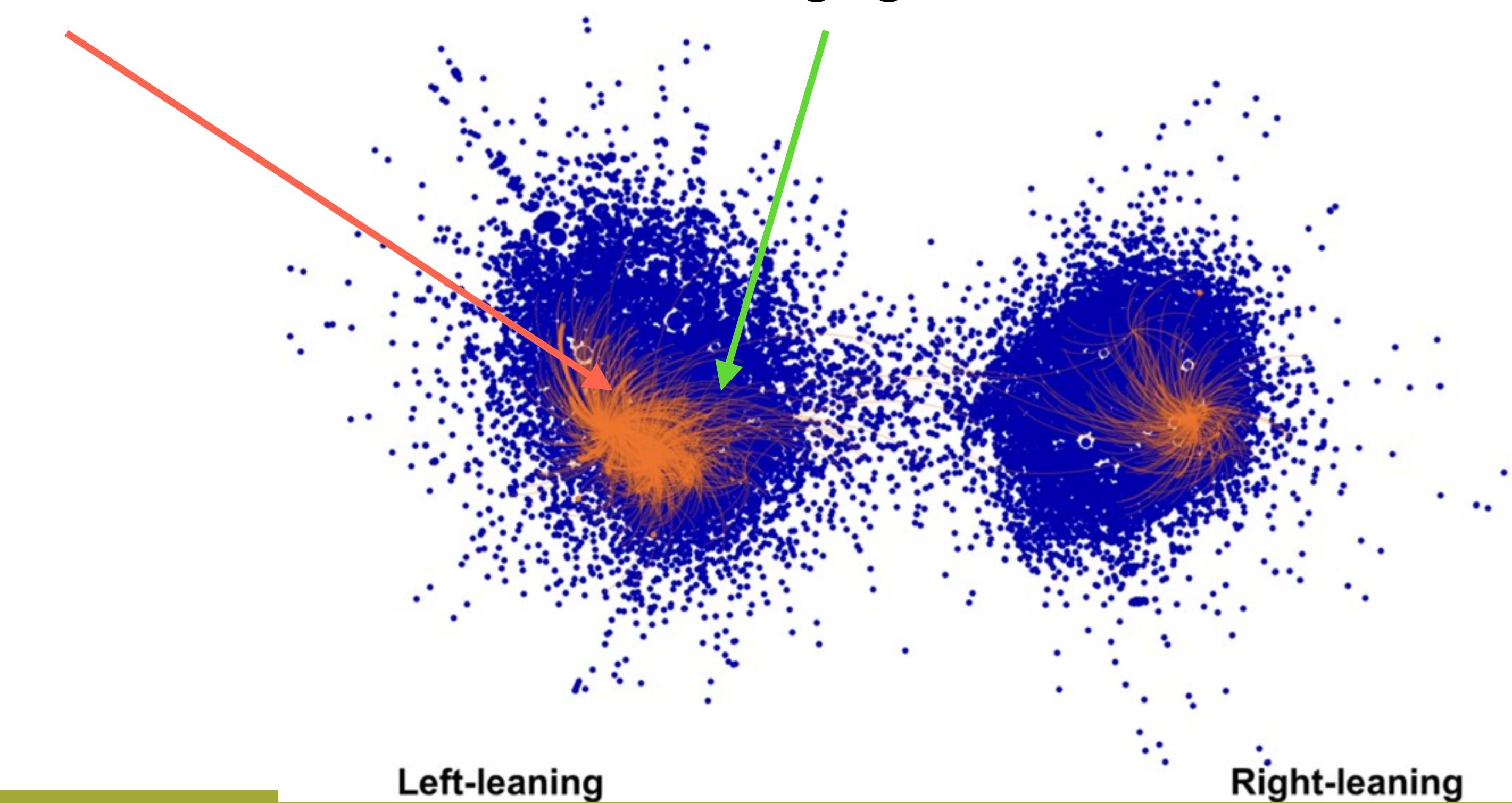
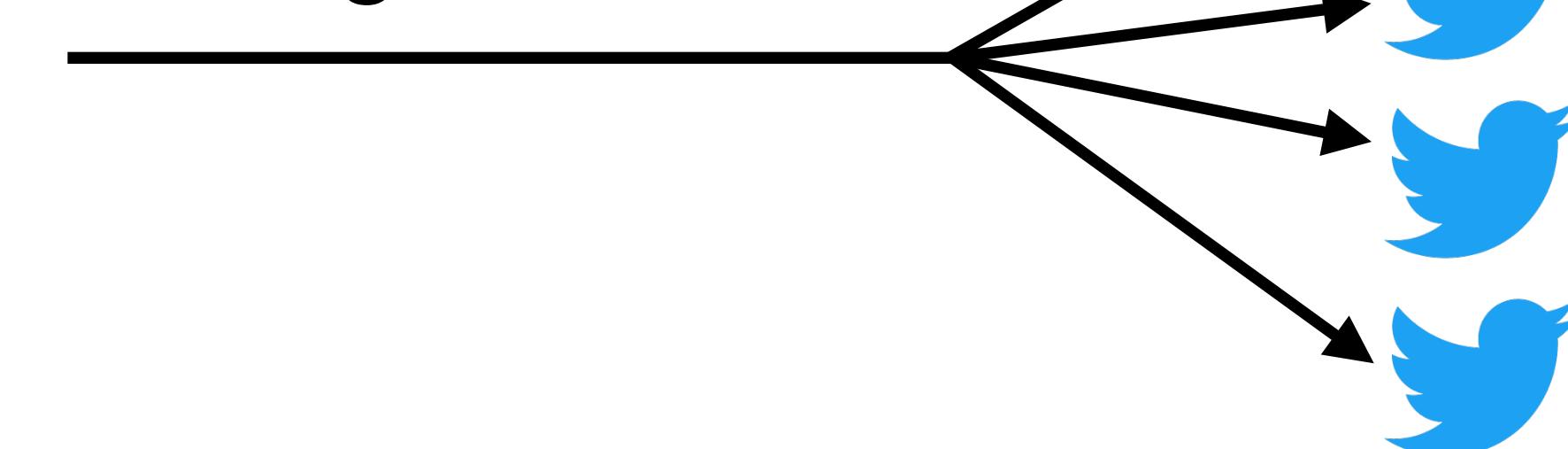
# Theory #1: it's trolls



Activist / motivated actor

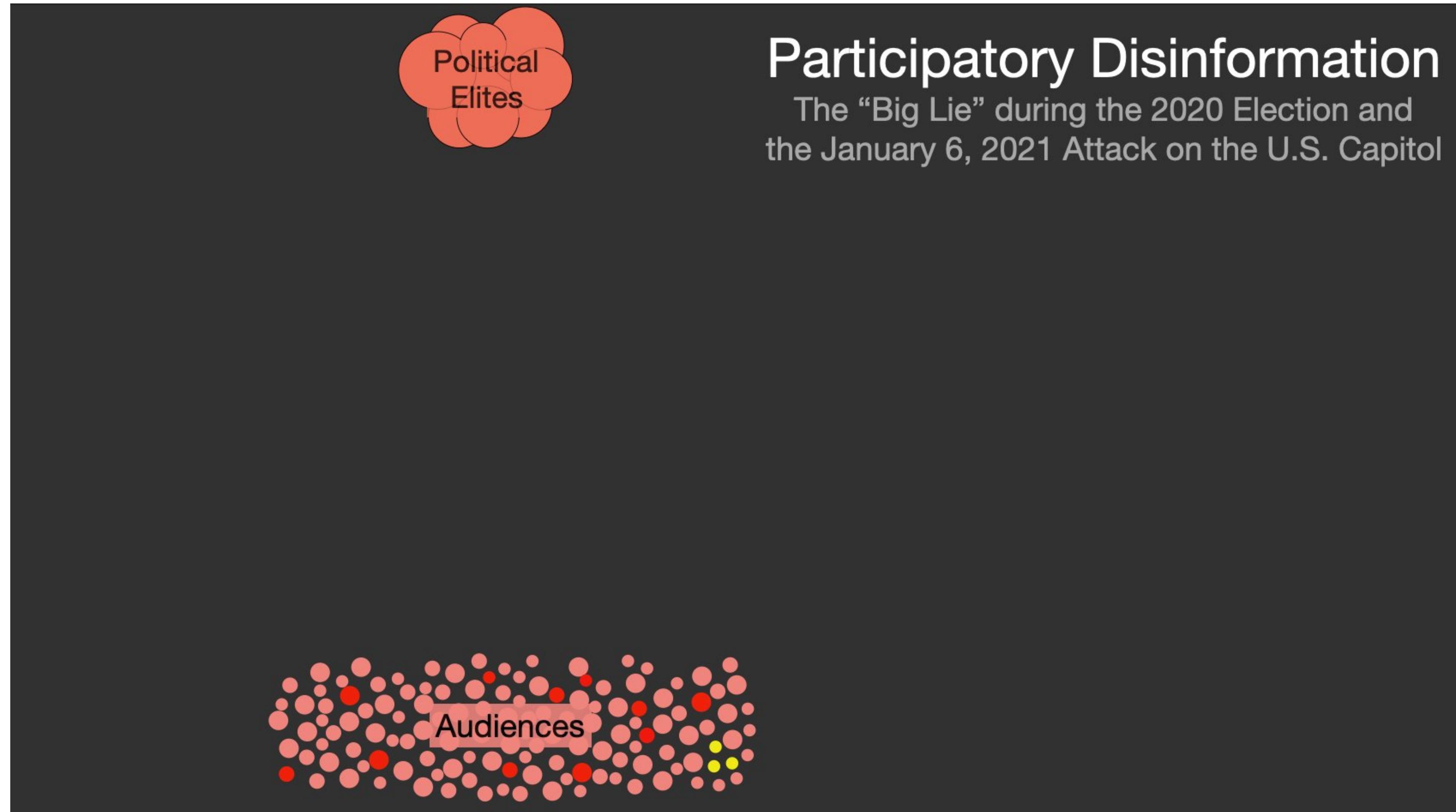
Unwitting agents

Reinforcing existing beliefs



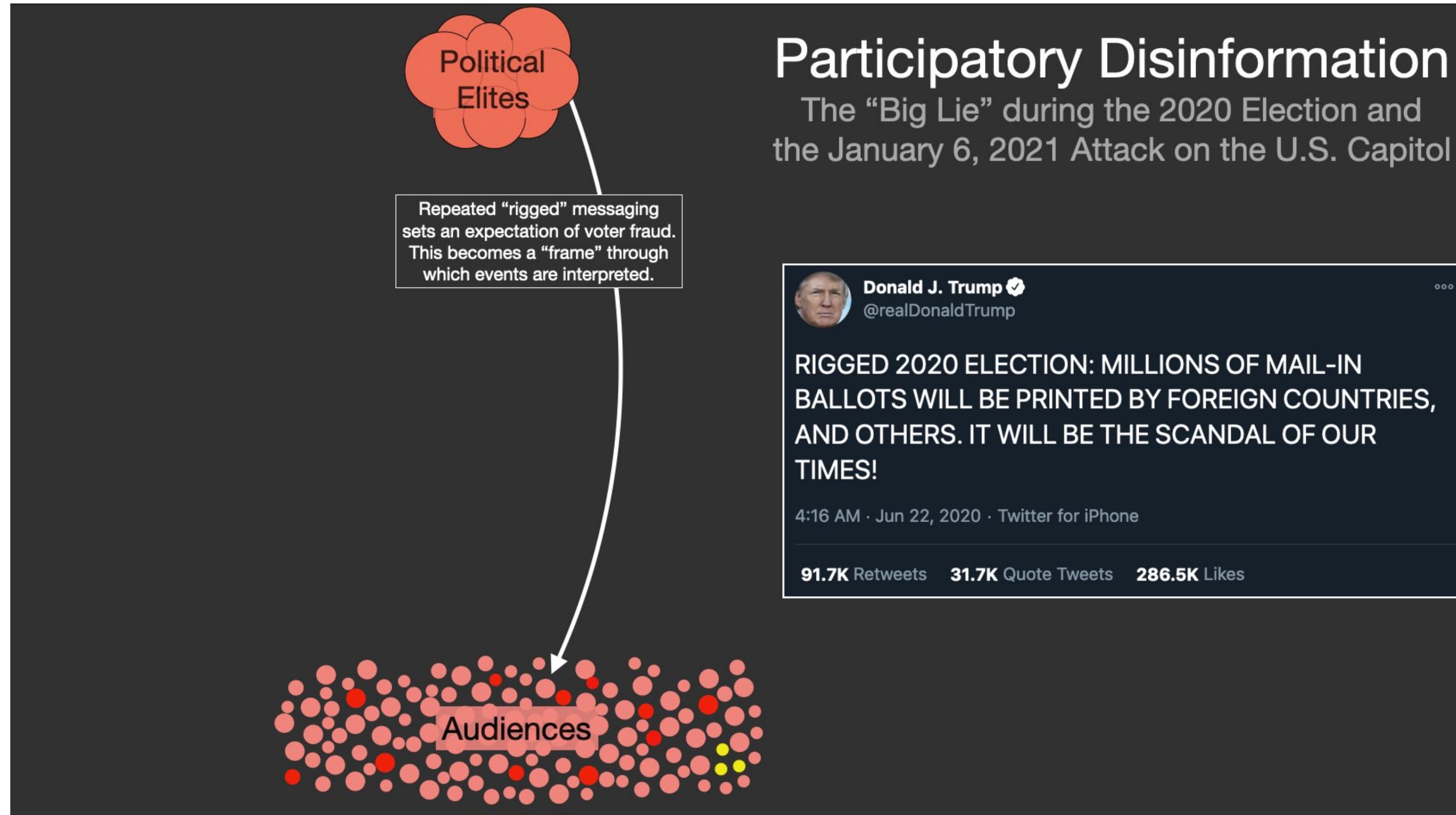
Orange arcs = blue accounts retweeting content from orange accounts = unwitting agents

# **Participatory disinformation and the 2020 election**



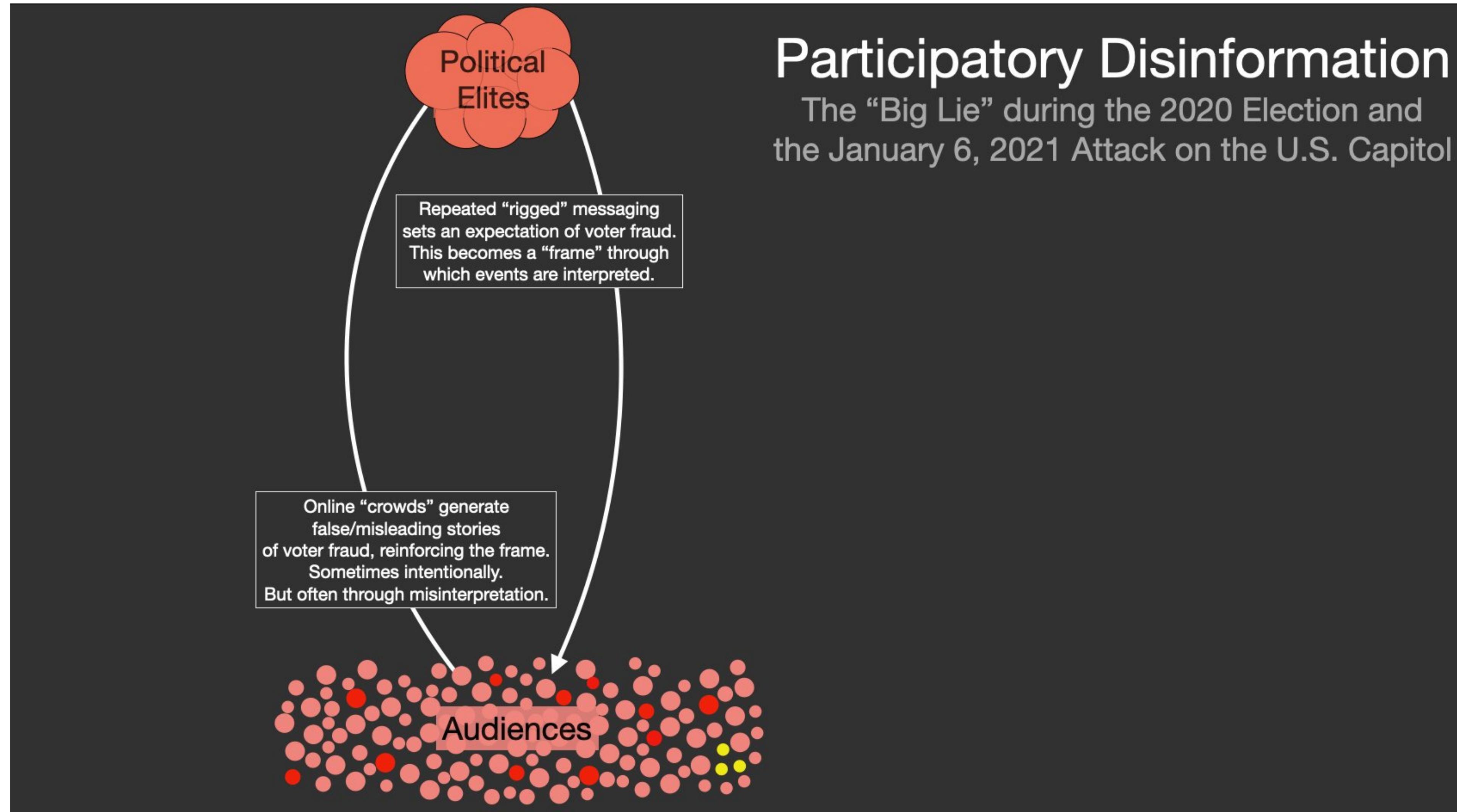
Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.



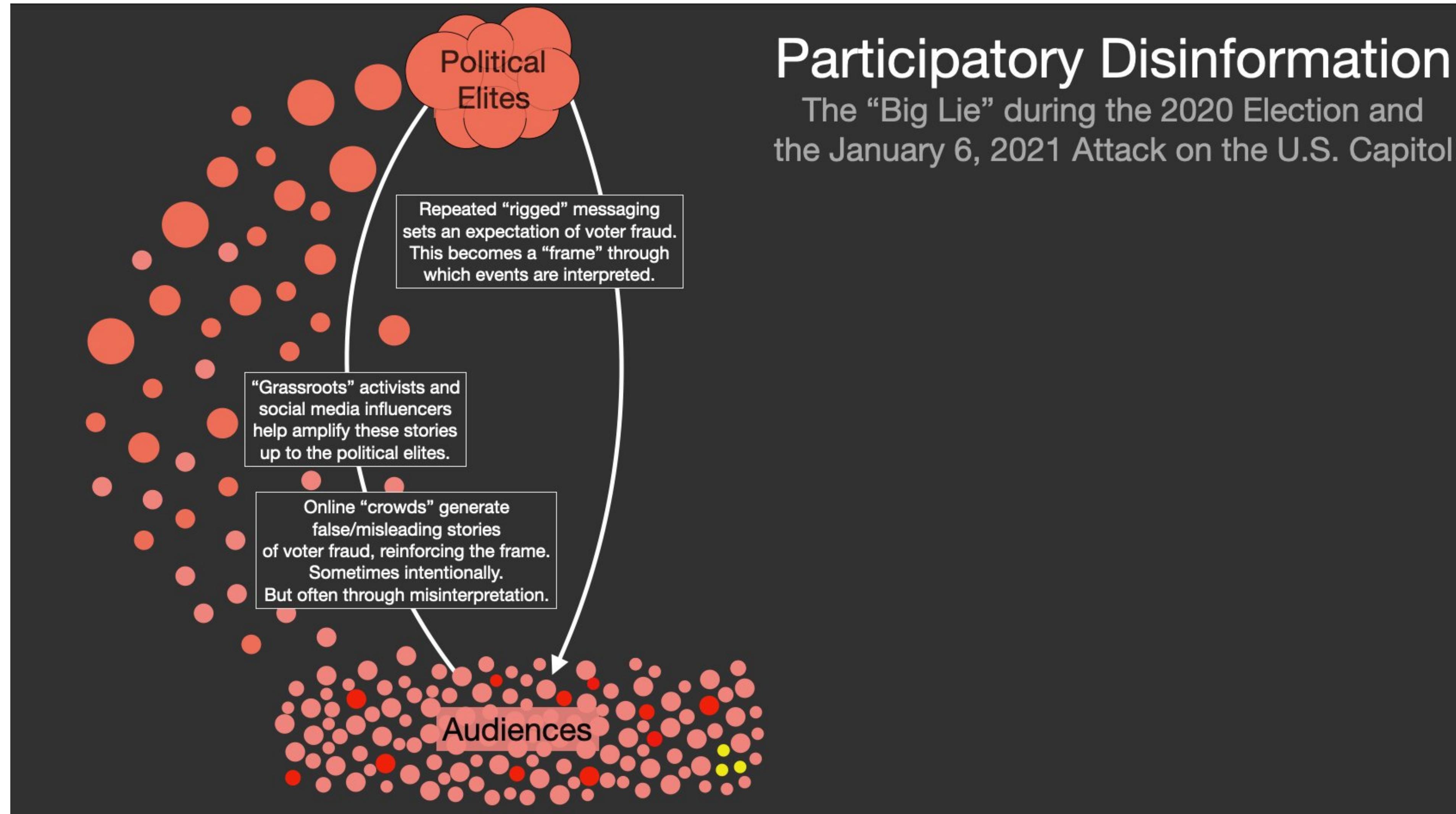
Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.



Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.

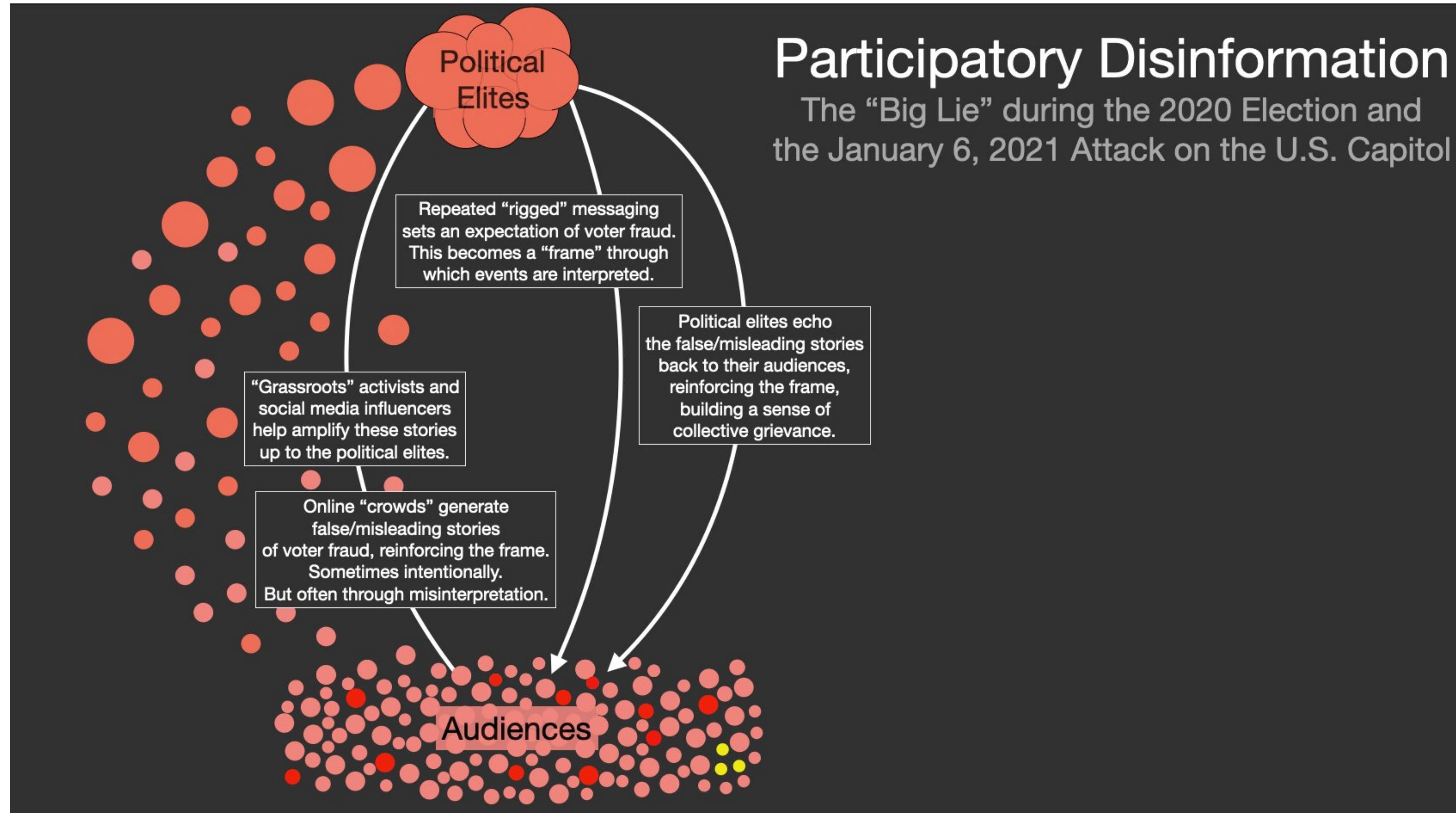


# Participatory Disinformation

The “Big Lie” during the 2020 Election and the January 6, 2021 Attack on the U.S. Capitol

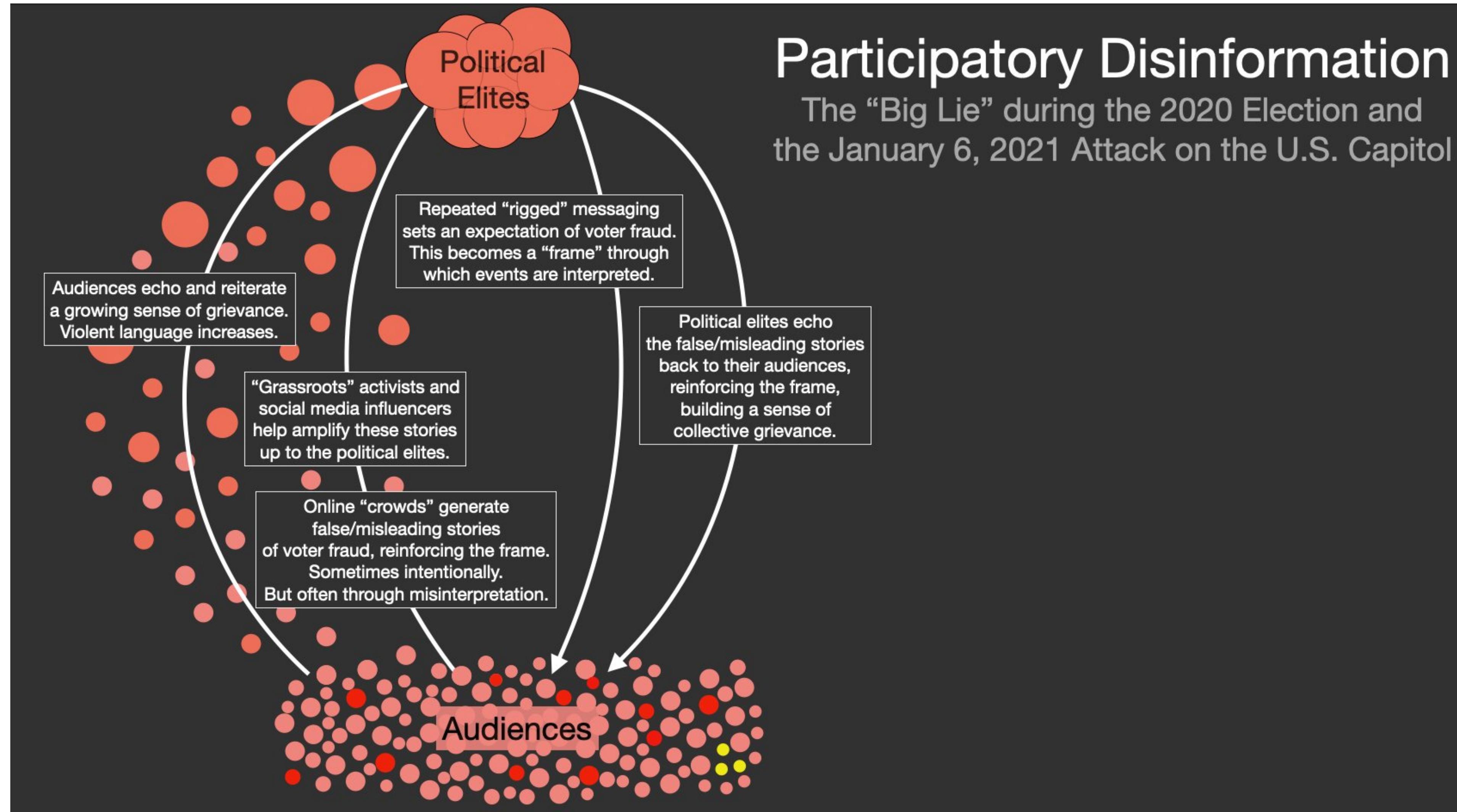
Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.



Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.

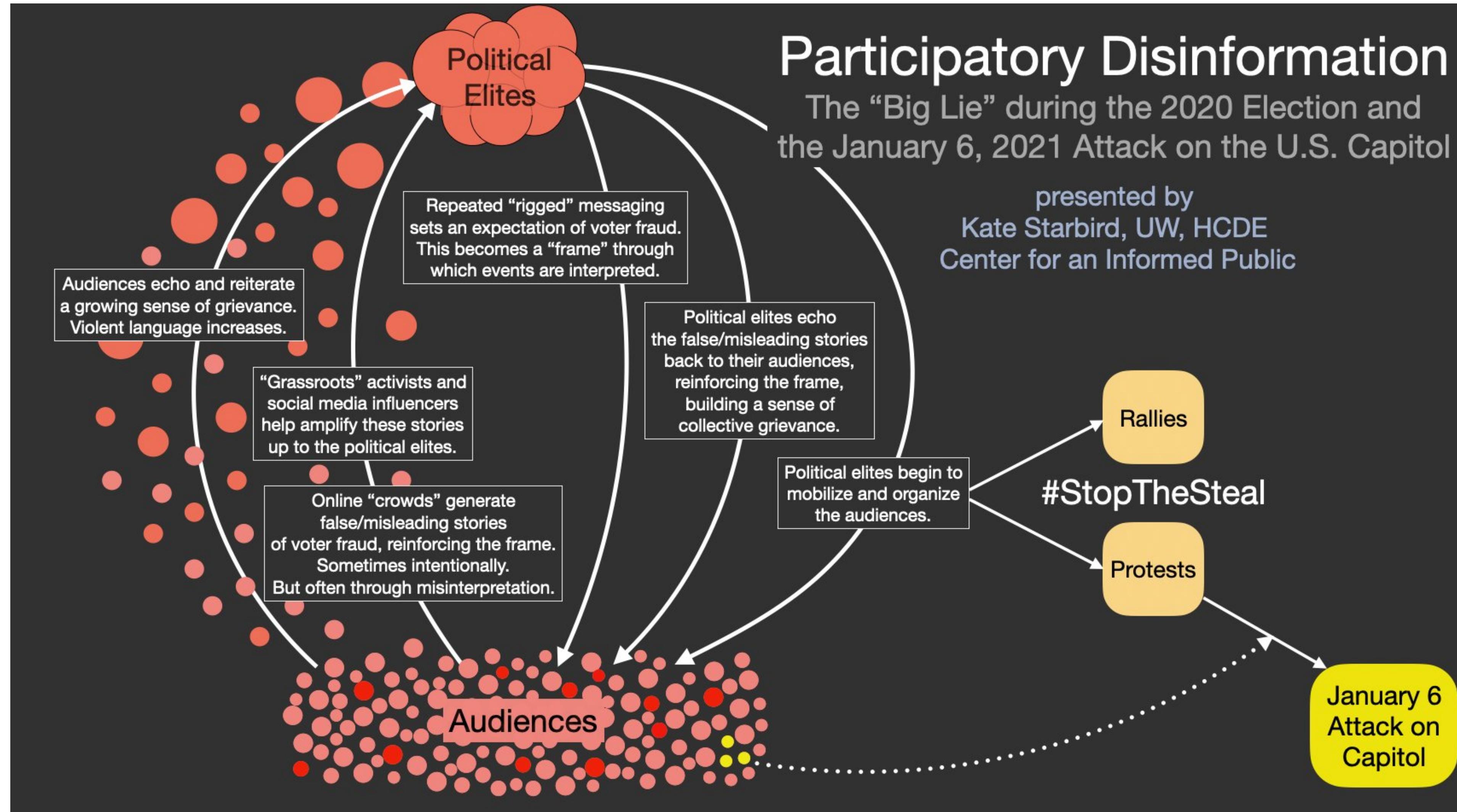


# Participatory Disinformation

The “Big Lie” during the 2020 Election and the January 6, 2021 Attack on the U.S. Capitol

Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.



Starbird, K., DiResta, R., & DeButts, M. (2023). Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society*, 9(2), 20563051231177943.

Prochaska, S., Duskin, K., Kharazian, Z., Minow, C., Blucker, S., Venuto, S., ... & Starbird, K. (2023). Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US Presidential election. *Proceedings of the ACM on human-computer interaction*, 7(CSCW1), 1-39.

**Main takeaway: Misinformation is a  
*collaborative effort* between producers  
and the audience**

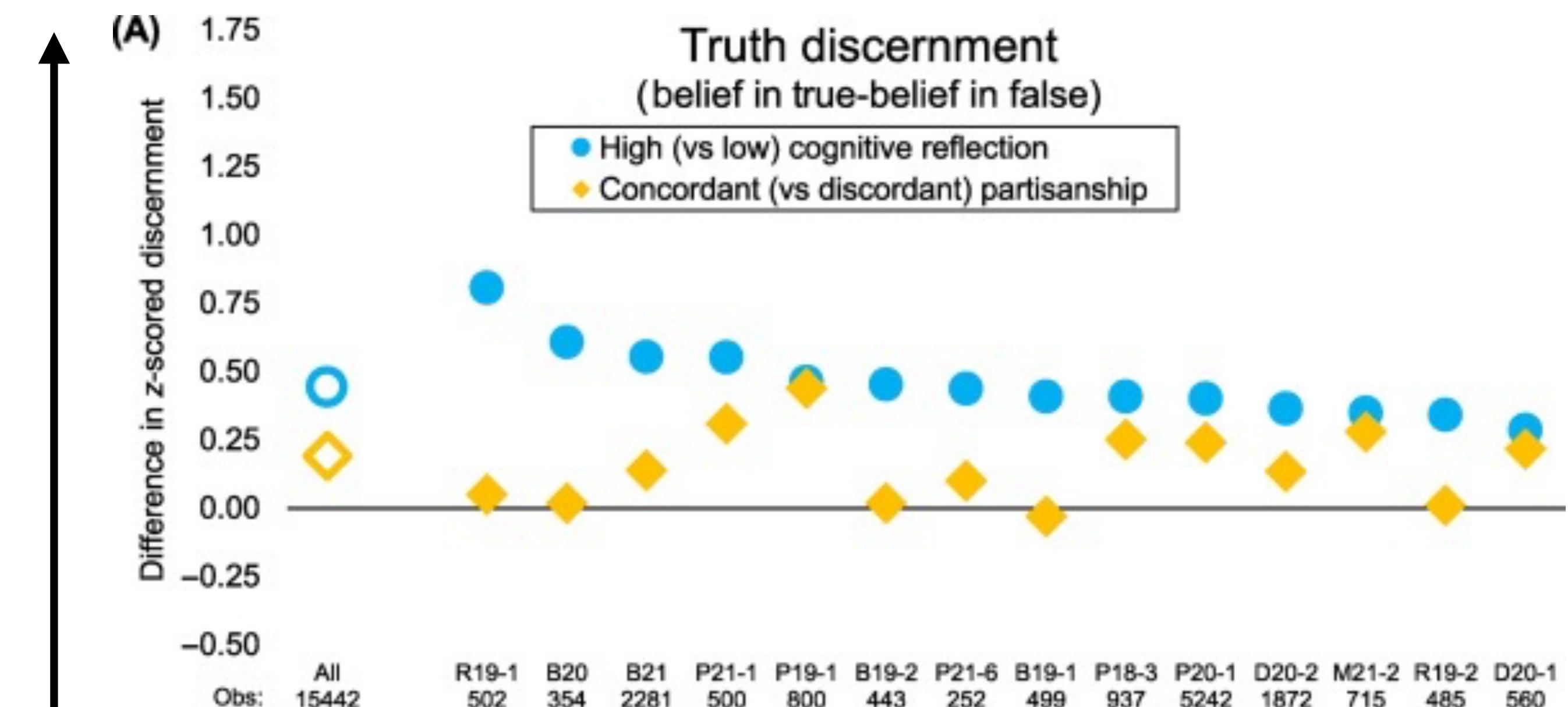
Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1-26.

# Theory #2: “Post truth”

Whether we share truthful content is more impacted by whether we are thinking deeply about the content we are sharing (using high-level reasoning) than whether content aligns with our beliefs

- While we are more likely to believe news that is concordant with our beliefs, the larger effect is whether we *engage with higher-level reasoning instead of automatic reasoning*

Better discernment  
of true vs. false  
news



Across a bunch of studies

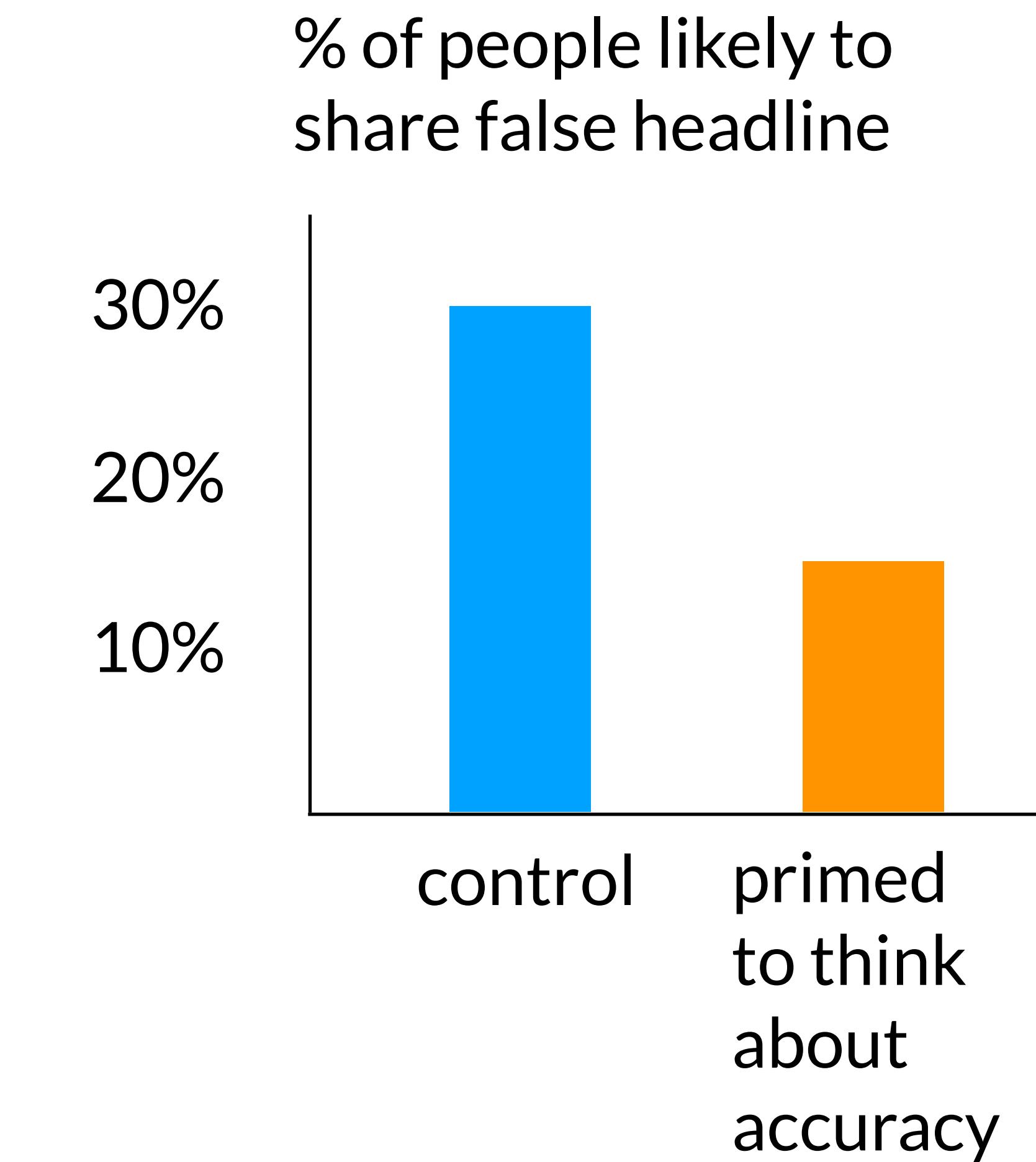
Engaged in higher-level reasoning (= bigger effect)  
vs.  
Saw belief-concordant news (= smaller effect)

# Theory #2: “Post truth”

- People rate accuracy as the single most important factor when deciding whether to share
- However, whether a headline is politically concordant has a much larger effect on sharing intention than the headline's accuracy
- Meaning, when deciding whether to share an article, people are more motivated by whether it aligns with their political beliefs than whether it's fully accurate
- So what gives?

# Theory #2: “Post truth”

- Theory: we don't pay attention to accuracy, and are more focused on pleasing followers or signaling group membership
- Evidence: focusing participants' attention on accuracy before seeing a headline reduces sharing of false headlines by half



Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature communications*, 13(1), 2333.

# How much do we consume?

- Most people rarely see misinformation.
- National sample of mobile, desktop, and TV consumption: misinformation is 0.15% of Americans' media diet
  - The average US adult saw ~1 misinformation story in the 2016 election
- Exposure to misinformation is highly concentrated: 1% of people account for 80% of exposures to misinformation
  - This exposure is typically pro-attitudinal

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science advances*, 6(14), eaay3539.

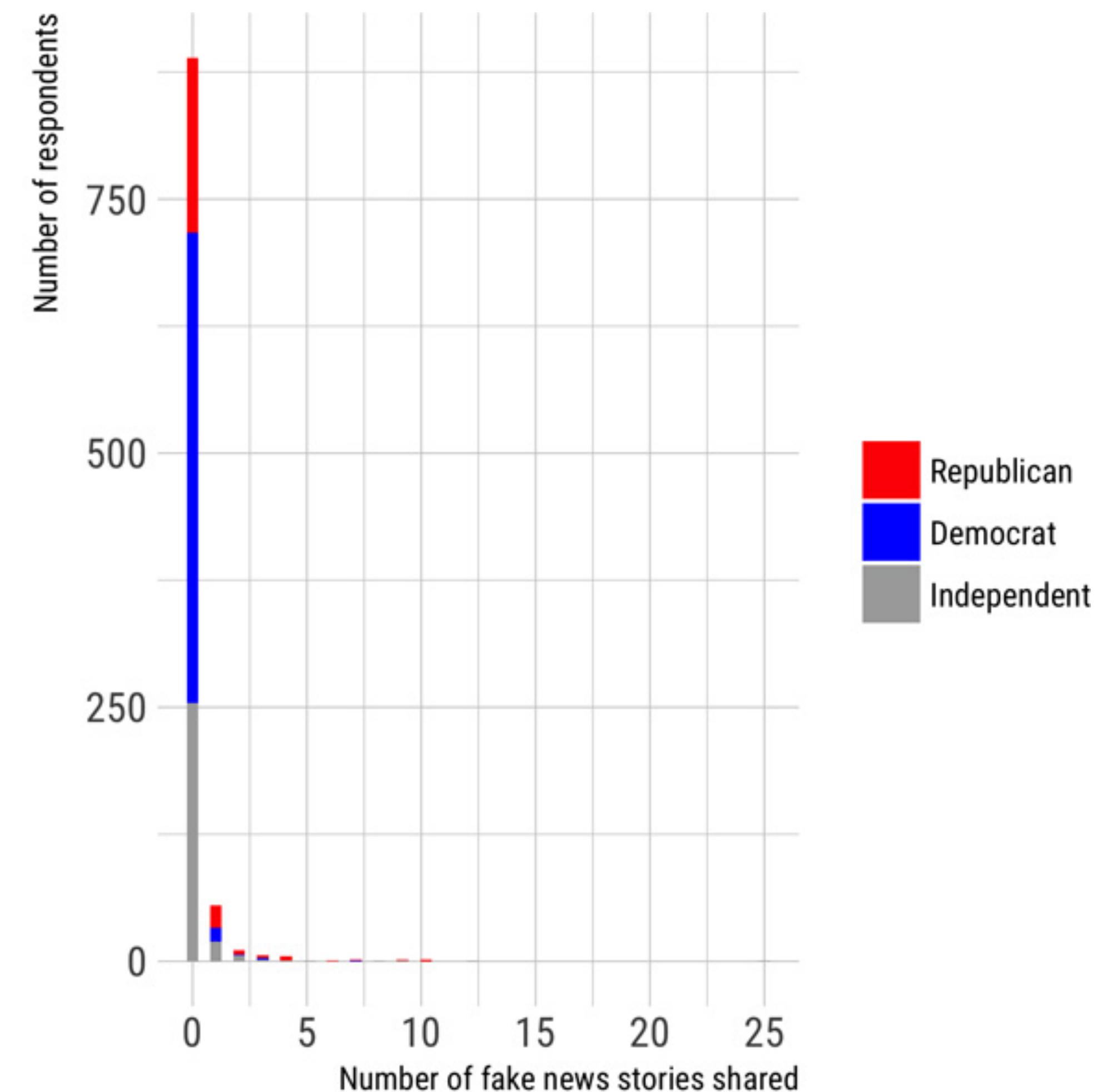
Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 211-236.

Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature human behaviour*, 4(5), 472-480.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, 363(6425), 374-378.

# How much do we share?

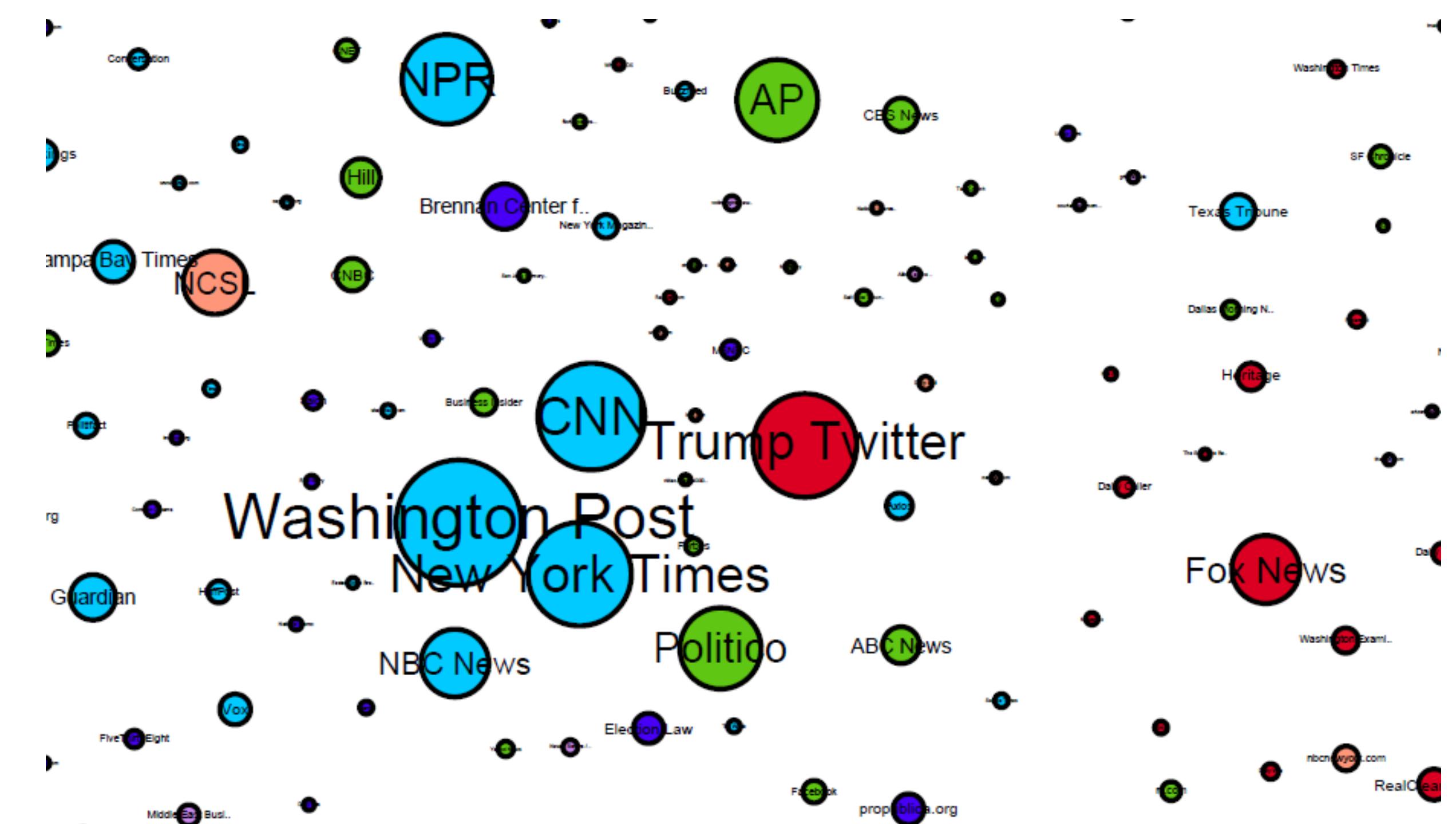
- It's rare: most never share disinformation
- In the 2016 election, >65 year olds were 7x more likely than teenagers to share disinformation
- And conservatives were more likely to share than liberals



Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances*, 5(1), eaau4586.

# Mainstream media amplifies the message

- Analysis of mail-in voter fraud disinformation suggested that social media played a smaller role than mainstream media in 2020:
  - 1) Trump tweets
  - 2) Mainstream media, trying to be neutral and avoid claims of anti-conservative bias, cover Trump's claims and thereby spread them



Benkler, Y., Tilton, C., Etling, B., Roberts, H., Clark, J., Faris, R., ... & Schmitt, C. (2020). Mail-in voter fraud: Anatomy of a disinformation campaign. Berkman Center Research Publication, (2020-6).

# So, where does misinformation come from?

- Theory #1: “It’s trolls”
  - In reality, it’s motivated actors, who activate unwitting agents
- Theory #2: “Post-truth”
  - In reality, people care about being accurate, but generally aren’t paying attention to accuracy when they share
- “It’s everywhere”
  - In reality, exposure and sharing is rare, but very concentrated

# **Classifying misinformation**

# Machine learning

- Some categories of misinformation can be flagged automatically
  - Near copies of flagged articles

ML APPLICATIONS

## Using AI to detect COVID-19 misinformation and exploitative content

May 12, 2020

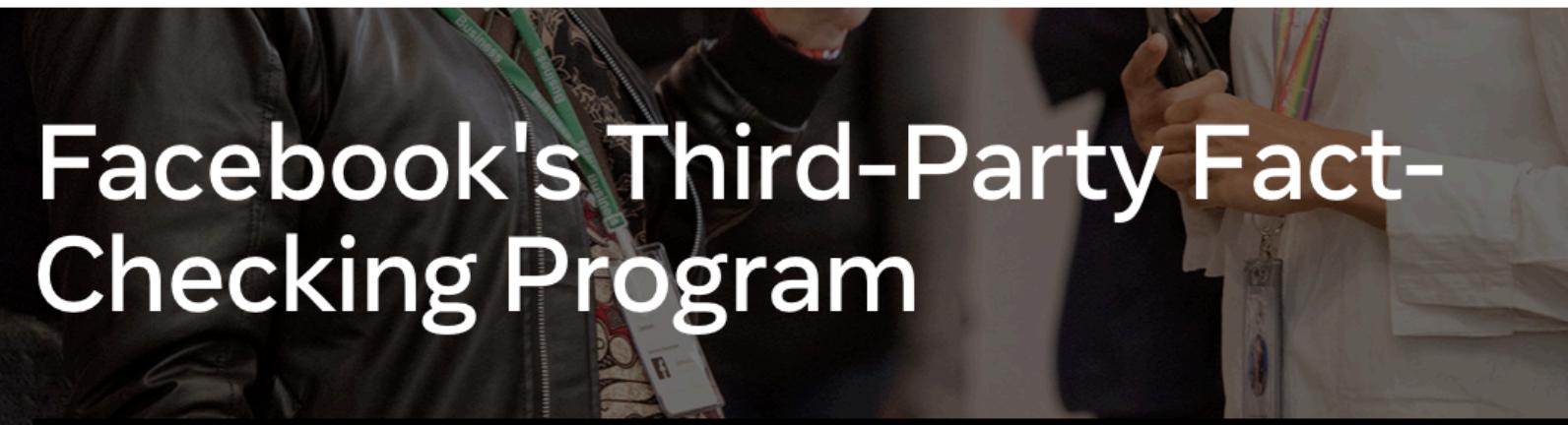
The COVID-19 pandemic is an incredibly complex and rapidly evolving global public health emergency. Facebook is committed to preventing the spread of false and misleading information on our platforms. Misinformation about the disease can evolve as rapidly as the headlines in the news and can be hard to distinguish from legitimate reporting. The same piece of misinformation can appear in slightly different forms, such as as an image modified with a few pixels cropped or augmented with a filter. And these variations can be unintentional or the result of someone's deliberate attempt to avoid detection. Furthermore, it is also important to avoid miscategorizing legitimate content as misinformation, because it could prevent people from being able to express themselves on our platforms.

AI is a crucial tool to address these challenges and prevent the spread of misinformation, because it allows us to leverage and scale the work of the independent fact-checkers who review content on our services. We work with over 60 fact-checking organizations around the world that review content in more than 50 languages. Since the pandemic began, we've used our current AI systems and deployed new ones to take COVID-19-related material our fact-checking partners have flagged as misinformation and then detect copies when someone tries to share them.

<https://ai.facebook.com/blog/using-ai-to-detect-covid-19-misinformation-and-exploitative-content/>

# Fact checkers

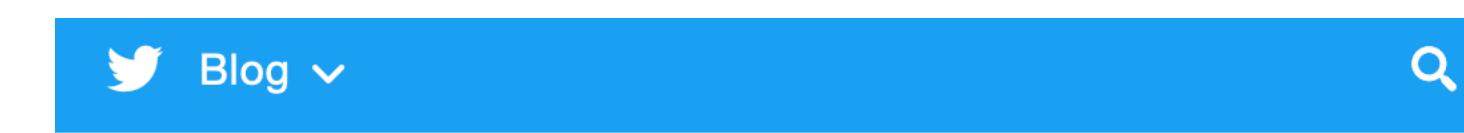
- Twitter and Facebook have started relying on third party fact checkers to decide whether an article is misinformation



**Facebook's Third-Party Fact-Checking Program**

Our Approach to Integrity on Facebook

Fighting misinformation is an ever-evolving problem and we can't do it alone. In 2016, we started our third-party fact-checking program, working with IFCN-certified fact-checkers around the world to rate and review the accuracy of content on our platform. The fact-checking program is one part of the three-part approach we take to addressing problematic content across the Facebook family of apps.



Product

## Updating our approach to misleading information

By [Yoel Roth](#) and [Nick Pickles](#)

Monday, 11 May 2020



In serving the public conversation, our goal is to make it easy to find credible information on Twitter and to limit the spread of potentially harmful and misleading content. Starting today, we're introducing new labels and warning messages that will provide additional context and information on some Tweets containing disputed or misleading information related to COVID-19.

# Fact checkers

- Twitter and Facebook have started relying on third party fact checkers to decide whether an article is misinformation
- However, this does not cover the long tail of articles shared
  - Facebook's partners are 26 fact checkers, who collectively review 200 articles per month
- Fact checkers can also take days to do the research
  - Given how quickly articles and videos spread, the damage could be already done

<https://thehill.com/policy/technology/478896-critics-fear-facebook-fact-checkers-losing-misinformation-fight/>

# Alternatives

- Representative samples of “crowd jurors” (users on the platform) can be as accurate as fact checkers, and much faster
- But, allowing anyone to participate can make the flagging another lever for politically motivated groups
  - Flagging your enemy’s content



Poynter.

Analysis Fact-Checking

**Analysis: Twitter’s crowdsourced fact-checking experiment reveals problems**

A Poynter analysis found that less than half of Birdwatch users include sources and many fact-checking notes contain partisan rhetoric.

Allen, J., Martel, C., & Rand, D. G. (2022, April). Birds of a feather don’t fact-check each other: Partisanship and the evaluation of news in Twitter’s Birdwatch crowdsourced fact-checking program. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-19).

# Why is classification of misinformation hard?

- Thoughtful fact checking requires thorough research, which takes time
- But the bigger picture: much of the disinformation out there isn't entirely made up
  - It's often a slant on a true story, and ambiguously on the line between "true" and "false"
- So what do we do, if most stories don't have a clear "true"/"false" line? What do we remove?



# Disinformation campaigns

- One strategy could be to classify disinformation campaigns, instead of individual pieces of content
  - 1) Is this campaign pushing a false narrative?
- Then, classify:
  - 2) Is this article a part of this disinformation campaign?
- This way, the decision can be made on the aggregation of actions, and a classifier only needs to assign the story to the campaign

Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1-26.

# **Technological interventions**

# Reduce the feed ranking

- Platforms can (temporarily) reduce the feed ranking of links that might be disinformation, slowing their spread while fact checkers review it
  - Ex: Article is lower in your Facebook feed, video is recommended less often on YouTube
  - Pros: walks a line between removal and unconstrained spread
  - Cons: opaque, unclear when it's happening, likely too late once other media start reporting on it

# Issue a correction

- Platforms could issue a correction, such as a pop-up adding more information if you watched a disinformation video/read an article
- Early research suggested that fact corrections could backfire and reinforce incorrect beliefs
- But, recent work fails to replicate the backfire effect and suggests that people do revise beliefs after a correction
- However, getting corrected leads people to decrease the quality and increase the political slant and toxicity of future content

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303-330.

Wood, T., & Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, 41, 135-163.

Guess, A., & Coppock, A. (2020). Does counter-attitudinal information cause backlash? Results from three large survey experiments. *British Journal of Political Science*, 50(4), 1497-1515.

Mosleh, M., Martel, C., Eckles, D., & Rand, D. G. (2021). Shared partisanship dramatically increases social tie formation in a Twitter field experiment. *Proceedings of the National Academy of Sciences*, 118(7), e2022761118.

# A better approach

- Intervene before the narratives begin



Laura Edelson @whiskeyocelot@mastodon.social  
@LauraEdelson2

...

ONE NEAT TRICK for making an information space hostile to misinformation: Flood the zone! The US government deserves credit for doing this early. Not leaving an information vacuum for your opponent to fill makes their job much, much, harder. 2/8



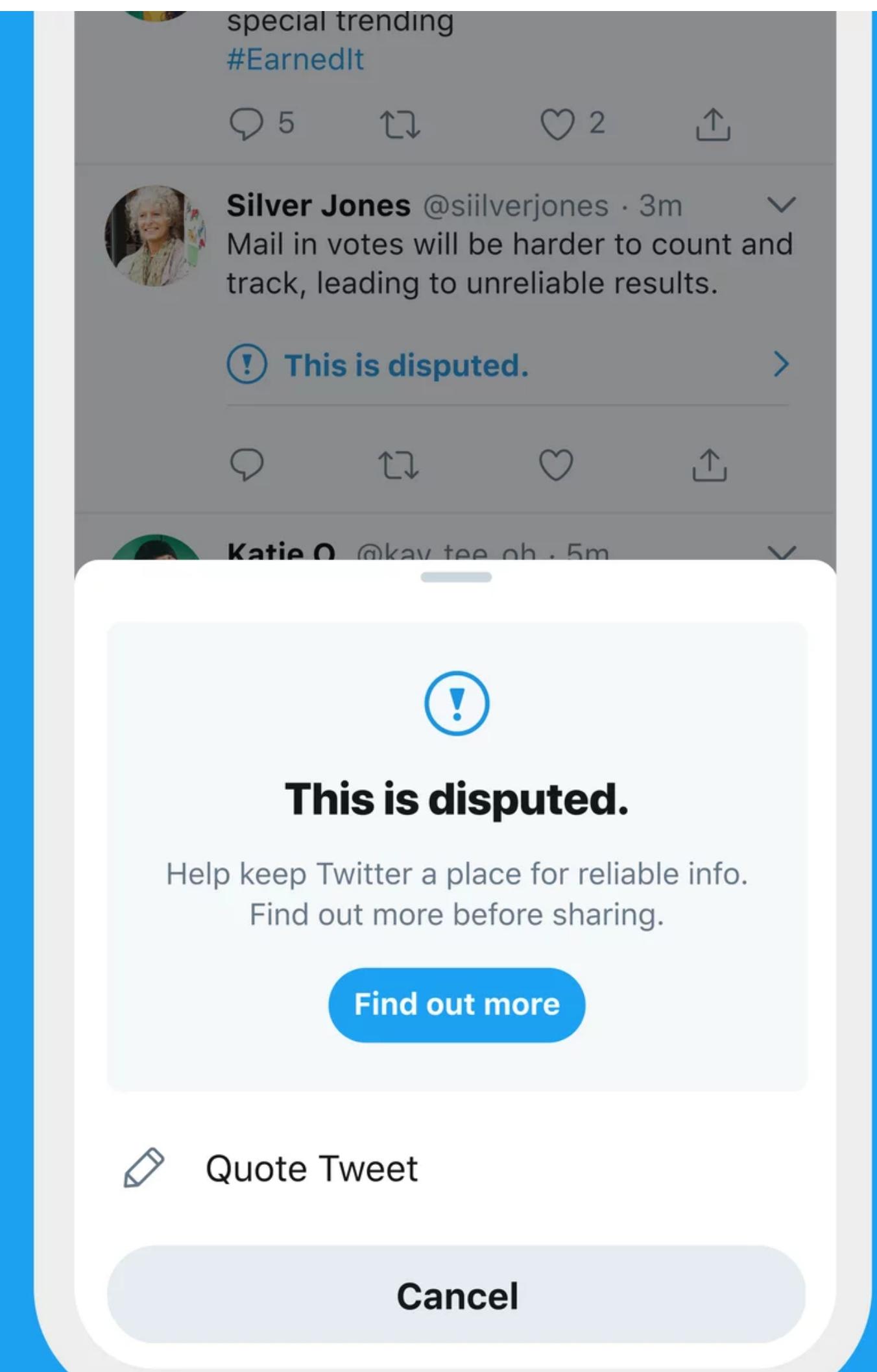
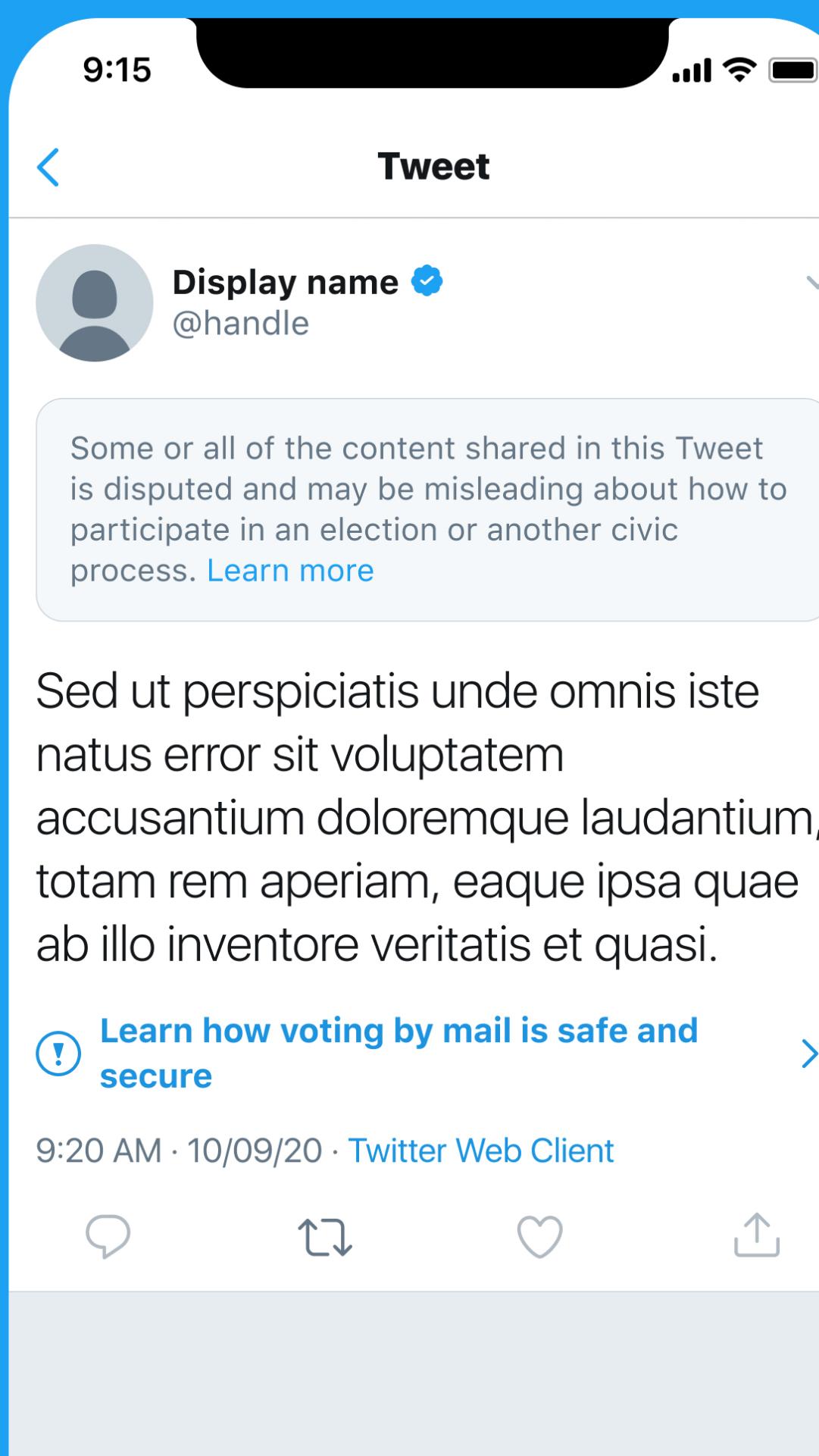
thehill.com

US employs unusual intel strategy to counter Putin

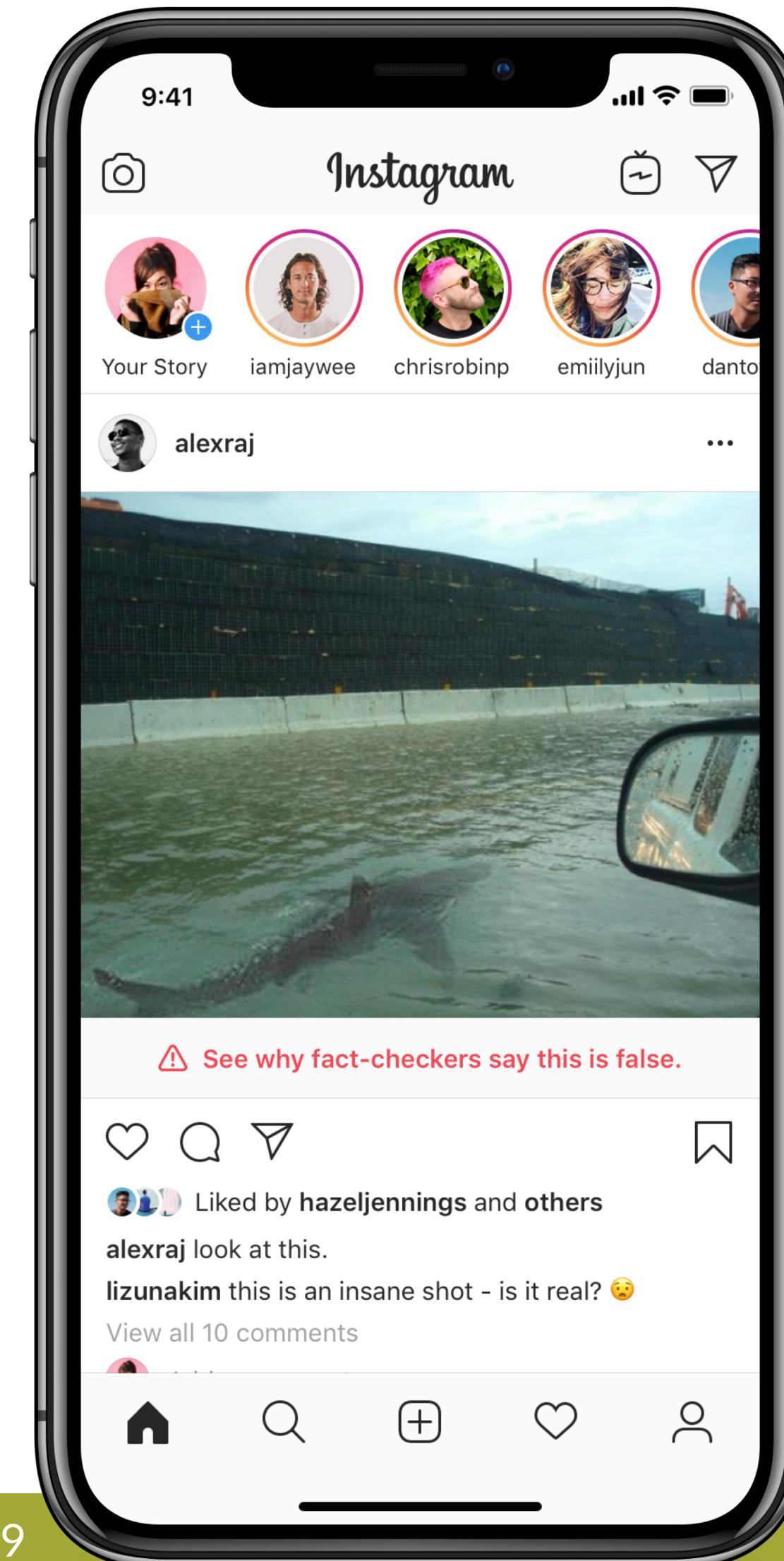
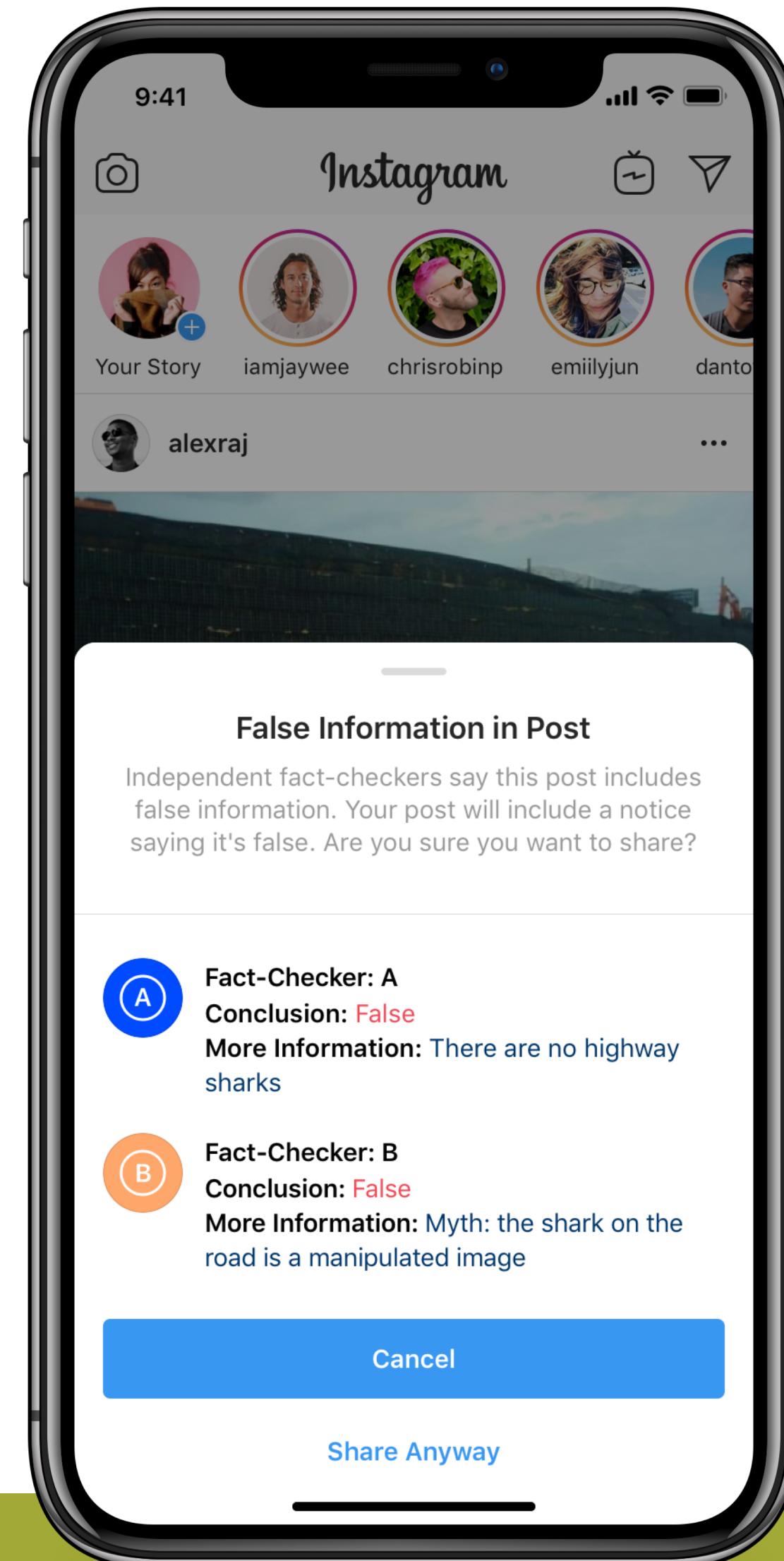
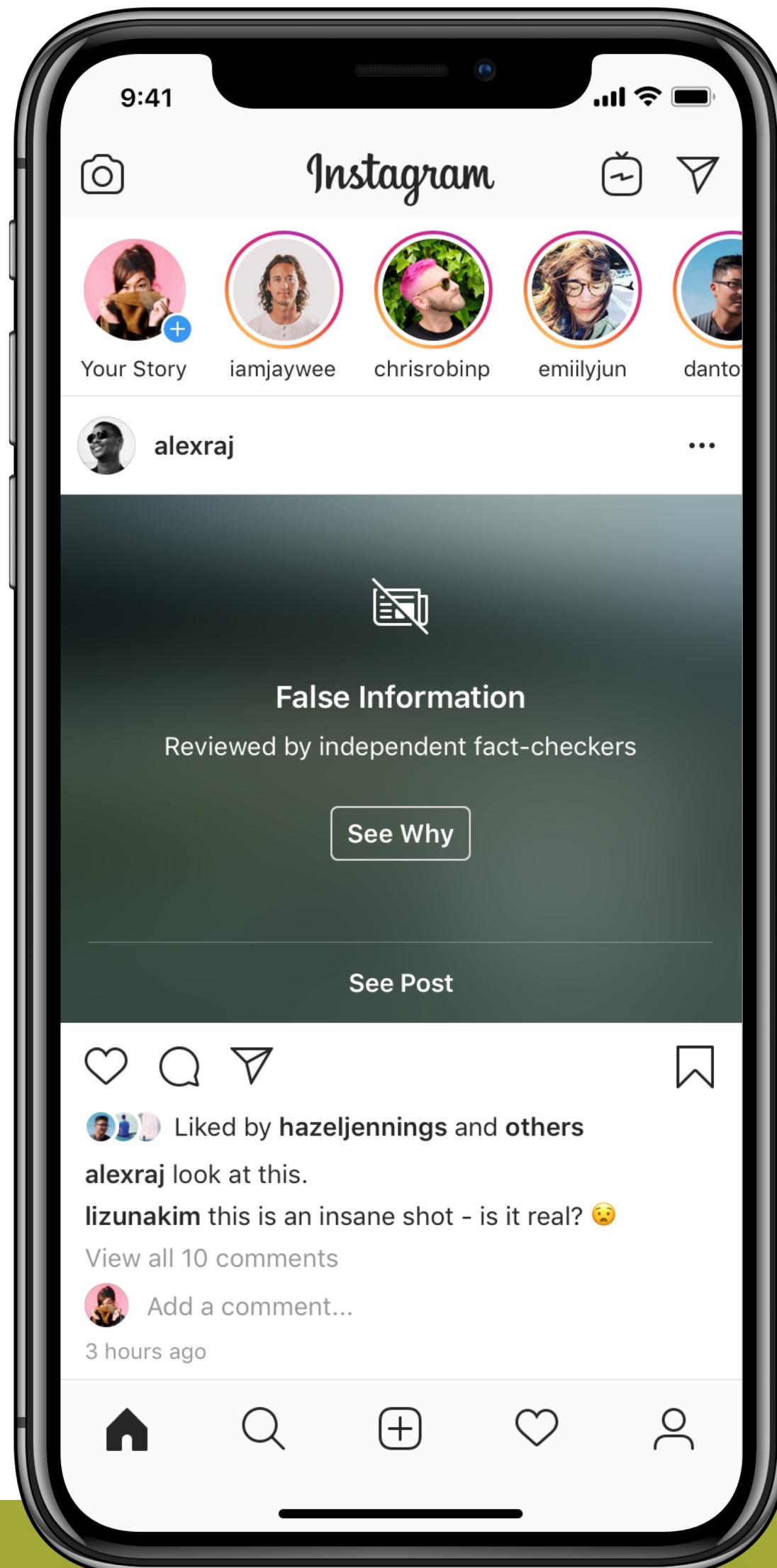
The White House has aggressively worked to rebut false narratives emanating from Russia about the crisis in Ukraine by proactively releasing intelligence ...

11:29 AM · Mar 1, 2022

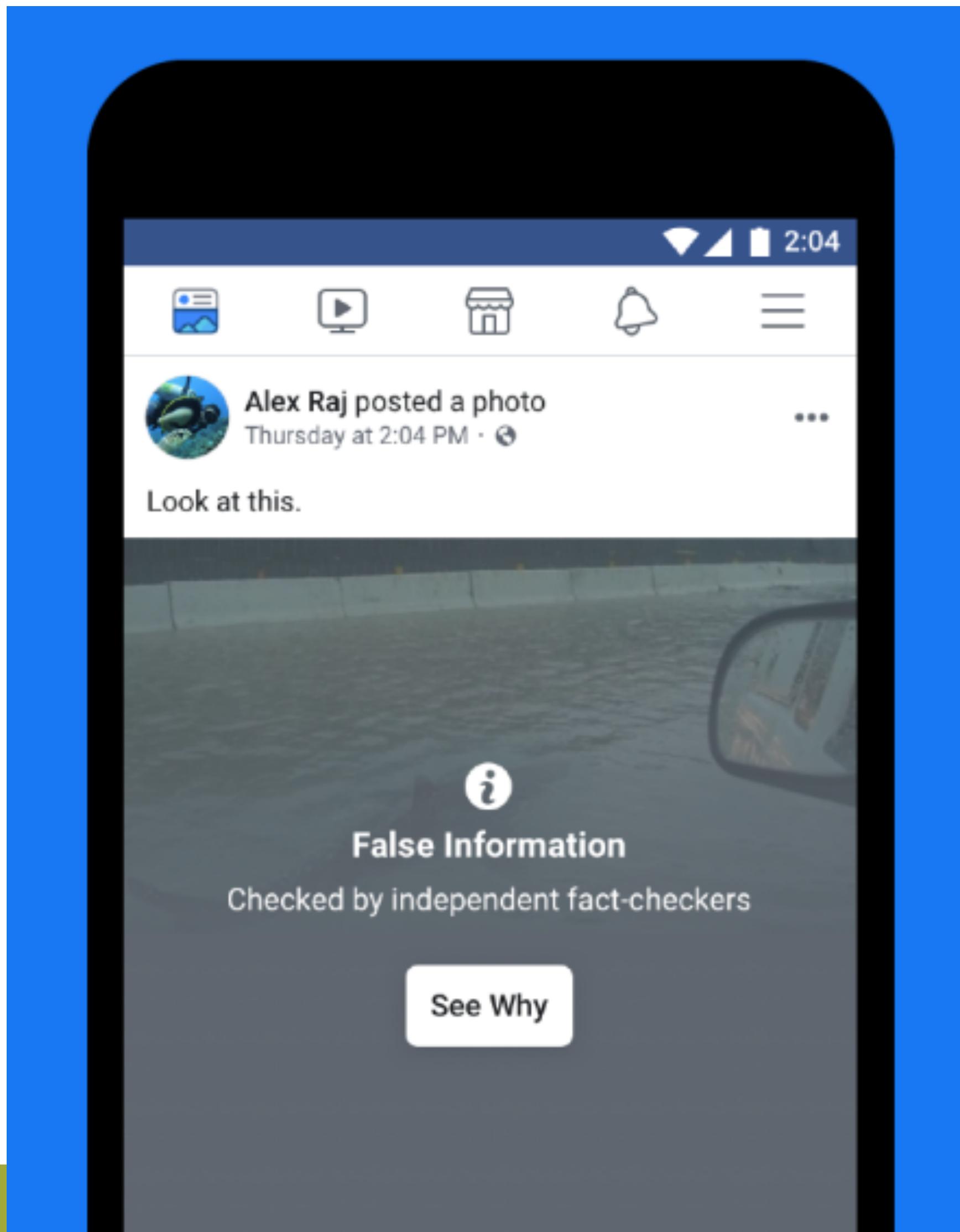
# Some interventions



# Some interventions



# Some interventions



# Some interventions



## U.S. elections

Robust safeguards help ensure the integrity of election results.

[LEARN MORE](#)

CHRISTINA BOBB: Trump won. MSM hopes you don't believe your eyes.

# Implied truth effect

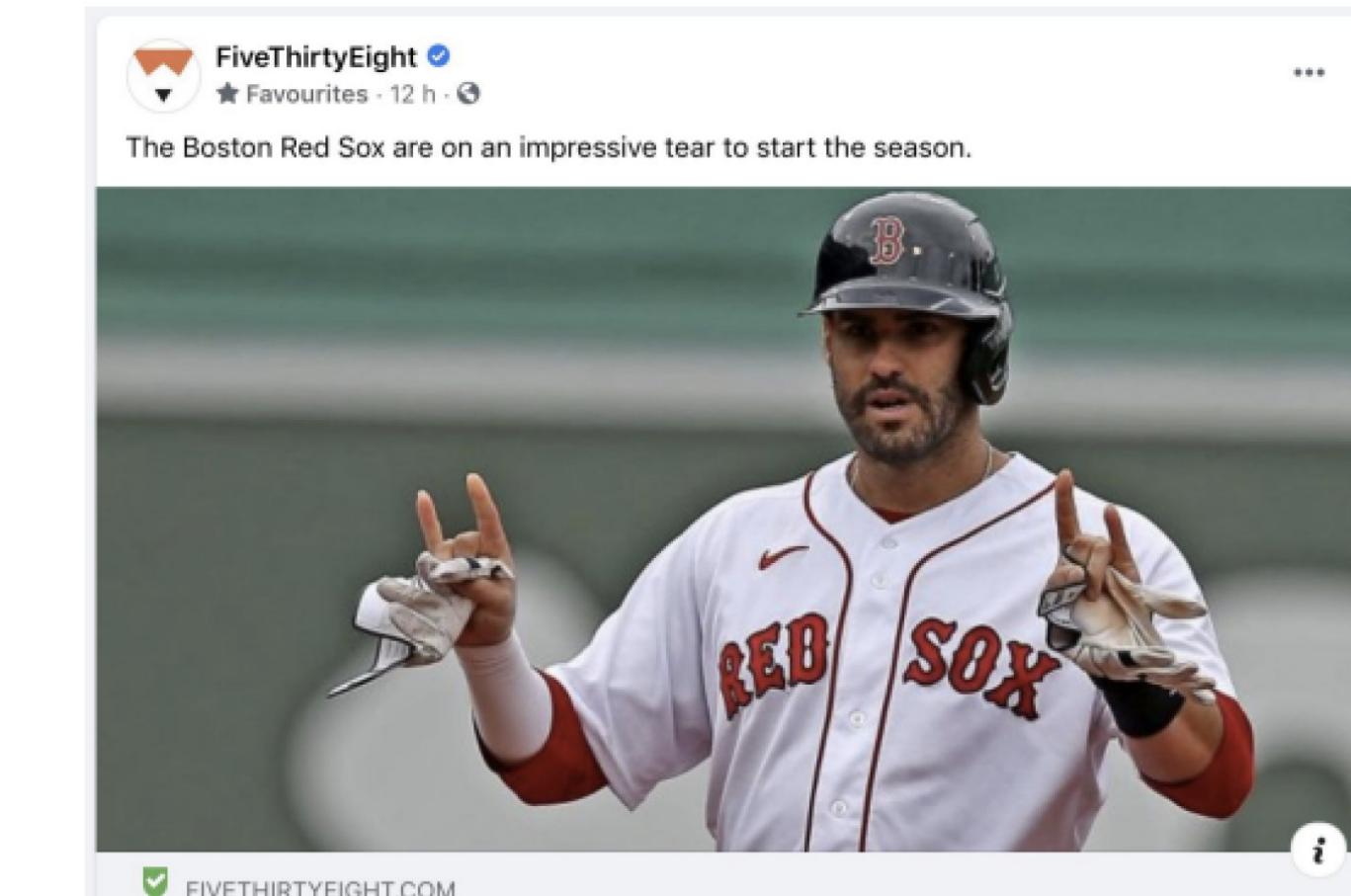
- However, labeling some stories as false leads people to believe that everything not explicitly labeled as false... is true.
- This is problematic when fact checkers can only check a tiny percentage of all content on the site



Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management science*, 66(11), 4944-4957.

# Source credibility signals

- Technique: flag source trust level (e.g., NewsGuard)
- Unfortunately, these signals do not change news consumption behavior



Aslett, K., Guess, A. M., Bonneau, R., Nagler, J., & Tucker, J. A. (2022). News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions. *Science advances*, 8(18), eabl3844.

# Priming accuracy

- Bringing attention to the accuracy of information shared on Twitter improves the quality of news shared later
- Why?
  - Because we're not in a post-truth world, where people don't care about accuracy. We instead tend to be more focused on other motivators, like pleasing our followers

HOW ACCURATE IS THIS HEADLINE?



Woman who had ovary frozen in childhood give...

She is believed to be the first woman in the world to have a baby after having ovarian tissue frozen befo...  
[surveycamel.com](http://surveycamel.com)

Thanks for following me! Can I ask you a favor? I'm wondering how accurate the above headline is, and I'm doing a survey to find out.

[surveycamel.com/ze/news/story5...](http://surveycamel.com/ze/news/story5...)

Based on the headline, do you think it is accurately describing something that actually happened?

Please rate as: 1=Not at all accurate, 2 = Not very accurate, 3= Somewhat accurate, 4 = Very accurate

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), 590-595.

# Priming accuracy at scale



IrrationalLabs  
@IrrationalLabs

We designed an intervention that reduced shares of flagged content on TikTok by 24% via a large scale RCT, thread  1/7



IrrationalLabs @IrrationalLabs · Feb 3

We put a short prompt on videos that reminded people to think about the accuracy of the content they were watching. And then - when people went to share the video - we reminded them again that the video was flagged & asked them if they were sure they wanted to share. 3/7

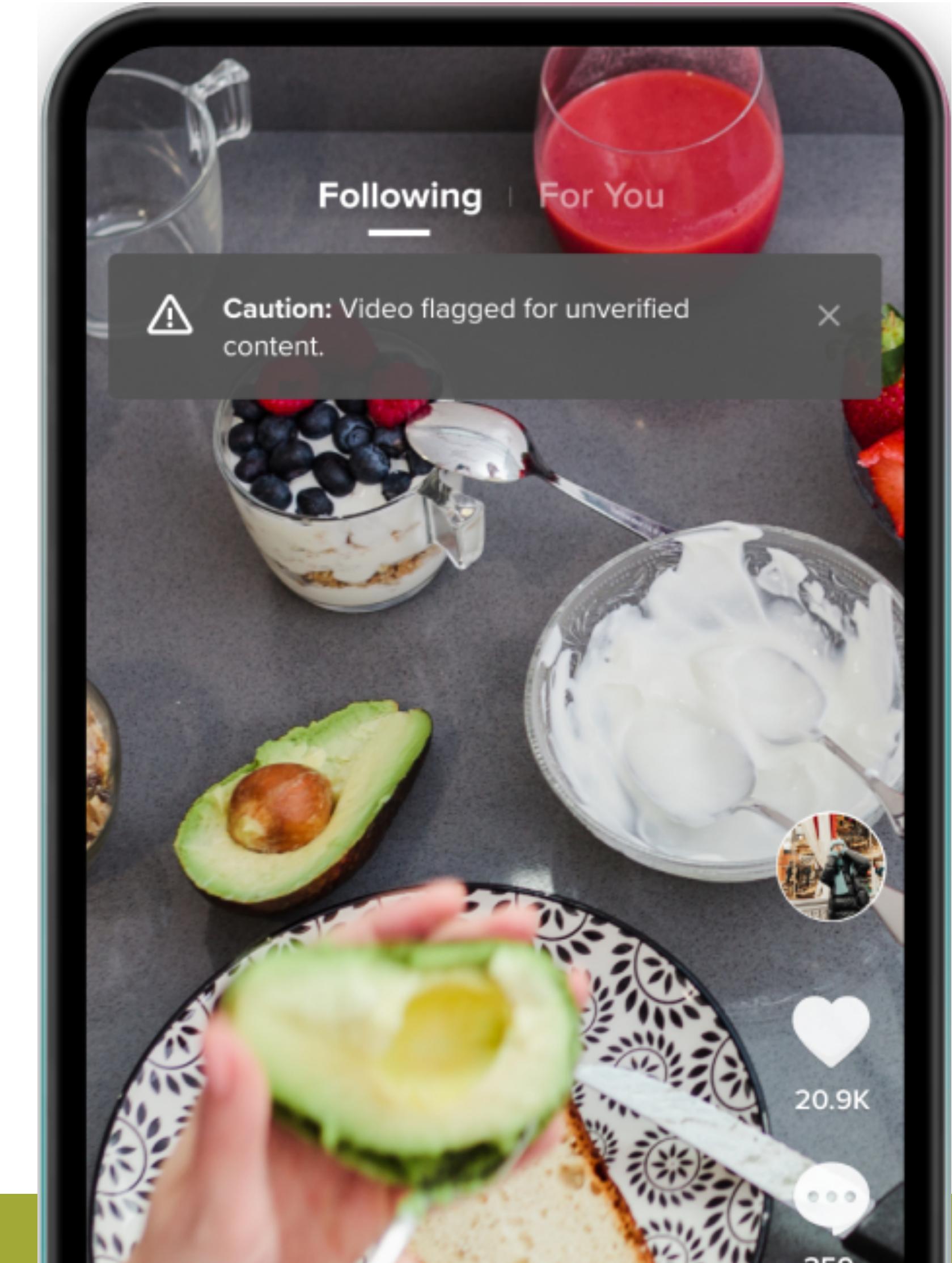


IrrationalLabs @IrrationalLabs · Feb 3

In addition to successfully reducing shares by 24%, our intervention also reduced likes by 7%, and views by 5%. 6/7

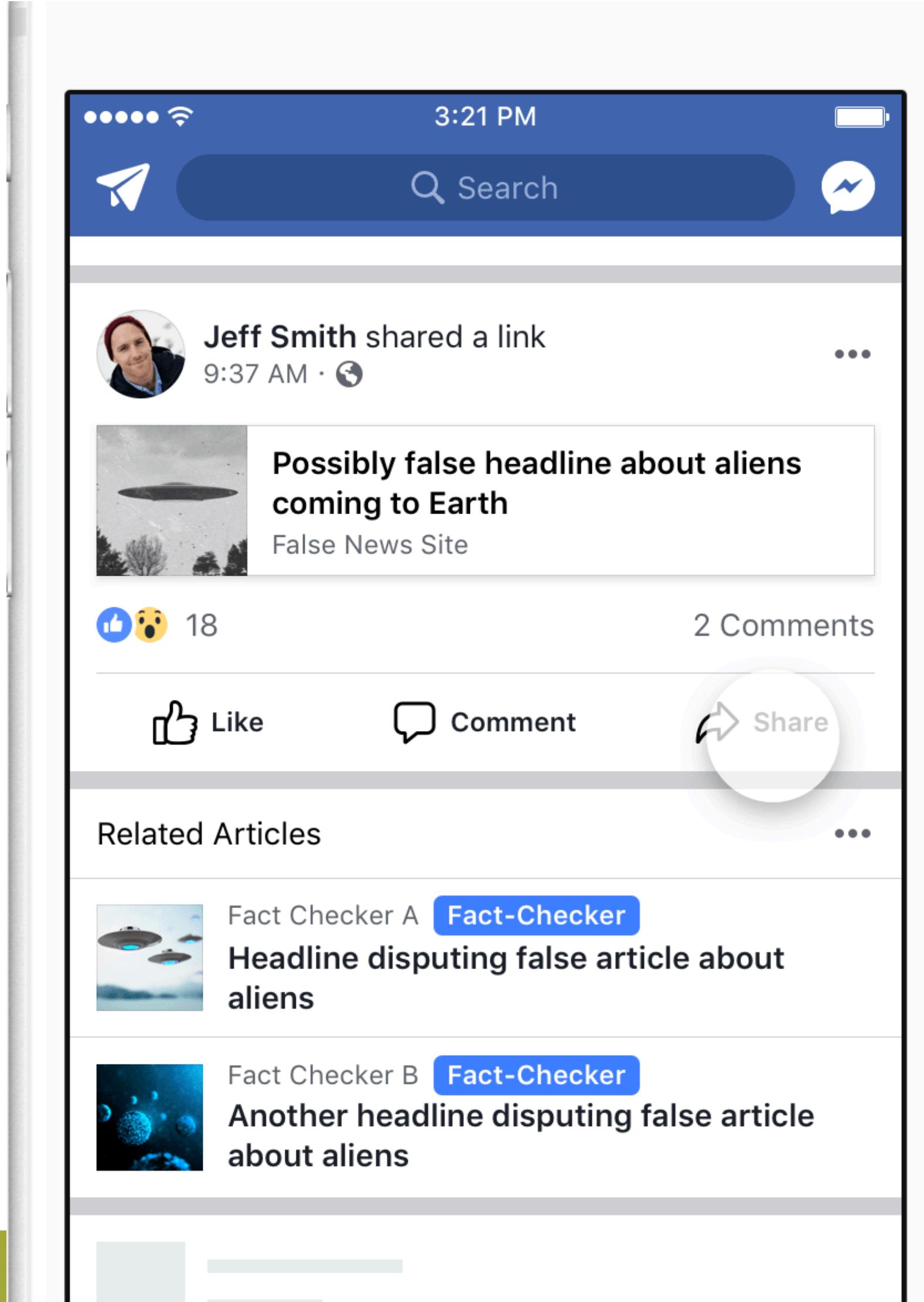
<https://twitter.com/IrrationalLabs/status/1357033901311451140>

...

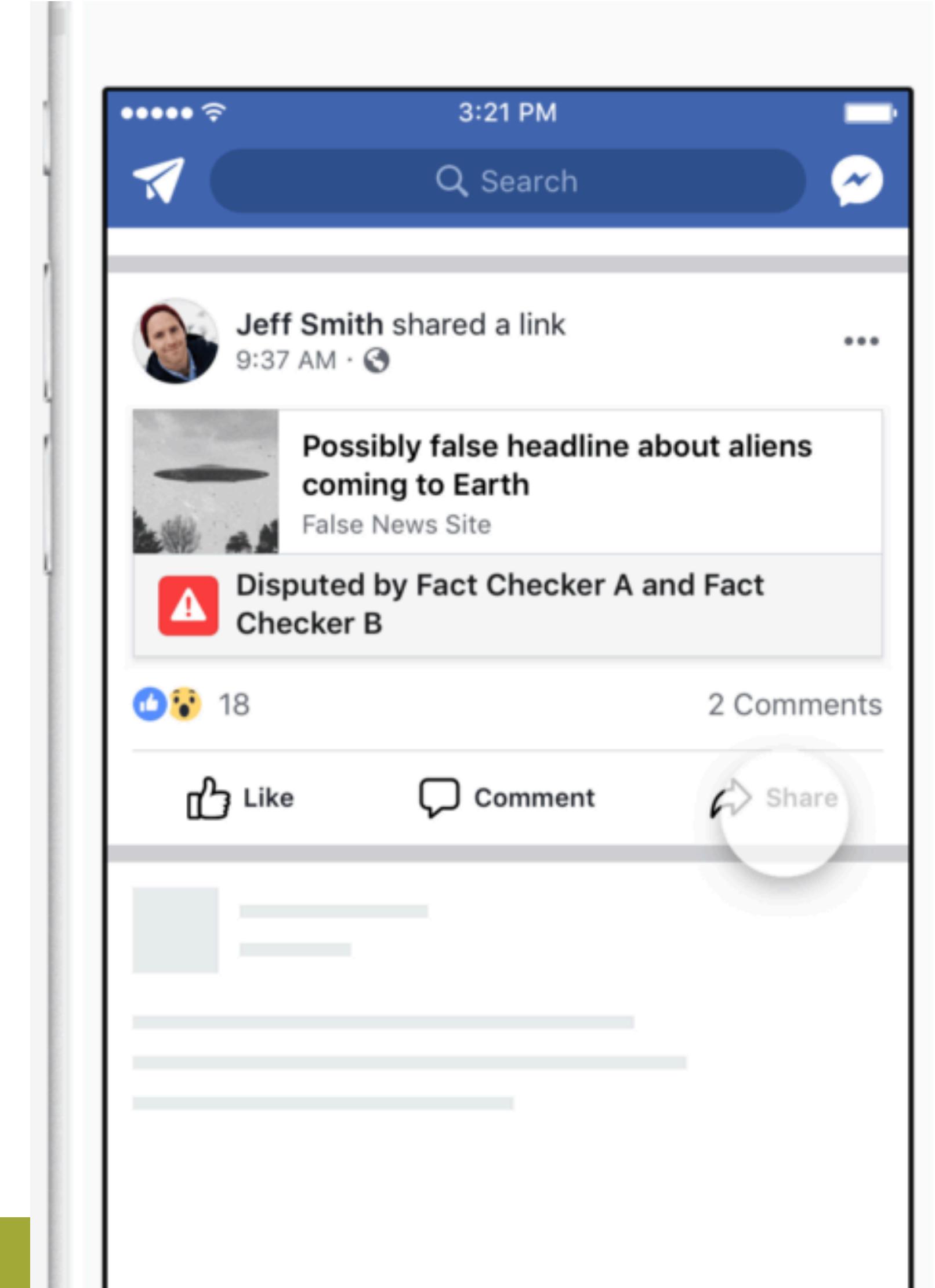


# Back to this question: Which of these will reduce misinformation more?

Related articles

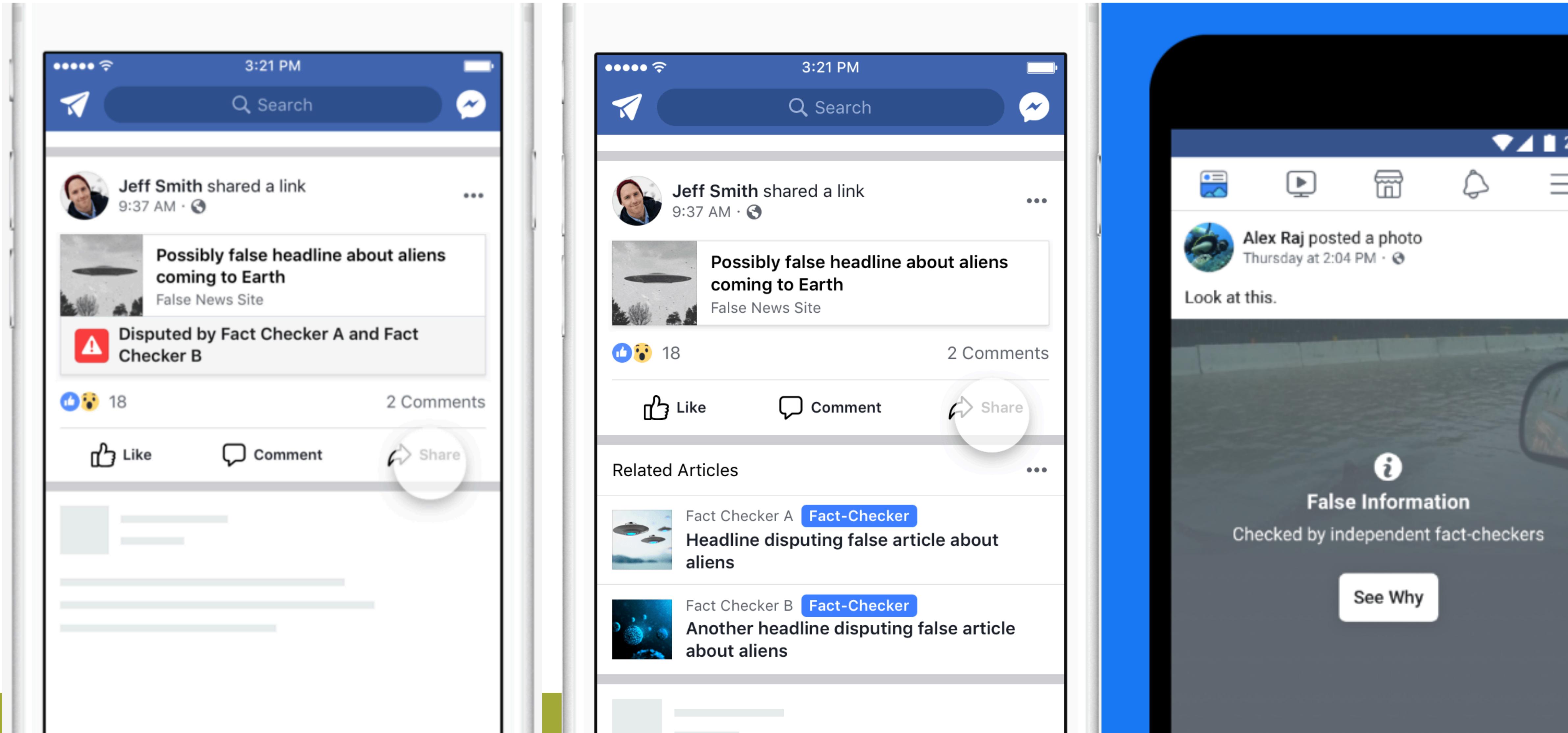


Fact check



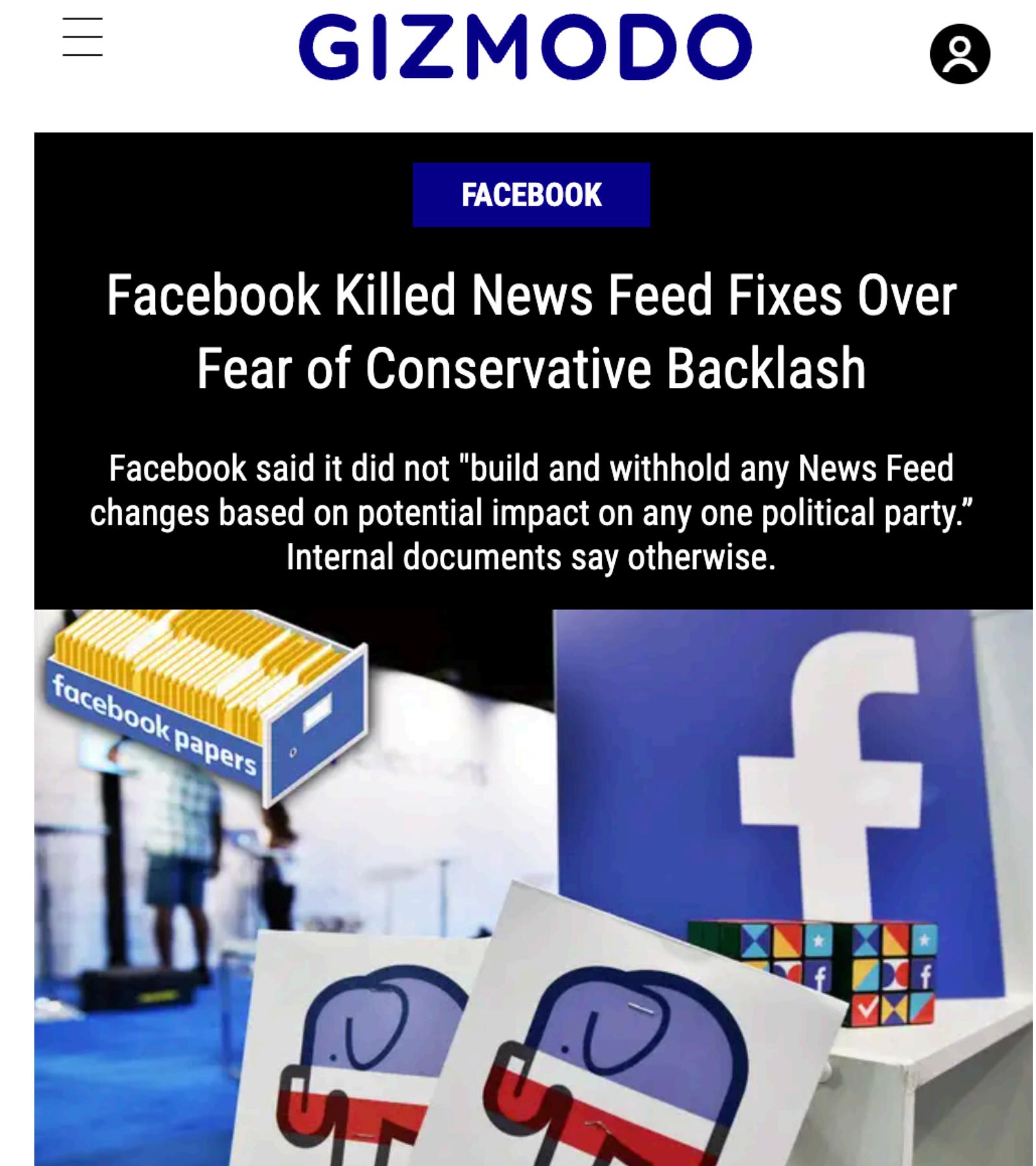
# Facebook's arc

Fact check → Related articles → Fact check



# No politically neutral option

- There exists vastly more conservative-leaning disinformation than liberal-leaning disinformation
- This difference persists even if the links are evaluated by balanced groups, or by groups of only conservatives
- So the issue is hot-button political, in addition to intersecting questions of freedom of expression



Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 211-236.

Mosleh, M., Yang, Q., Zaman, T., Pennycook, G., & Rand, D. (2022). Trade-offs between reducing misinformation and politically-balanced enforcement on social media.

# What should platforms do about the political bend of misinformation?

---

Nobody has responded yet.

Hang tight! Responses are coming in.

# Summary

- Misinformation is not the same as disinformation
- Disinformation is often created and amplified collectively by motivated actors and their audience
- People share misinformation when they are not paying enough attention to accuracy cues
- Fact checkers are often slow
  - But, it might be possible to classify campaigns rather than information items

# Today's goals

By the end of today, you should be able to...

- Differentiate misinformation from disinformation
- Reason through theories around why misinformation spread is more prominent in the digital era
- Articulate the role of varied stakeholders such as conventional media, bystanders, and fact checkers in advancing and minimizing spread
- Explain technological design strategies for minimizing the spread of misinformation on social media

# **IN4MATX 153: CSCW**

**Class 16:**  
**Health Forums**

Professor Daniel Epstein  
TA Dennis Wang  
Reader Weijie Du