The Double-Edged Sword of Banning Extremists from Social Media

Sam Jackson

University at Albany

 @sjacks26

jack ✔ @jack · Jul 21
Fundamentally, we need to focus more on the conversati
Twitter. We haven't paid enough consistent attention here
more context, helping to identify credibility, ease of use.

Challenging work and would love to hear your thoughts a

💬 1.4K   ⟲ 135   ♡ 806   ✉

jack ✔ @jack · Jul 21
Been thinking more about this. Has anyone done a rigorous, nonpartisan,
dispassionate retrospective of coverage? I ask out of intent to learn, not to
blame or single out channels or mediums. Reply with links if any.

💬 183   ⟲ 43   ♡ 201   ✉

Follow

Replying to @jack

You're incapable of learning. Ban the Nazi or
fuck off

k ✔ @jack · Jul 21
ndamentally, we need to focus more on the conversational dynamics within
tter. We haven't paid enough consistent attention here. Better organization,
re context, helping to identify credibility, ease of use.

allenging work and would love to hear your thoughts and ideas

jack ✔ @jack · Jul 30
An update! We've selected 2 partners from 230 idea submissions. Our first goal
is working to measure the "health" of public conversation, and that
measurement be open and defined by third parties (not by us).

Follow

Replying to @jack

Ban the Nazis you fucking Nazi

1:25 PM - 22 Jul 2018

jack ✔ (
We're co
civility o
towards

💬 7.6K

@jack when are you going to ban the Nazi's?

2:40 PM - 20 Dec 2017

Follow

rk with to create health metrics for Twitter

blog.twitter.com

💬 1.2K   ⟲ 150   ♡ 318   ✉

Replying to @jack

JUST.
GET.
RID.
OF.
THE.
NAZIS.

11:38 AM - 1 Mar
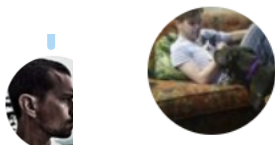
Replying to @jack

Or maybe ban Nazi's and people who
support Nazi's. Stop normalizing white
supremacists, homophobes, and
misogynists.

10:53 AM - 22 Jul 2018

engin

1.4K

Follow

Replying to @jack

ack

Less nazi's on here would be
greeeeeeaaaaat

get rid of all the Nazi/fascist accounts

Jul 2018

Debates about banning extremists

1. Civil liberties versus public safety/security

2. Technical challenges of identify extremists, appeals process

3. Consequences of banning versus tolerating

What is the goal of banning extremists

1. Reducing extremist violence
2. Reducing extremism

What approaches can be taken to extremists

1. Removing extremists
2. Tolerating extremists

	Goal	
	Reduce Extremist Violence	Reduce Extremism
Approach Tolerate extremist users		
Remove extremist users		

Goal: Reducing extremist violence

Approach: tolerating extremists

- Extremists exposed to differing opinions, no echo chamber of awful, perhaps less chance of individuals becoming more willing to use violence
  - Idea X seems pretty extreme

Approach: removing extremists

- Extremists pushed into marginal spaces (like gab.ai and 4chan), less interaction with non-extremists, more echo chamber of awful
  - Idea X seems less extreme
- Being banned increases their sense of persecution, fueling grievances and the desire to fix grievances

Goal: Reducing extremism

Approach: tolerating extremists

- More opportunities for recruitment and mainstreaming of ideas (Overton window)
- Possibilities for deradicalization or disengagement through interaction with non-extremists

Approach: removing extremists

- Fewer opportunities for recruitment or mainstreaming of ideas
- Fewer opportunities for deradicalization and disengagement

	Goal	
	Reduce Extremist Violence	Reduce Extremism
Approach Tolerate extremist users	?	X
Remove extremist users	X	✔

Caveats

Banning extremists might make an individual more likely to use violence, while tolerating extremists might make an individual less likely to use violence.

- **BUT** tolerating extremists might mean a larger population of people who might use violence, so even if individual likelihood goes down, overall likelihood could go up
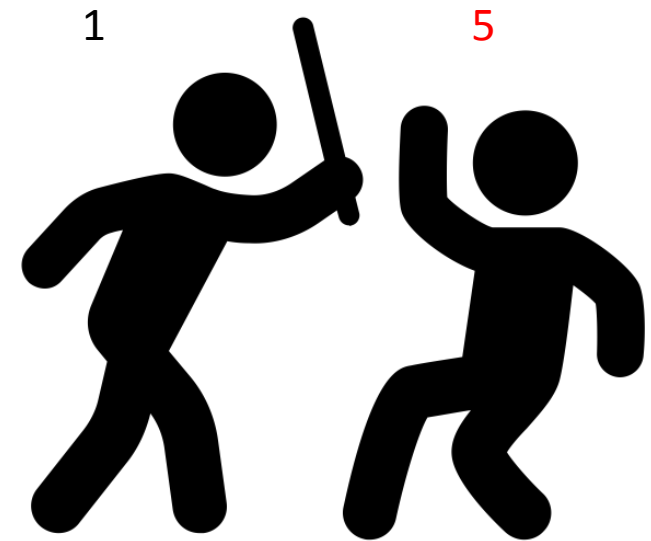
Consequences of banning extremists

1   5

Created by Adrien Coquet
from Noun Project

1   5
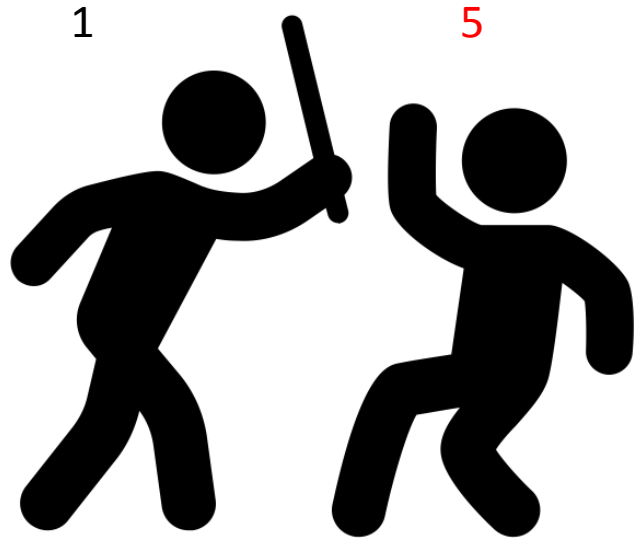
Created by Adrien Coquet
from Noun Project

X 2 extremists

10

Consequences of tolerating extremists

X 200 extremists

20

Banning extremists

Likelihood of violence = individual propensity for violence x number of individuals

Likelihood of violence given banning = increased individual propensity for violence x decreased number of individuals

@jack when are you going to ban the Nazi's?

2:40 PM - 20 Dec 2017

Decision makers need to decide what their goal is before creating policy to pursue that goal.

- Reduce extremist violence
- Reduce extremism

The Double-Edged Sword of Banning Extremists from Social Media

Sam Jackson

University at Albany

@sjacks26