IFIP 10.4 Workshop
Dependable News
Issues in Counter Deception

O. Sami Saydjari, Cyber Defense Agency, Inc.

23 June 2017

All warfare is based on deception.

The skillful leader subdues the enemy's troops without any fighting; he captures their cities without laying siege to them; he overthrows their kingdom without lengthy operations in the field.

. -Sun Tzu, The Art of War

Some Definitions

- Misinformation—unintentionally incorrect information
- Disinformation—intentionally incorrect information as part of deception
- Deception—actions (including spreading disinformation) taken to intentionally mislead
- Information warfare—use of information and information technology to gain an advantage (includes deception)
- Psychological Operations (PsyOps)—action by psychological methods to cause a planned psychological reaction
- Influence operations—coordinated effort to alter target's attitudes, decisions and behaviors toward influencers interests

Purpose

- View as Countering disinformation/deception campaigns
- Take on a very hard problem
  - **Step 1: Understand the problem**
  - Step 2: Propose Solutions
  - Step 3: Evaluate Solutions
  - Step 4: Implement Solutions (before 2020)
- Saying the problem is "insoluble" is useless fatalism
- Resisting "Not my job"
  - Resisting technology people who say this in solely a psycho-social problem
  - Resisting psycho-social people saying that it's a technology problem

Background

- What's at Stake—Why this is worth my time (and yours)
  - Democracy, worldwide
    - Possible disinformation campaign to undermine trust in democratic process itself
  - Science, worldwide
  - Freedom, worldwide
- Input from Many Subject Matter Experts
  - Cybersecurity experts
  - National Security Experts
  - Policy wonks
  - Journalists
  - Technologists
  - Psychologists
  - Sociologists
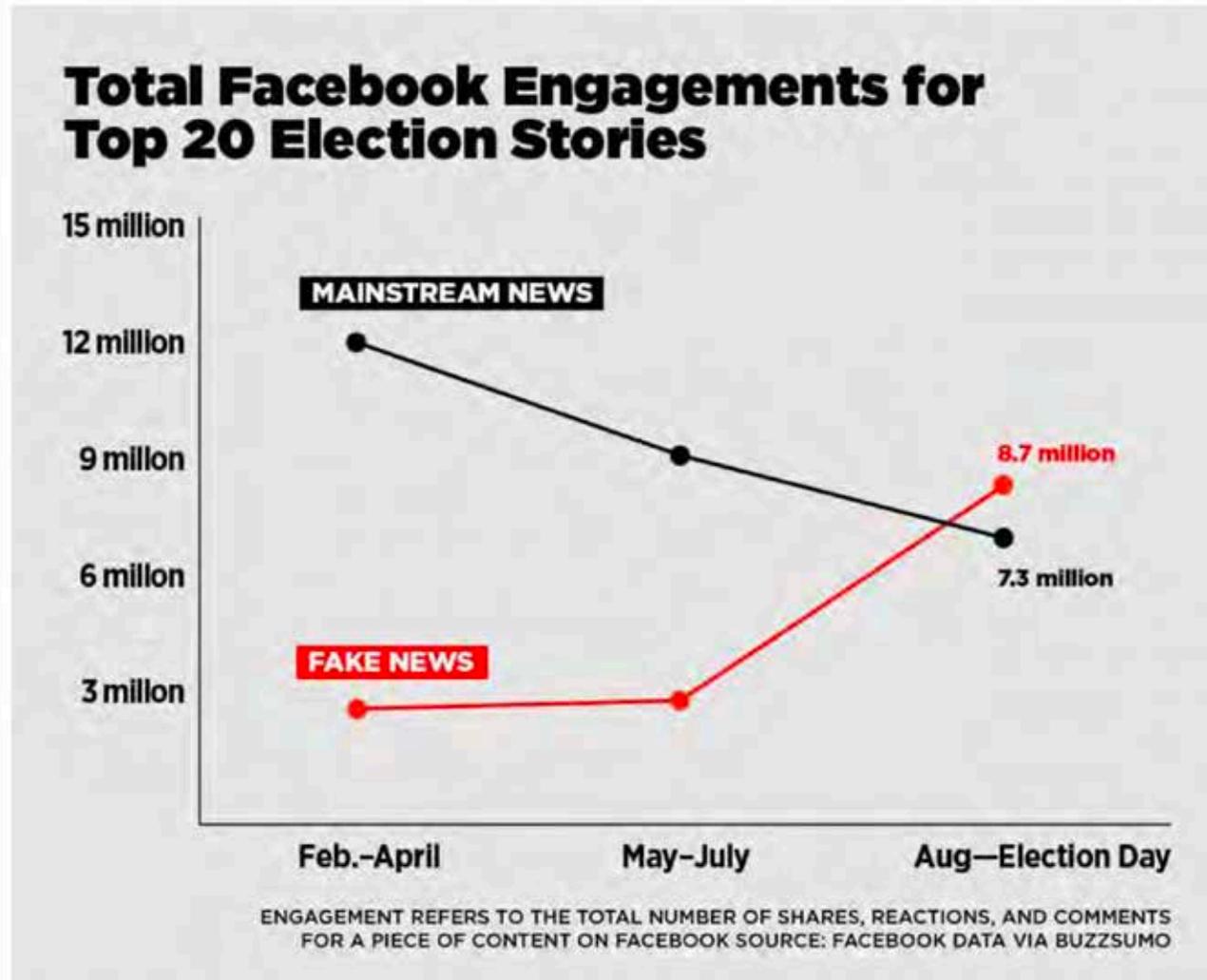
Key background re fake news

- Low cost of widespread info sharing
- Proliferation of news sources
- Collapse of trust in media on the right (47% in 2000 to 14% now)
- Monetization of attention on Internet
- *Biased assimilation of information*
- *Lack of "real world" feedback if your beliefs are wrong*
- *People don't know a lot about politics (~30% didn't know who was VP, 65% a single SC justice, etc)*

Source: David Lazer, Northeastern University/Harvard

Goals

- Understand vulnerability to influence
- Understand vulnerability to *behavior* change
  - Can belief changes change votes or decision of whether to vote?
  - How big of a behavior change is possible?
  - How to measure the effect (IC Report)
- Understand underlying reasons for these vulnerabilities
- Understand what counter strategies work and don't and why
- Understand the vector of the problem and how bad bad can get
- Create a DARPA investment program??

"Fake news" distribution

Total Facebook Engagements for Top 20 Election Stories

MAINSTREAM NEWS

FAKE NEWS

8.7 million

7.3 million

15 million

12 million

9 million

6 million

3 million

Feb.–April   May–July   Aug—Election Day

ENGAGEMENT REFERS TO THE TOTAL NUMBER OF SHARES, REACTIONS, AND COMMENTS FOR A PIECE OF CONTENT ON FACEBOOK SOURCE: FACEBOOK DATA VIA BUZZSUMO

(Silverman 2016)

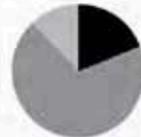# Exposure to/belief in "fake news"

"Have you seen or heard about the following story in the past few weeks?"

FBI Agent Suspected in Hillary Email Leaks Found Dead in Apparent Murder-Suicide
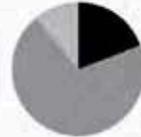(1,809 RESPONDENTS)
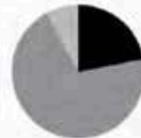
22% Yes
70% No
8% Unsure

Donald Trump Protester Speaks Out: "I Was Paid $3,500 to Protest Trump's Rally"
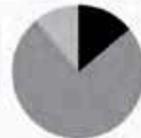(1,809 RESPONDENTS)

19% Yes
69% No
12% Unsure

Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement
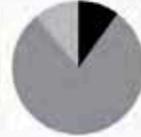(1,809 RESPONDENTS)

19% Yes
69% No
11% Unsure

Donald Trump Sent His Own Plane to Transport 200 Stranded Marines
(1,809 RESPONDENTS)

14% Yes
75% No
11% Unsure

FBI Director Comey Just Put a Trump Sign On His Front Lawn
(1,809 RESPONDENTS)

10% Yes
79% No
11% Unsure

Fake Headline Perceptions of Accuracy

The percentage of respondents who recalled seeing the headline and who believe it to be very or somewhat accurate.

FBI Agent Suspected in Hillary Email Leaks Found Dead in Apparent Murder-Suicide
(389 RESPONDENTS)

72%

Donald Trump Protester Speaks Out: "I Was Paid $3,500 to Protest Trump's Rally"
(348 RESPONDENTS)

79%

Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement
(330 RESPONDENTS)

64%

Donald Trump Sent His Own Plane to Transport 200 Stranded Marines
(263 RESPONDENTS)

84%

FBI Director Comey Just Put a Trump Sign On His Front Lawn
(186 RESPONDENTS)

81%

(Silverman and Singer-Vine 2016)

Useful Analogies

- Environmental Biosphere
  - Misinformation sludge polluting into the "infosphere"
  - Detect and eliminate major sources
  - Liability to polluters
  - Global damage from local pollution

- Epidemiology, biological
  - Memes spread like biological viruses
  - Vulnerable communities tend to aggregate
  - Homogeneity can be devastating
  - Notion of inoculation—expose recipients to counter argument in advance

- Dependable Systems
  - Creators, distributors, consumers
  - Faults, Errors, Failures (see next slide)

- Warfare/Conflict Models
  - Information Warfare → War on/with information
  - Military supremacy no longer sufficient (uncomfortable) [Art of War, Sun Tzu

Misinformation $\xrightarrow{\text{activation}}$ **Misbelief** $\xrightarrow{\text{propagation}}$ **Misbehavior**

Fault $\xrightarrow{\text{activation}}$ **Error** $\xrightarrow{\text{propagation}}$ **Failure**

Misinformation Versus Misperception

- Misinformation is false information

- Misperception concerns mistaken beliefs

- Some misinformation leads to misperceptions in some people
  - Virus analogy—not everyone will get sick when exposed to a virus

- Some misperceptions are pre-existing

People Are Vulnerable to Disinformation

- All people are vulnerable
  - Selective Exposure--we prefer attitude consistent information
  - Directionally Motivated Reasoning–more likely to believe attitude-consistent info
  - People prefer wrong models to incomplete models {Lewandosky17]
- Even if Disinformation is Corrected
  - Correction fades over time
  - There is a "belief echo"—feel negatively despite correction [Thorson17]
- Vulnerability is not obviously correlated to education level!
- Teaching critical thinking may help, but experts seem to doubt it

Believability Amplifiers

- What you saw first
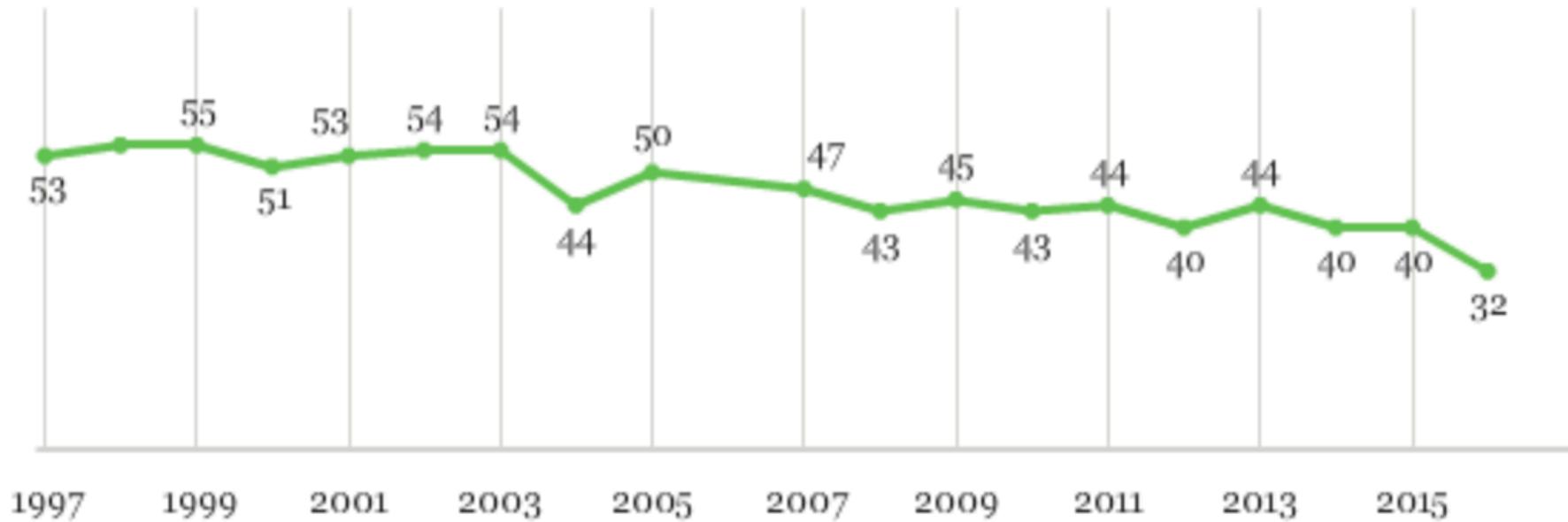 - Retractions are ineffective
- Emotional saliency, particularly outrage
- What you believe most people believe
- *What is easiest to read*
- What sounds plausible
- What is easiest to remember—KISS principle
  - "If it doesn't fit, you must _____"
  - Use small words, common vocabulary, short sentences
- What you have heard most
  - Even repeating misinformation to debunk it, reinforces it

Repetition Reinforces (Mis)information

- Repetition reinforces (mis)information
- Repetition reinforces (mis)information
- Repetition reinforces (mis)information
- Repetition reinforces (mis)information
- Repetition reinforces (mis)information
- Repetition reinforces (mis)information
- Repetition reinforces (mis)information
- Repetition reinforces (mis)information

Uncertainty—A Root Cause

- Uncertainty is a Disinformation Warrior's Friend
- Social media creates a lower barrier to entry for news stories
- Contradictory Stories Sows Uncertainty
- Attacks on Truth Amplify Uncertainty—Attack Truth Sayers
  - Attack the News media, particularly fact checkers
  - Attack the Intelligence Organizations
- Hijack Terminology to create even more confusion
  - Call everything you don't like "fake news"

Trust in Mass Media

Trust in Mass Media, by Party
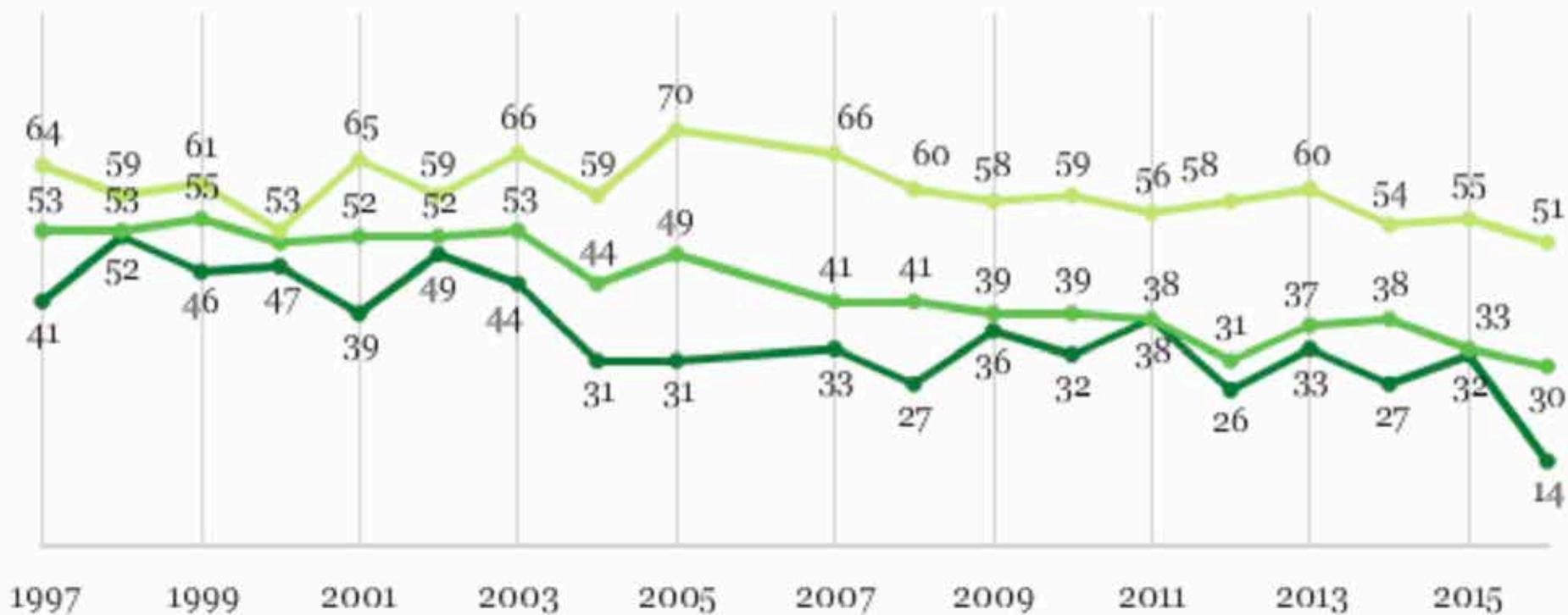
% Great deal/Fair amount of trust

- **Republicans**
- **Independents**
- **Democrats**

Republicans: 41, 52, 46, 47, 39, 49, 44, 31, 31, 33, 27, 36, 32, 38, 26, 33, 27, 32, 14

Independents: 53, 53, 55, 53, 52, 52, 53, 44, 49, 41, 41, 39, 39, 38, 31, 37, 38, 33, 30

Democrats: 64, 59, 61, 53, 65, 59, 66, 59, 70, 66, 60, 58, 59, 56, 58, 60, 54, 55, 51

1997 1999 2001 2003 2005 2007 2009 2011 2013 2015

GALLUP

Trust in Mass Media, by Age

% Great deal/Fair amount of trust

■ 18 to 49 years old ■ 50 and older

18 to 49 years old: 55, 54, 55, 45, 53, —, 42, 42, 46, 43, 44, 36, 40, 35, 36, 26

50 and older: 50, 53, 53, 44, 45, —, 54, 46, 44, 42, 43, 46, 47, 45, 45, 38

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
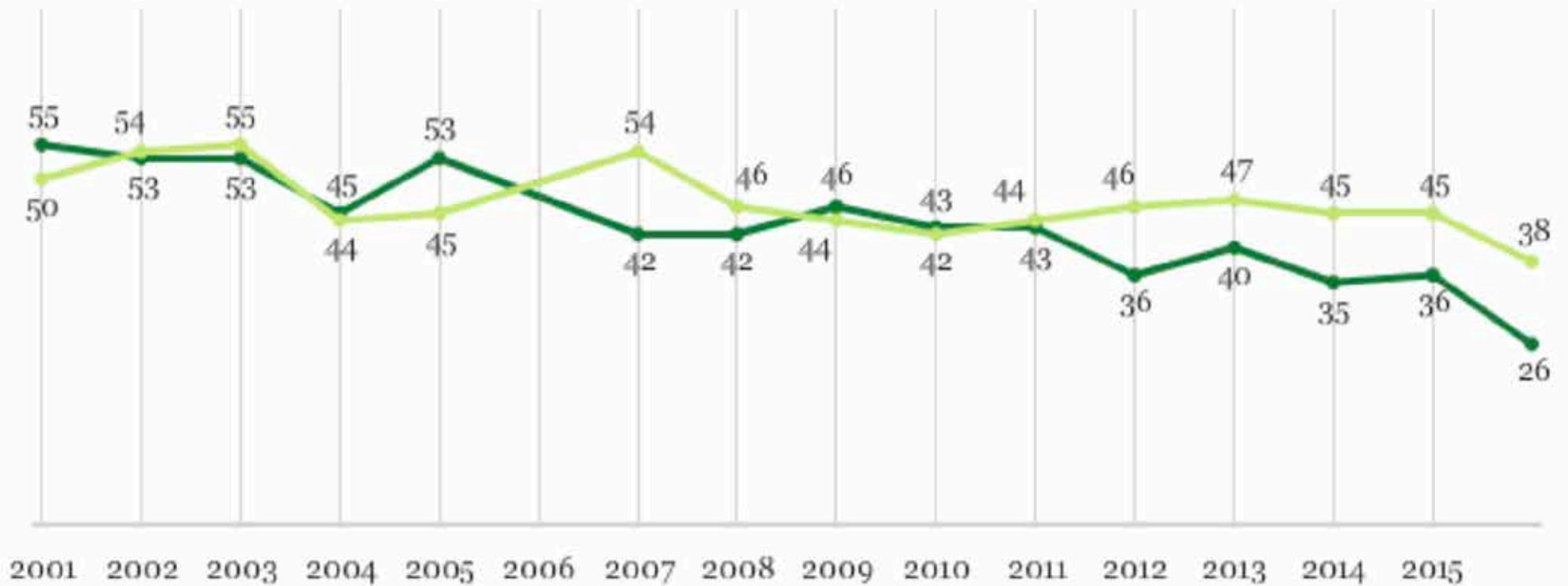
GALLUP

Antipathy—Another Root Cause

- Antipathy is a disinformation warrior's friend

- Antipathy makes galvanizing belief into action easier

- Antipathy has been increasing worldwide between parties
  - Fomenting divisiveness is in attacker's best interest, against defender's
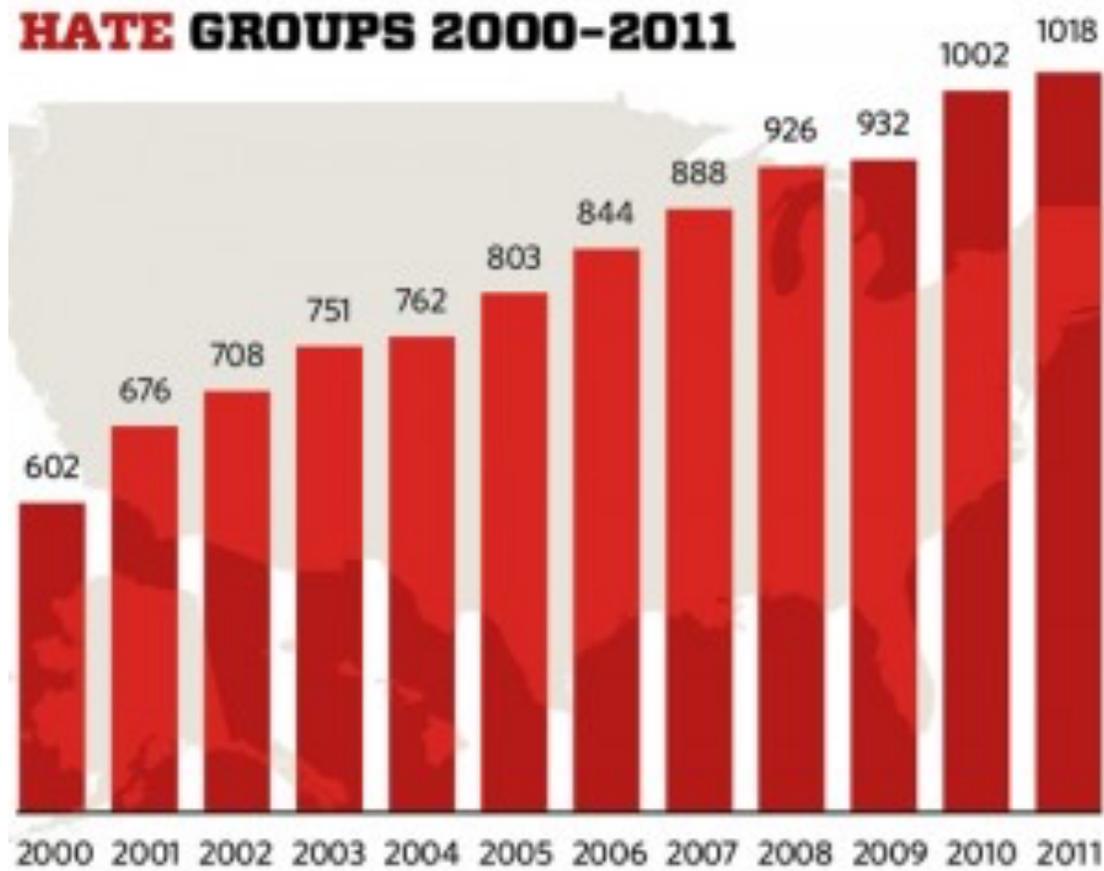 - Funding *both* sides of opposing hate groups would be a effective

Race Relations Worries Are Growing in U.S.

Percentage of Americans who worry "a great deal" about the problem of race relations
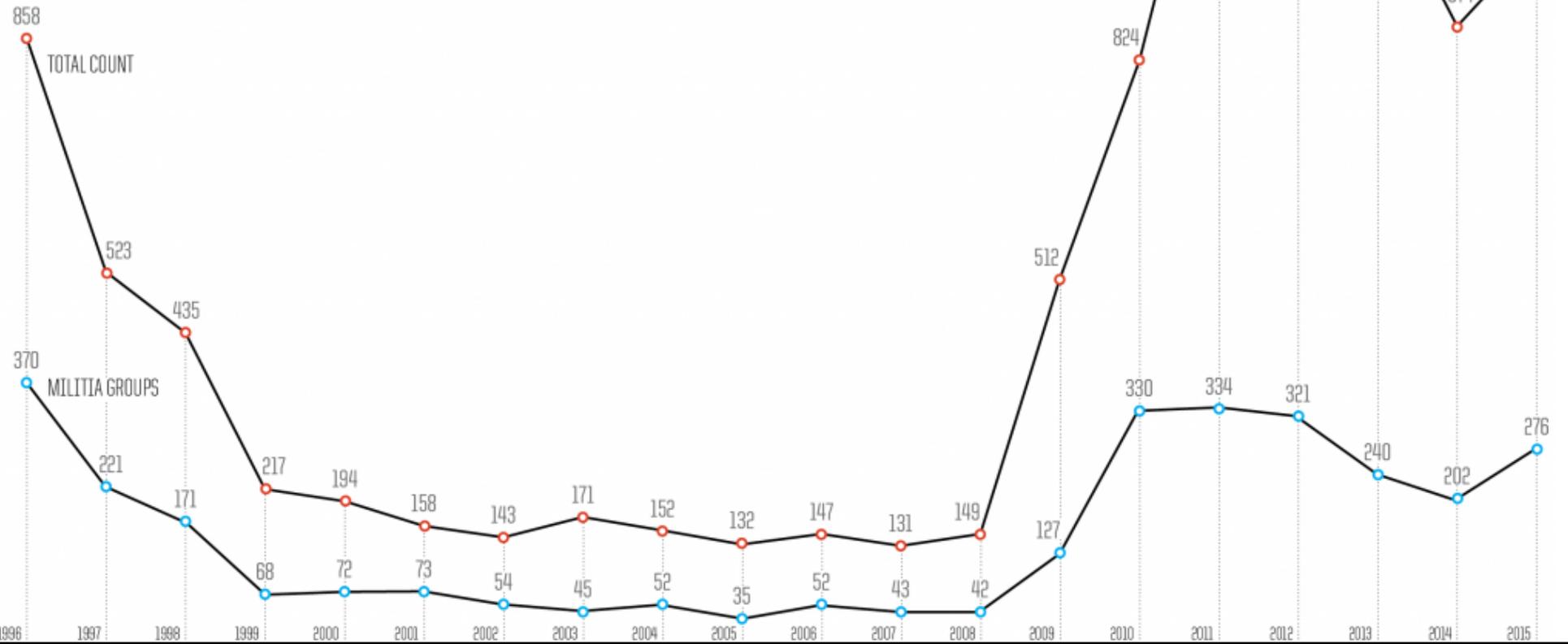
GALLUP

Hate is on the rise

HATE GROUPS 2000-2011

Year	Count
2000	602
2001	676
2002	708
2003	751
2004	762
2005	803
2006	844
2007	888
2008	926
2009	932
2010	1002
2011	1018

Source: Southern Poverty Law Center

ANTIGOVERNMENT 'PATRIOT' GROUPS 1995-2015

TOTAL COUNT

858
523
435
217
194
158
143
171
152
132
147
131
149
512
824
1274
1360
1096
874
998

MILITIA GROUPS

370
221
171
68
72
73
54
45
52
35
52
43
42
127
330
334
321
240
202
276

1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

Actors

- Botnets and Water Armies
- Sometimes its just a few users using automation

Chinese Water Armies

- Armies of 10,000 people who, for a fee, will propagate a meme of your choosing on social media
  - Hired through cut-outs
  - "iPhones are passé"
  - "Candidate X is a crook and a suspected pedophile"
  - "Company XYZ is poorly managed", then sell company short
- Botnets can also be hired to do the same
  - Human-emulating bots based on key words
  - Some flame wars can be bots yelling at each other

The Vector and End Game

- It's going to get way worse before it gets better
- Disinformation works
- Ability to do disinformation is increasingly easy
- Tools to stop it are minimal

Solution Challenges

- Freedom of Speech
- Need to Embrace Diversity of Opinion as Healthy
- Proscription Against Govts Using Propaganda on own population
- For every measure, there will be a countermeasure
- People are highly susceptible to manipulation
- People naturally gravitate towards tribalism
- Liability: Act of War? (Be careful!)

Warnings
(Ecker, Lewandowsky, & Tang, 2010)
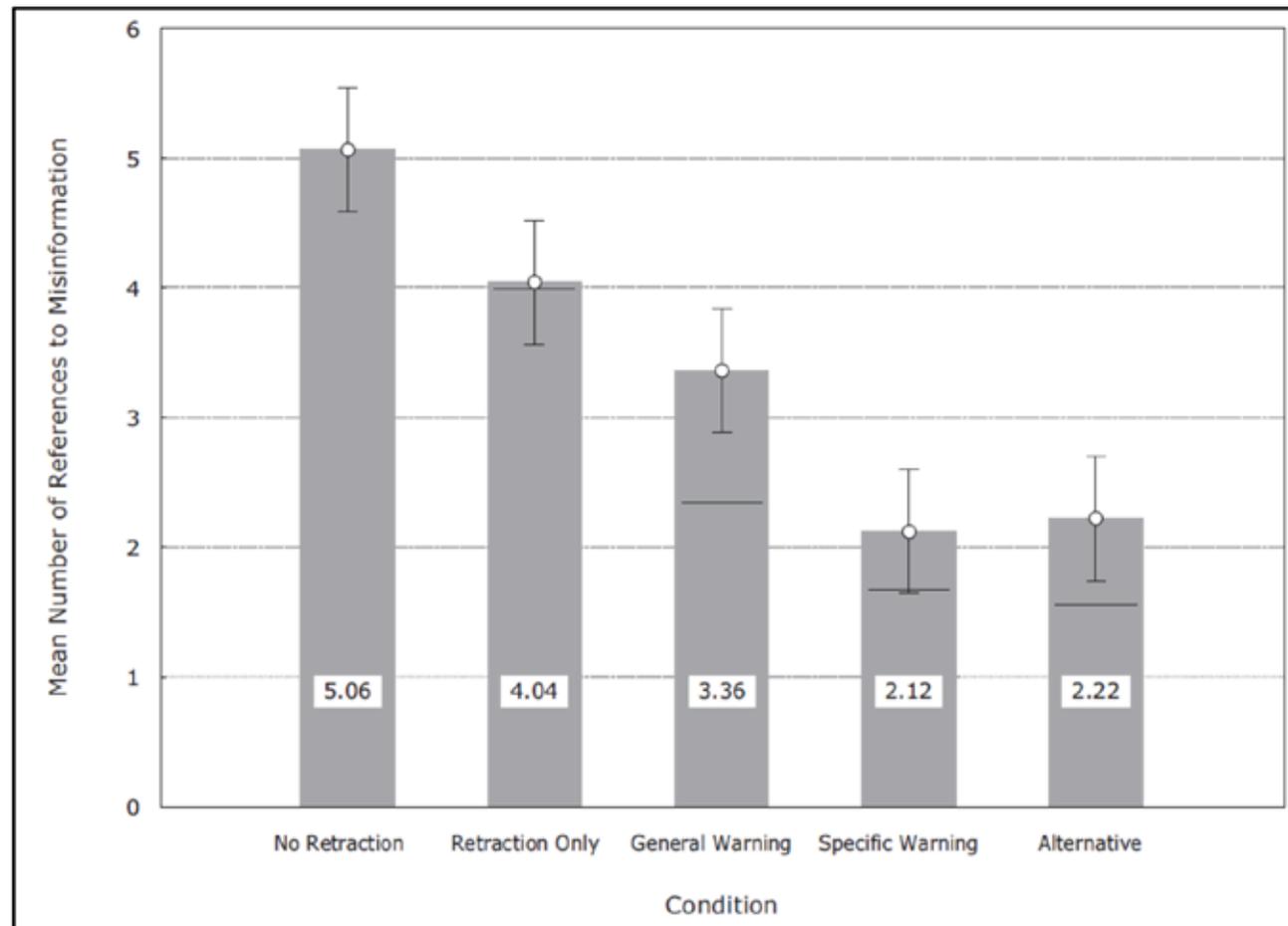
the
DEBUNKING
handbook

John Cook
Stephan Lewandowsky

Also available in
German, French,
Dutch, Polish,
Swedish, Spanish,
and Italian
http://sks.to/debunk

What's Different today

- Speed of propagation of information
  - Speed of feedback on actions taken (re-posts, likes)
  - Think in terms of control theory
- Easier to make fake news appear real
  - New York Times Printing Press, versus garage typewriter
- Increased Uncertainty and Antipathy
- Disenfranchisement
- Resistance to Correction based on belief of how many share it
  - Echo chambers distort how widespread belief

Swayable (Election) Outcomes

Impervious Left

Impervious Middle

Impervious Right

Swayable Middle

Galvanizable Swayed
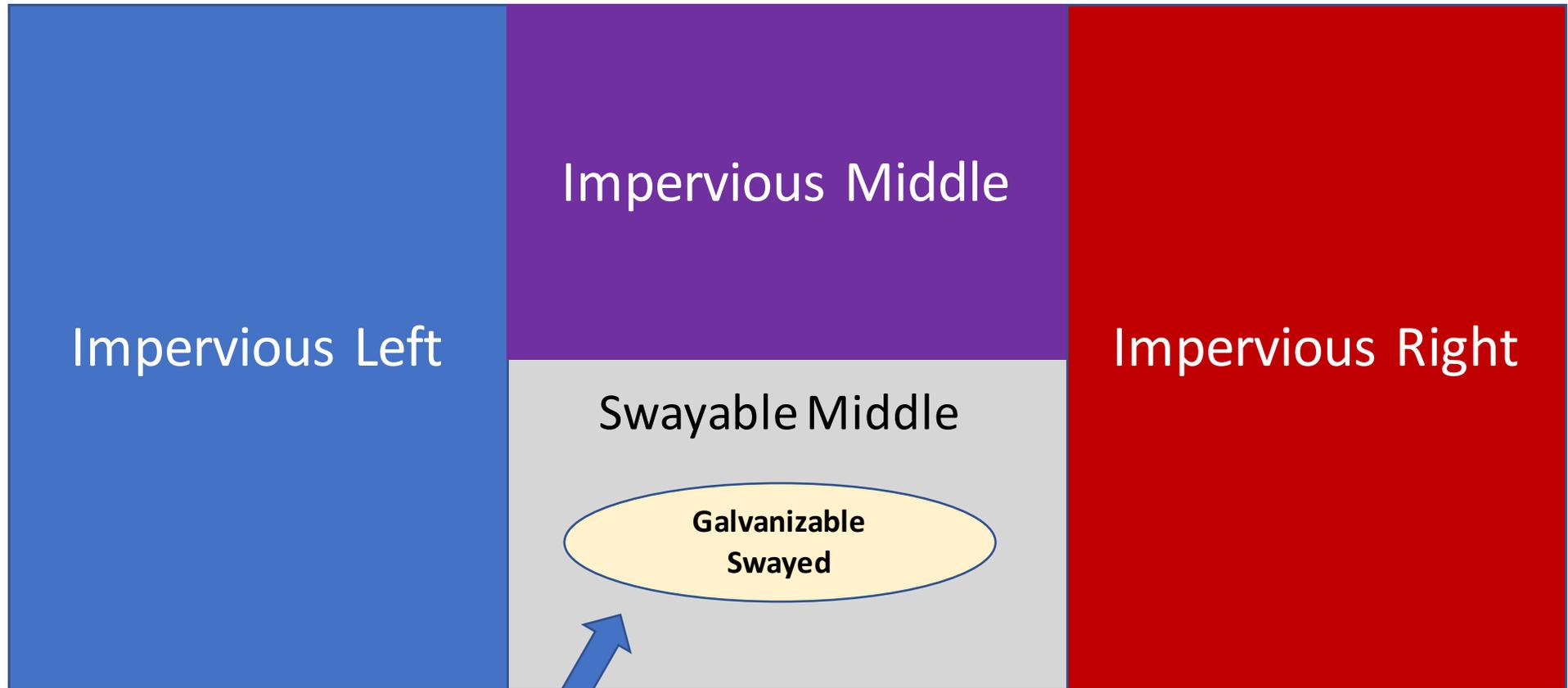
How big are these subsets and what influences their size

* Not to Scale!

Excellent Report—Must Read

Combating Fake News: An Agenda fro Research and Action, May 2017. Lazer, et. al.

https://www.northeastern.edu/nulab/combating-fake-news/

Counter-strategy Concepts

- Repeating an idea, even to debunk is a bad idea
- Inoculation may work, but it must come from a trusted source
  - Must provide alternative to non-factual information
  - Repeat Correction
- Maybe better to target credibility of source of misinformation
  - Plant basis of skepticism of source (e.g. conflict of interest)
- Filter small number of bad players [Lazer]
  - Filter bots and "cyborgs"
- Easier fact checking may help, but distrust in news media is problem
  - News Media needs to activity campaign to regain trust