

“I need to respond to some harmful content. What should I do?”

Insights and strategies from Kelsey Suter, [Drive Agency](#)

Learn More: [Race Class Narrative \(RCN\) Combatting Disinformation Fact Sheet](#)

What is counter-messaging?

A message that disrupts harmful content by offering a more proactive, truthful message to address the core fear or falsehood being spread. For example:

THEY SAY...	THEY MEAN...	WE SAY...
“She’s radical.”	<i>She doesn’t care about you.</i>	She’s spent her career helping those in need.
“She’ll do anything for political power.”	<i>She didn’t earn her place.</i>	She has a proven track record.
“She’s not ‘really’ Black.”	<i>Her nomination isn’t historic or meaningful.</i>	She’s the next part of a generational legacy of civil rights leaders.

Before responding: three factors to consider

1 REACH

To what extent has a new claim or piece of content **broken out of its original echo chamber**? Is it reaching the people we care about?

2 IMPACT

To what extent do we care if it has broken out of its echo chamber? Is it going to change the offline situation at all? **Who does it impact?**

3 RESPONSE

What’s the best plan for **responding** to impacted audiences? How do we mitigate the threat of amplification?

1 Assess REACH using the breakout scale



LEVEL 1

Spreading within a single community on a single platform, not yet broken out of niche group.



LEVEL 2

Breaking out of niche on a single platform or spreading to niche groups on multiple platforms.



LEVEL 3

Broken out of niche groups on multiple platforms, but still on social media only.



LEVEL 4A

Featured on conservative TV, newspaper, and radio media outlets only.



LEVEL 4B

Featured on mainstream TV, newspaper, and radio media outlets also.



LEVEL 5

Spread by influencers or celebrities with a large general audience reach.

Source: Dewey Square Group, learn more [here](#).

2 Assess IMPACT across audiences

HIGH

- Risk of severe, **real shift** in offline situation.
- **Short-term**, immediate shift.
- **Directly** impacts key audiences.

MEDIUM

- Risk of **moderate** shift in offline situation.
- **Medium-term** (weeks/months).
- **Less direct** impact to key audiences.

LOW

- **Limited risk** of offline shift.
- **Longer term** (months/years).
- **Very indirect** or no impact on key audiences.

3 Decide on the best RESPONSE

DIRECT RESPONSE:

Amplification Risk: HIGH

Types of direct response:

- **PREBUNKING:** the process of debunking lies, tactics, or sources before they strike.
- **DEBUNKING:** to show or explain that a lie is not true (e.g., a fact check).



Risks:

- Prebunking risks giving something more fuel before it would otherwise spread.
- Debunks don't reach as many people as misinformation.
- **Factual corrections (a.k.a. fact checks) don't reliably erase falsehoods from our brains over the long term.**

INDIRECT RESPONSE:

Amplification Risk: LOW

Types of indirect response:

- **INOCULATION:** the process of “building immunity” to a harmful narrative, so that users know to be skeptical of that narrative when they hear it.
- **COUNTER-MESSAGING:** turning up the volume on YOUR proactive message on a given topic.

Notes:

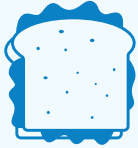
- Can be done without restating or amplifying harmful content.
- Often helpful to offer a “logic correction” instead of a “factual correction”.
Logic correction focuses on:
 - **Motives**– why disinformation is spread, who is behind disinformation (exposing and naming bad actors/villains), what do they gain from spreading lies.
 - **Networks**– calling out and naming the connected or related networks/channels that coordinate to spread lies.
 - **Tactics**– how disinformation is spread, how confirmation bias and our emotions are weaponized to spread lies.

PASSIVE RESPONSE:

Amplification Risk: LOW

Types of passive/non-response:

- Do nothing.
- Set up a monitoring flag system (social media monitoring, keyword search).
- Prepare talking points in case something moves further.
- Submit takedown requests / flag as spam or disinformation on the platform.
- Flag for important stakeholders.
- Engage the media to avoid reporting.



CONSIDER A “TRUTH SANDWICH” FOR INOCULATION MESSAGING.¹

Our brains remember what we hear first and most often!

- 1 POSITIVE:** Address the core fear/concern with a shared value.
- 2 NEGATIVE:** Offer “logic correction” by focusing on bad actors’ motives, tactics, or networks.
- 3 POSITIVE:** Pivot to your proactive message.

4 Consider how to integrate trust-building into your work

Trust is the ultimate form of inoculation against disinformation.

Trust can be reinforced with...



FEELING LISTENED TO



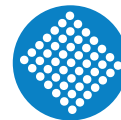
EMOTIONAL CONNECTIONS



RELATABLE CONTENT



CRITICAL THINKING



REPETITION



AUTHENTICITY

CONSIDER MORE OF THIS...

- ✓ Building relationships with influencers, surrogates.
- ✓ Repetition over longer periods of time.
- ✓ Two-way communication that rewards participation.

...INSTEAD OF THIS

- ✗ Only communicating through our owned channels.
- ✗ “Winning” the internet today.
- ✗ Over-relying on one-way paid ads.

1. Inoculation messaging attempts to build resistance against mis/disinformation in general. The idea is that preemptively exposing people to a weakened persuasive argument builds people’s resistance against future manipulation. Learn more: [A New Way to Inoculate People Against Misinformation](#).

CONSIDER MORE OF THIS...

- ✓ Networked engagement—liking, commenting, sharing content from other values-aligned partner accounts.
- ✓ Feeling frustrated and sad with your audience.
- ✓ Acknowledge that EVERYONE has been deceived by disinformation, and everyone is susceptible—but we can all do our part to fight disinformation.
- ✓ Using many channels to funnel people to one place.
- ✓ More experimentation and creativity with metrics.

...INSTEAD OF THIS

- ✗ Multiple accounts all sharing the exact same graphic.
- ✗ Always trying to get people to see more of the positive.
- ✗ Limiting the number of platforms we use.
- ✗ Basing experiments only on the metrics we have [views/likes].

Acknowledging challenges on social media

- Platforms make organic reach really hard... and it's getting harder!
- Platforms reward inflammatory content. We can—and should—call this out, demand accountability, and push for policy changes.
- Platforms increasingly prioritize professionalized content creators over grassroots networks and connections.
- Platforms don't always give us the data and metrics we need to assess impact well.
- It takes time, staff or volunteers, and resources to set up and run smaller, community-focused spaces on social media.
- It can be hard to get organizations to engage with each other.

Interested in more ways to address disinformation?

[LEARN MORE](#)