# Exploring Why We Fall for Misinformation (and What We Can Do About it)

Madeline Jalbert

Center for an Informed Public

ERS Quarterly Meeting

March 7th, 2024

CENTER FOR AN INFORMED PUBLIC

UNIVERSITY of WASHINGTON

# How do we decide if something is true?

The unicorn is the national animal of Scotland

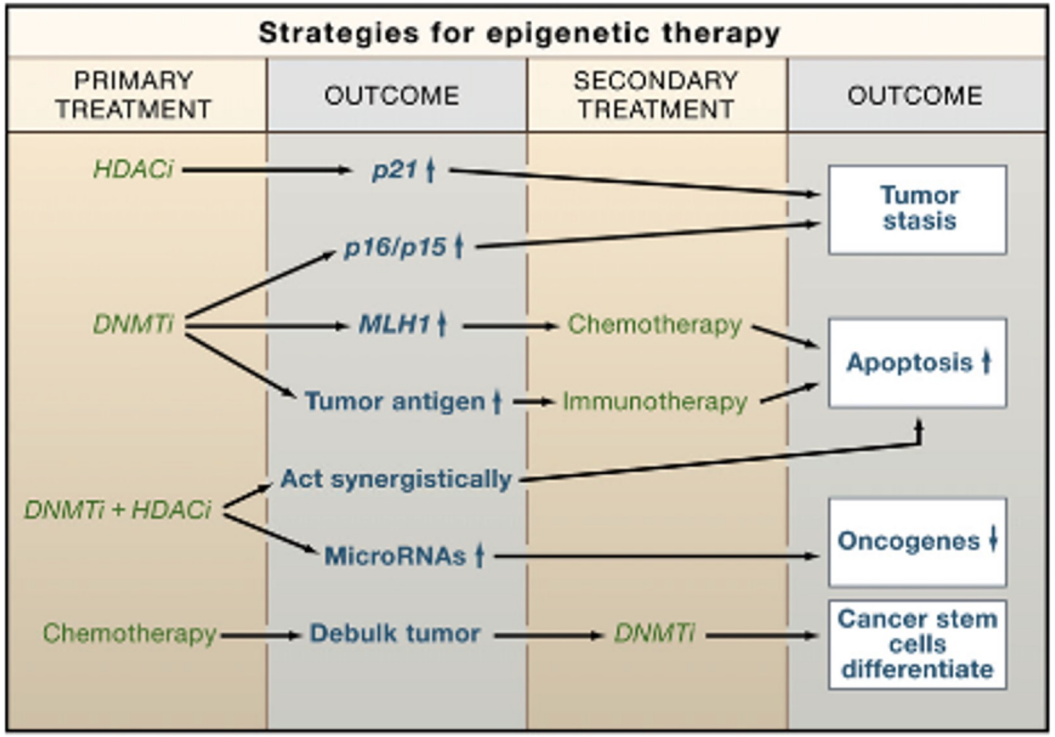
- When judging truth, people can rely on:
  - The content of the information
  - Whether or not something *feels* true

# Feelings of truth

> Information can feel easy or difficult to process

## The Epigenomics of Cancer

Peter A. Jones[1,*] and Stephen B. Baylin[2,*]
[1] Department of Urology, Biochemistry, and Molecular Biology, USC/Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA
[2] Cancer Biology Program, The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins Medical Institutions, Baltimore, MD 21231, USA



# Feelings of truth

> Information can feel easy or difficult to process

BuzzFeed

33 Products That'll Help You Get Your Health And Fitness Back On Track

# Feelings of truth

- When information feels easy to process, it also feels more true
- Ease of processing can be a valid cue that something is true
  - The arguments flow logically
  - It's consistent with what you already know
  - You've heard the information before from a credible source
- But processing ease can also result from factors unrelated to truth

# For example:

> Print font

The unicorn is the national animal of Scotland

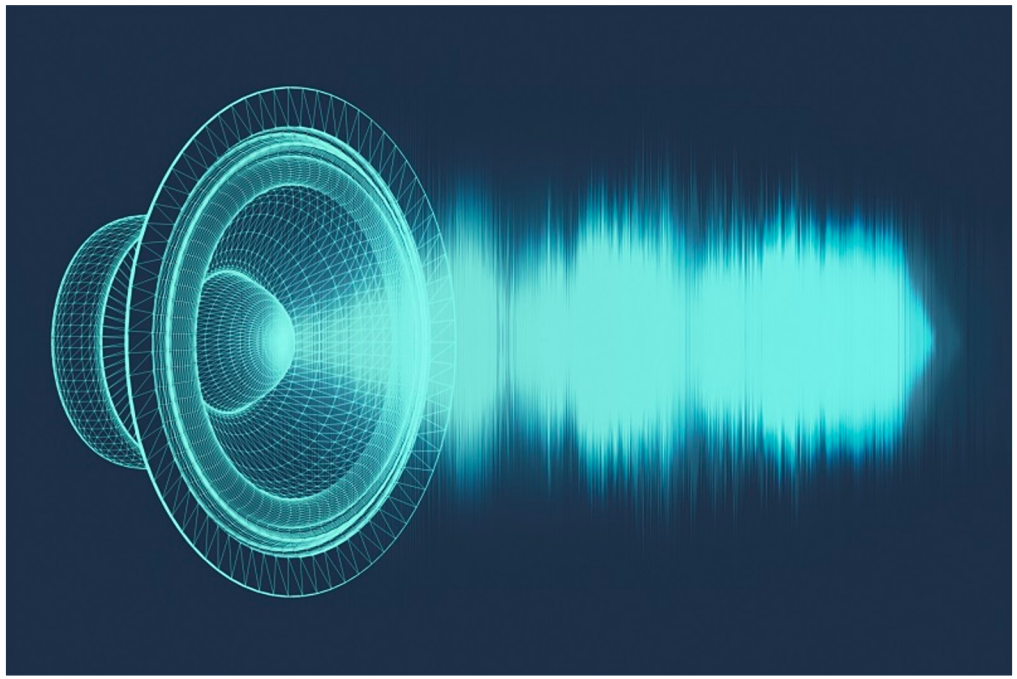
*The unicorn is the national animal of Scotland*

Song & Schwarz, 2008

# For example:

> Audio quality

The unicorn is the national animal of Scotland



Newman & Schwarz (2018)

# For example:

> Photos

The unicorn is the national animal of Scotland

Newman, Jalbert, Schwarz, & Ly, 2020;
Newman & Zhang, 2020 for a review

# For example:

> Repetition

The unicorn is the national animal of Scotland

The unicorn is the national animal of Scotland

The unicorn is the national animal of Scotland

> “Illusory Truth Effect”

Jalbert et al., 2020*;*
Hasher et al., 1977

**Question:**

How many animals of each kind did Moses take on the Ark?

# Intuitive and analytical processing

- We generally assume information is true
- **Intuitive processing** is often the default - we nod along if things feel right
- **Analytical processing** requires motivation and cognitive resources
- Distrust or suspicion can give us a reason to switch from intuitive processing to analytic processing
  - Something feels wrong
  - Explicit warnings
  - Suspicion for someone’s motives

# Truth Criteria

- **Compatibility:** Is it compatible with other things I know?
- **Coherence:** Is it internally coherent?
- **Credibility:** Does it come from a credible source?
- **Consensus:** Do other people believe it?
- **Evidence:** Is there supporting evidence?

- You can use intuitive or analytical processing to evaluate each of these criteria

Schwarz & Jalbert, 2020

# **Consensus:** Do other people believe it?

- Intuitive evaluation:
  - Does it feel familiar?

- Analytical evaluation:
  - What do my friends say?
  - What do the opinion polls say?

# **Credibility:** Does it come from a credible source?

- Intuitive evaluation:
  - Does the source *feel* familiar and trustworthy?

- Analytical evaluation:
  - Does the source have the relevant expertise?
  - Does the source have a vested interest?
  - Is the source trustworthy?

# Why do we believe misinformation?

- Assuming information is true the default
- There may not be cues to analyze information more carefully
- Misinformation may feel easy to process
- Misinformation may meet (or feel like it meets) truth criteria like source credibility or social consensus
- We may not remember or use source information
- We don’t have a better explanation

# Why do we share misinformation?

- There are many reasons why people read and share information
  - To inform others
  - Entertainment
  - To build social bonds
  - To feel like part of the group
- Truth may not be relevant

The unicorn is the national animal of Scotland

# What can we do?

- Draw attention to truth when misinformation may be present
- Make the truth *feel* true
- Appeal to truth criteria when sharing the truth
  - Present factual information in a way that is **compatible** with the audience's prior knowledge
  - Make your story **coherent**
  - Draw attention to the **credibility** of the information's source
  - Point to the social **consensus** for that information
  - Share **evidence** for your claims

# The unicorn is the national animal of Scotland

- This is true!
- National Unicorn Day: April 9th

# Online misinformation

- What can users do when they encounter misinformation online to reduce its impact?

- Fact checks can be effective
  - e.g., Bode & Vraga, 2018; Vraga & Bode, 2021

- But…
  - Fact checks take time and effort
  - Not all information is fact-checkable
  - Users may be hesitant to correct peers, especially publicly

# Alternatives


- Democracy Yethu
  - NGO in South Africa
  - Run through the Centre for Analytics and Behavioural Change (CABC)

- Developed an intervention to combat election misinformation online
  - Based on dialogue facilitation techniques

- Volunteers trained to reply to problematic posts following specific guidelines

# Alternatives


- Example replies
  - “Have you ever been there, how do you know that?”
  - ”Do you have any proof for what you’re saying?”
  - ”Where is the evidence for this?”

- Noticed that many example replies drew attention to truth

- Reminded us of interventions that involve shifting attention to accuracy
  - Pausing (Fazio, 2020)
  - Accuracy nudges (Pennycook et al., 2021)
  - Warning of falsehoods (Jalbert, Newman, & Schwarz, 2021)

- Also involves communicating social consensus of peers

# Social Truth Query Studies

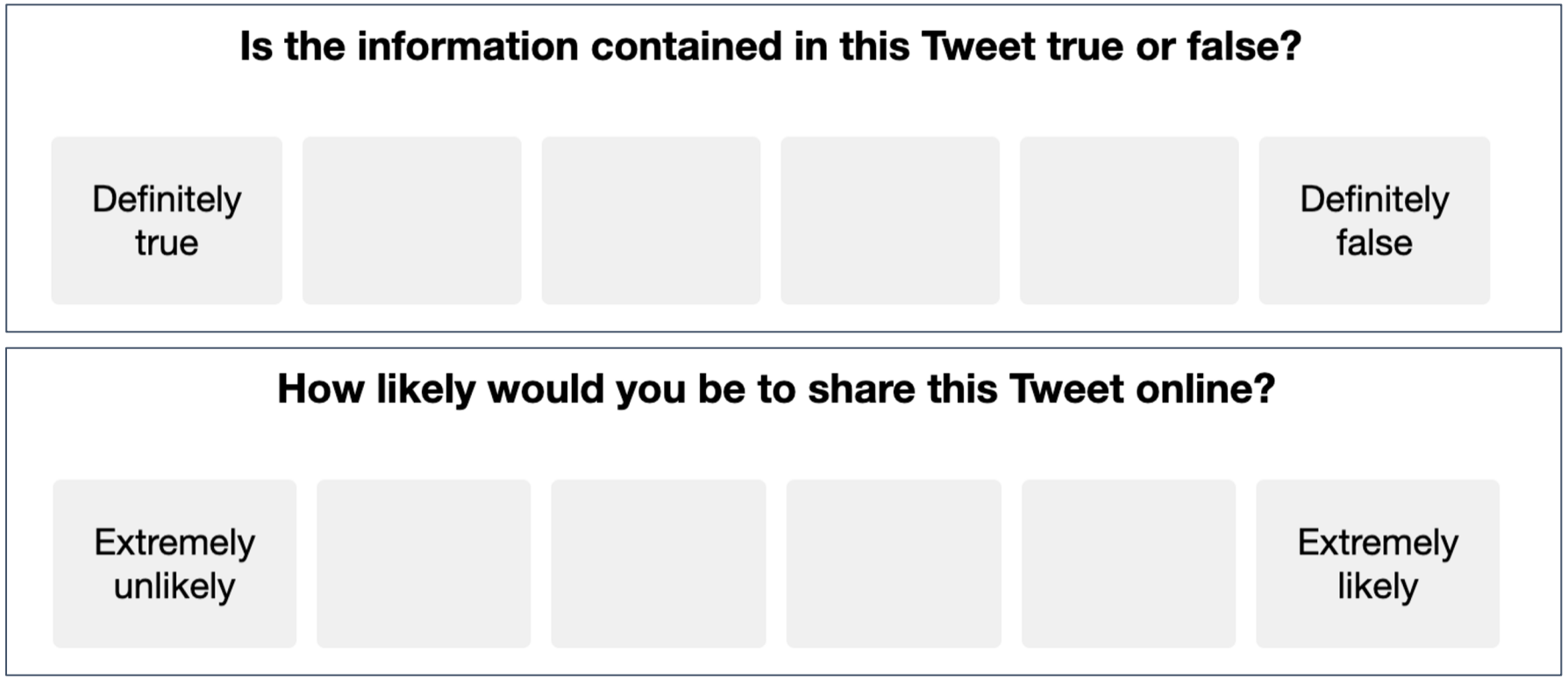
- Truth queries = Questions that appeal to truth or truth criteria
- Can social “truth queries” posted by users reduce the impact of false information posted on online?
- Conducted three initial experiments to test this

# Methods

- Participants
  - Twitter users on Mturk (Exp. 1-2) and Prolific (Exp 3)
  - *N*s = 200, 600, 600, respectively

# Methods

- Viewed a series of Tweets
- Rated truth or sharing for each Tweet



# Methods

- Tweets
  - Key Tweets containing false information
    - 8 Tweets, 6 Tweets, 24 Tweets, Exp. 1-3 respectively
  - 4 fillers containing true information

- Tweet content
    - Myths fact-checked on Snopes (or similar websites)
    - Chosen based on norming to not be obviously true or false

# Methods


Tommy Brennan
@Tommyfbrennan
Good to be reminded of this with the warm weather ahead: A bottle of hand sanitizer can spontaneously combust if left in a hot car.
2:05 PM · Apr 3, 2022 · Twitter for iPhone

# Methods


Amanda Gallagher
@MandyGallagher
One more reason to celebrate on Mother's Day: there's less crime than any other day of the year!
8:22 AM · May 8, 2022 · Twitter for Web App

# Methods

- Tweets appeared in one of three reply conditions (within-subjects):
  - With no reply
  - With a user reply containing a truth query
  - With a neutral reply unrelated to truth (Exp. 2+3)

# Methods

- Condition:
  - **With a user reply containing a truth query**

# Methods

- Condition:
  - **With a user reply containing a truth query**

# Methods

- In Experiments 1 and 2, always paired the same truth query reply with the same Tweet

- In Experiment 3, more systematically tested different truth queries
  - Tweets appeared with different truth queries (between-subjects)

# Methods: Experiment 3 Truth Queries

| **Truth Criteria**<br>(Schwarz & Jalbert, 2020) | **Truth Query** |
|---|---|
| Compatibility:<br>Is it compatible with other things I know? | 1. Does that make sense given everything else you know? |
| Credibility:<br>Does it come from a credible source? | 2. Where did you learn this?<br>3. How do you know that? |
| Consensus:<br>Do other people believe it? | 4. Do other people believe that? |
| Evidence:<br>Is there supporting evidence? | 5. What evidence is there for that?<br>6. Is there proof of that? |
| General appeal to truth | 7. Why would that be the case?<br>8. How do you know this is true? |

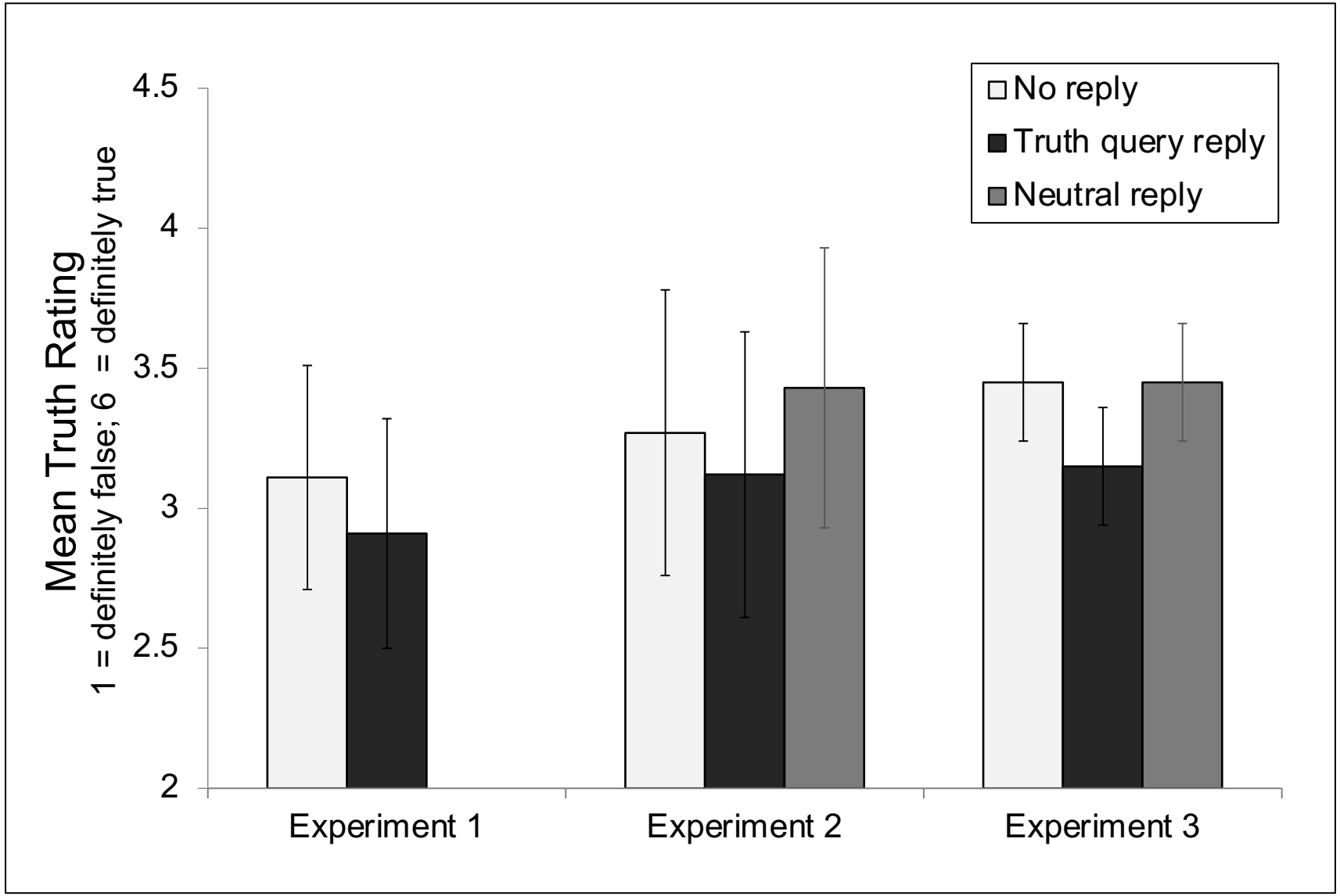
# Methods

- Condition:
  - **With a neutral reply unrelated to truth**

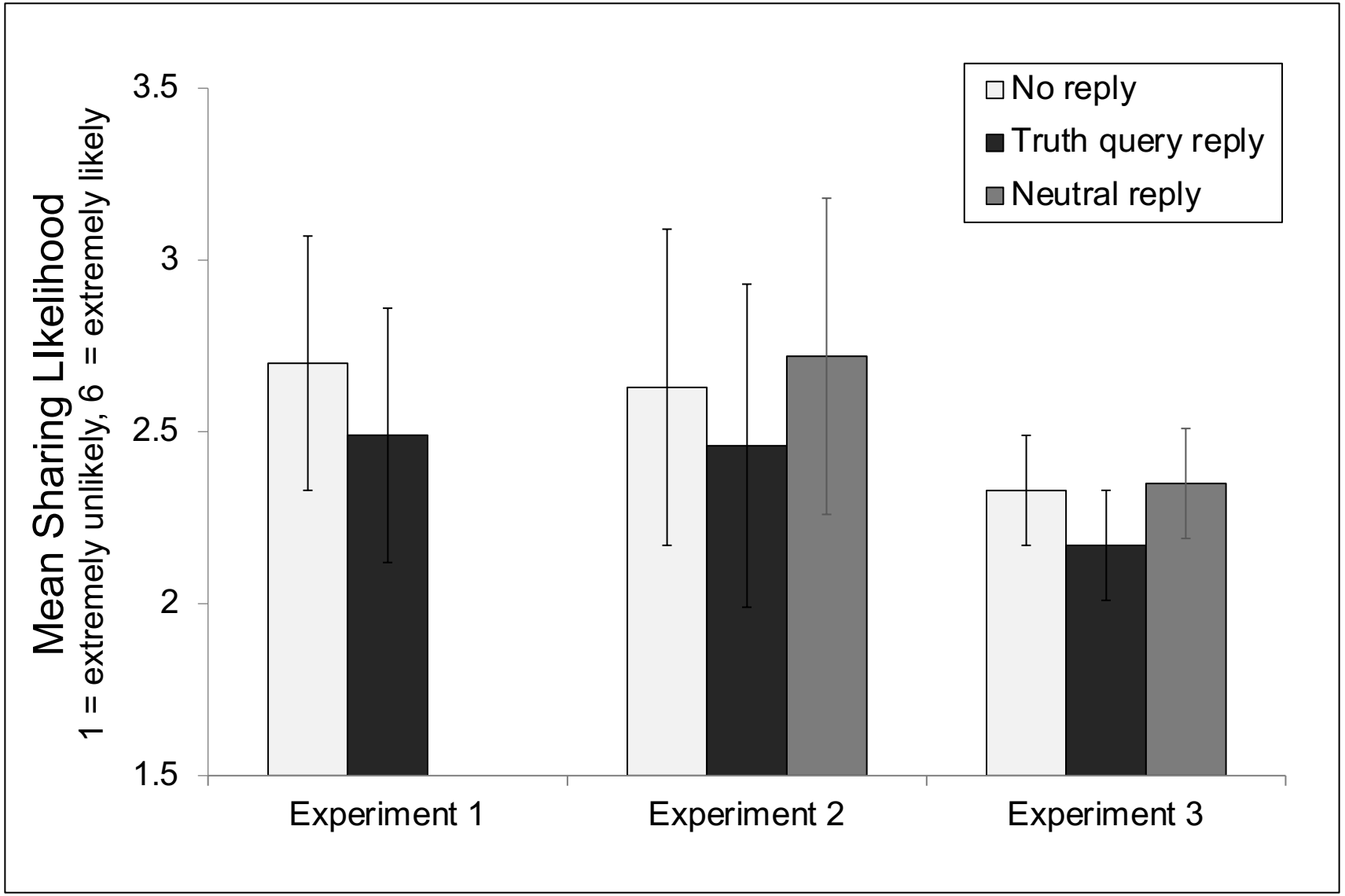
# Methods

- Condition:
  - **With a neutral reply unrelated to truth**

# Truth Ratings

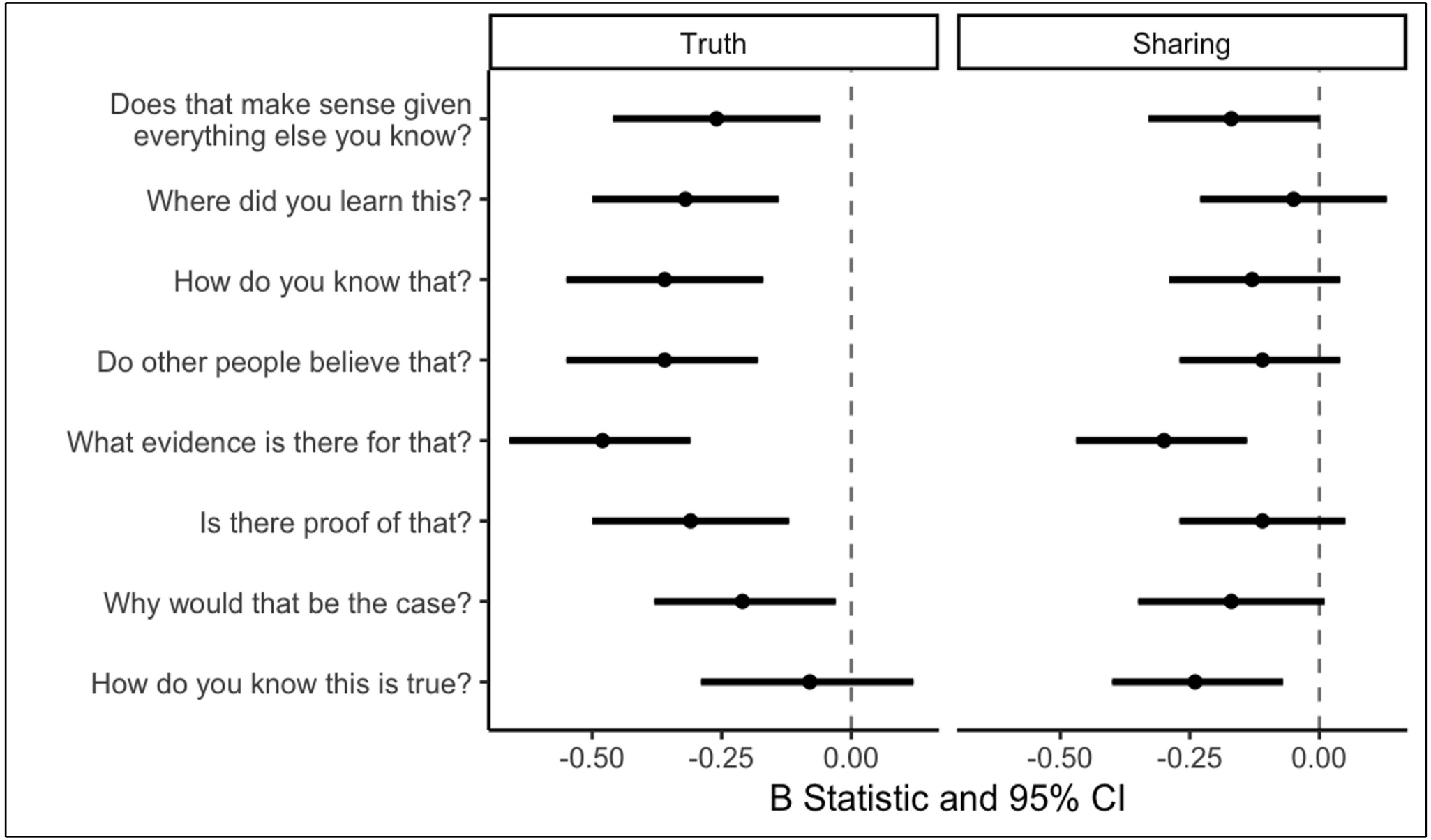


# Sharing Ratings



# Individual Truth Queries (Exp. 3)

Truth Query Reply vs. No Reply



# Discussion

- We find initial evidence for truth queries as flexible, user-driven strategy for addressing online misinformation

- Several potential reasons for this effectiveness
  - Attention to accuracy
  - Communicating a lack of social consensus from peers
  - Social presentation

- A broad range of truth queries appealing to different truth criteria were effective

- Truth queries may be most be the effective for people who are the most susceptible to believing and sharing false information

CENTER FOR AN INFORMED PUBLIC
UNIVERSITY of WASHINGTON

# Discussion

- Still many unknowns about when truth queries are the most likely to be effective

- Other considerations include:
  - Impact on true information
  - The presence of other replies
  - The specific affordances of the social media platform
    - Algorithms that prioritize posts with more engagement
    - Comment order
    - Whether comments appear with shared posts

- Lots of possibilities for follow-ups!

# Thank you!

Email: mjalbert@uw.edu

Morgan Wack
Clemson University

Pragya Arya
University of Southern California

Luke Williams
Pomona College