

# AI Headlines

---

Group 9

# The Paper

- Title: People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation
- Authors: Sacha Altay and Fabrizio Gilardi

# Overview of paper

- 2 studies
  - Study 1: "how labeling headlines as AI-generated influences both the perceived accuracy of the headlines and participants' intention to share them"
  - Study 2: "The mechanisms responsible for these effects"



# Study 1

- Sample size: August 31 2023 - September 4 2023, 1,976 recruited participants from the United States
  - 992 women and 987 men
  - 997 Independents, 498 Republicans, 484 Democrats
  - Median age 40.30
  - Participants were paid for taking part
- Hypothesis: "headlines labeled as AI-generated would be rated as less accurate and receive lower sharing intentions"

# Study 1

## Methodology

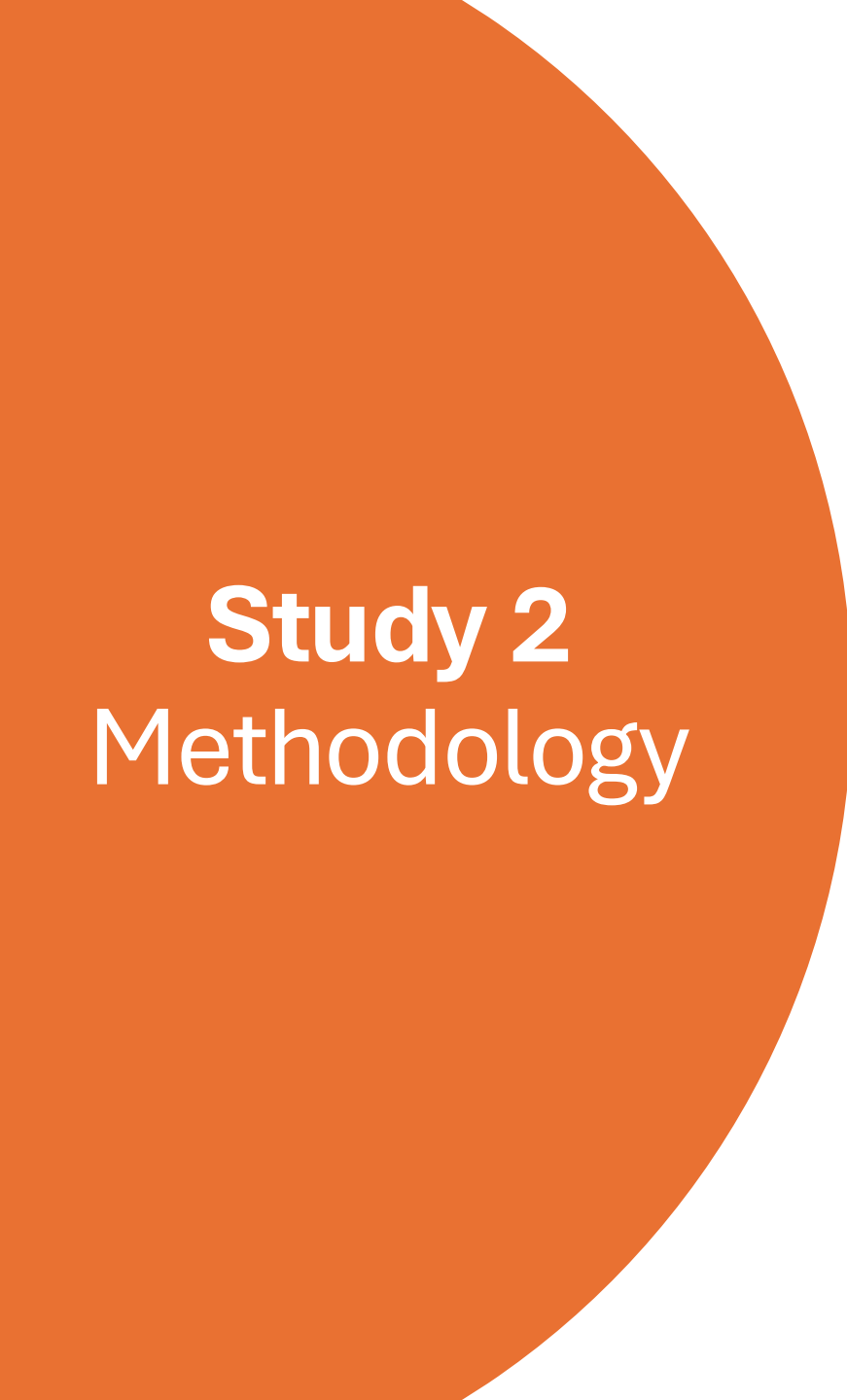
- Each participant saw 16 headlines
- Half reported the trustworthiness of the headline, half reported how willing they would be to share the headline.
- Real headlines used, CHATGPT (version 3.5) to generate fake headlines based off articles, the exact prompt used was “Can you create a headline and a 1-sentence led based on this news article: [copy paste the full text of the article]”



## Study 2

- Study 2 builds off study 1. This time, they try to ask why they got the results they did. "why are people skeptical of headlines labeled as AI-generated?"
- Three new conditions added: Weak, medium, strong



A large orange circle on the left side of the slide, partially cut off by the edge.

## Study 2 Methodology

- March 21st 2023 - March 20th 2024
- 3,003 participants in the United States and United Kingdom
  - 1,498 women and 1,499 men
  - Median age 40.30
  - Equal political leanings
- Participants saw 10 headlines



# Study 1: Control Condition (no labels)

	AI-generated headlines		Human-generated headlines	
	True headlines	False headlines	True headlines	False headlines
Control	No label	No label	No label	No label

- No label AI headlines were compared to no label human headlines
- Human headlines **rated less accurate** than AI
- No difference in the sharing intentions



# Study 1: Labeling AI headlines

	AI-generated headlines	
	True headlines	False headlines
Control	No label	No label
Correct labels	All labeled as AI	All labeled as AI

- Comparing no label AI headlines to labeled AI headlines
- **Labeling as AI-generated** ↓ perceived accuracy & shareability

# Study 1: Labeling AI headlines

- Reduction effect held true in an:
  - **Ideal environment** (all headlines labelled correctly)
  - **Realistic environment** (some mislabeled/labels missing)



# Study 1: Mislabeling Human Headlines

- **Mislabeling as AI-generated** ↓ perceived accuracy & shareability



## Study 2: Similar Results to Study 1

- **Labeling as AI-generated** ↓ perceived accuracy & shareability
- Effect occurred **regardless of headline type**



# Study 2: AI Involvement Conditions

- **Strong condition** (AI chose topic & wrote article)
  - biggest ↓ perceived accuracy & shareability
- **Medium condition** (AI drafted article)
- **Weak condition** (AI made refinements)
  - smaller reductions





# Takeaways

- **Labeling as AI-generated** reduces perceived accuracy & shareability
- Effect holds for **human or AI content, true or false**
- Impact on perceived accuracy greater than impact on sharing



# Takeaways

- Labeling content as **false** has 3x the impact of labeling it as **AI-generated**
- Participants don't equate AI generated with false
- But the **AI generated label increases skepticism** (↓ perceived accuracy)



# Takeaways

- Without definitions, **people assume “AI-generated” = full automation** (AI wrote everything)
- With definitions, stronger AI involvement → lower perceived accuracy & shareability ratings



# Limitations



Focused only  
on news headlines



US/UK audiences



Unclear which AI task  
caused skepticism



Test on harmful content and  
combined labels.

# Future Work



Study sharing reasons



Writing style and topic  
vs. Label



Explore future dimensions



- Emotional reaction
- Engagement behavior




Model choice could  
affect results



Diversity

# Discussion

- Similar background, still varied opinions on labeling
- When a headline is labeled AI-generated, what do you think that means?
- Willing to share news 
  - Privacy
  - Fear of judgment
- Label vs. Content
- Transparency trade-off
- Results may shift over time
- AI-generated label can reinforce AI skepticism

## Broader Context and Conclusion



Why does this research matter outside of academia?



Flagging is increasingly common on social media platforms



Mislabeling erodes trust



Public misunderstanding of "AI-Generated"

# Conclusion



## Education User

Educate users to understand AI labels clearly and make informed judgements.



## Transparency Platform

What does "AI-Generated" mean? Platforms must be transparent about how and why content is flagged.



## Label With Care Policymakers

Does a false positive do more harm than a correct one does good?

# Questions?

A thick, hand-drawn orange line that spans the width of the text above it, positioned below the word "Questions?".

	AI-generated headlines		Human-generated headlines	
	True headlines	False headlines	True headlines	False headlines
Control	No label	No label	No label	No label
Correct labels	All labeled as AI	All labeled as AI	No label	No label
Missing labels	Half labeled as AI	Half labeled as AI	No label	No label
Noisy labels	Half labeled as AI	Half labeled as AI	Half labeled as AI	Half labeled as AI
False labels	No label	All labeled as false	No label	All labeled as false