

When AI Rewrites the News: How Sentiment, Framing, and LLM Disclosure Shape Perceptions

PRERANA KHATIWADA, VARUN PAPPU, BENJAMIN BAGOZZI, & MATTHEW LOUIS MAURIELLO



13. April 2026

Presented by Varun Pappu

SENSIFY LAB

COMPUTER &
INFORMATION SCIENCES



DEPARTMENT OF
POLITICAL SCIENCE & INTERNATIONAL RELATIONS

Content

1 Background and Motivation



2 Research Goals



3 How We Transformed News Articles



4 Results Overview



5 Insights, Limitations, Future Work



Content

1

Background and Motivation



Motivation

Despite the growing adoption of **AI-assisted writing**, relatively little is known about how algorithmic editorial changes, such as **altering sentiment or reframing arguments** affect audience perceptions of fairness, credibility, and trustworthiness

Background

Prior research shows that **media framing and emotional tone** influence how audiences interpret political information, shaping perceptions of bias, fairness, and responsibility in news coverage (*Entman et al. 1993; Lecheler et al. 2013*).

Background

Studies in political communication demonstrate that sentiment and emotional cues such as anger or negativity can amplify framing effects and influence audience engagement, attitudes, and emotional responses to political content

(Soroka et al.2015; Nelson et al.1997)

Background

Computational approaches using **sentiment analysis** and **linguistic modeling** have been widely used to analyze political discourse, audience reactions, and misinformation, though **sentiment alone often fails to capture deeper evaluative meaning or stance** (*Khoo et al.2012; Mannan et al.2023*).

What We Did

We aim to investigate how LLM-driven transformations in news—specifically sentiment, framing, and disclosure of LLM modification shape readers' perceptions of bias, trust, and emotional response.

This involves generating controlled AI-modified news articles, conducting experimental user studies to evaluate audience reactions, and deriving design implications for transparent and trustworthy AI-mediated news systems.



Content

2

Research Goals/Questions



Understand How Sentiment and Framing Shape News Perceptions

- Dataset Curation
- LLM-Based Article Transformation
- Sentiment & Framing Validation

RQ1



How do emotional tone (sentiment) and news framing influence readers' perceptions of bias, trust, and emotional engagement with news?

Isolate the Effect of Framing on News Perceptions

- Controlled Experimental Design
- Participant Evaluation Surveys
- Comparative Analysis of Framing Effects

RQ2



When sentiment is held constant, what role does framing play in shaping perceptions of bias, trust, and balance?

Examine How LLM Disclosure Shapes Reader Perceptions

- LLM Disclosure Conditions
- Perception and Trust Measurement
- Statistical and Mixed-Method Analysis

RQ3



Extending beyond textual features, how does disclosure that an article was modified by an LLM affect reader perceptions of bias, trust, and engagement?

Contributions to the Intervention Design

Guiding Questions

If framing drives misinterpretation → surface alternative frames or highlight one-sided narratives.

1. What if framing matters more than sentiment?
2. How should interventions adapt when neutral tone still carries bias?
3. When should explanations vs. contrastive perspectives be shown?

If neutral articles still bias readers → provide missing perspectives or contrastive rationale cues.

Content

3

How We Transformed Articles





**ORIGINAL
NEWS ARTICLE**

Validation Rubric

Human Review

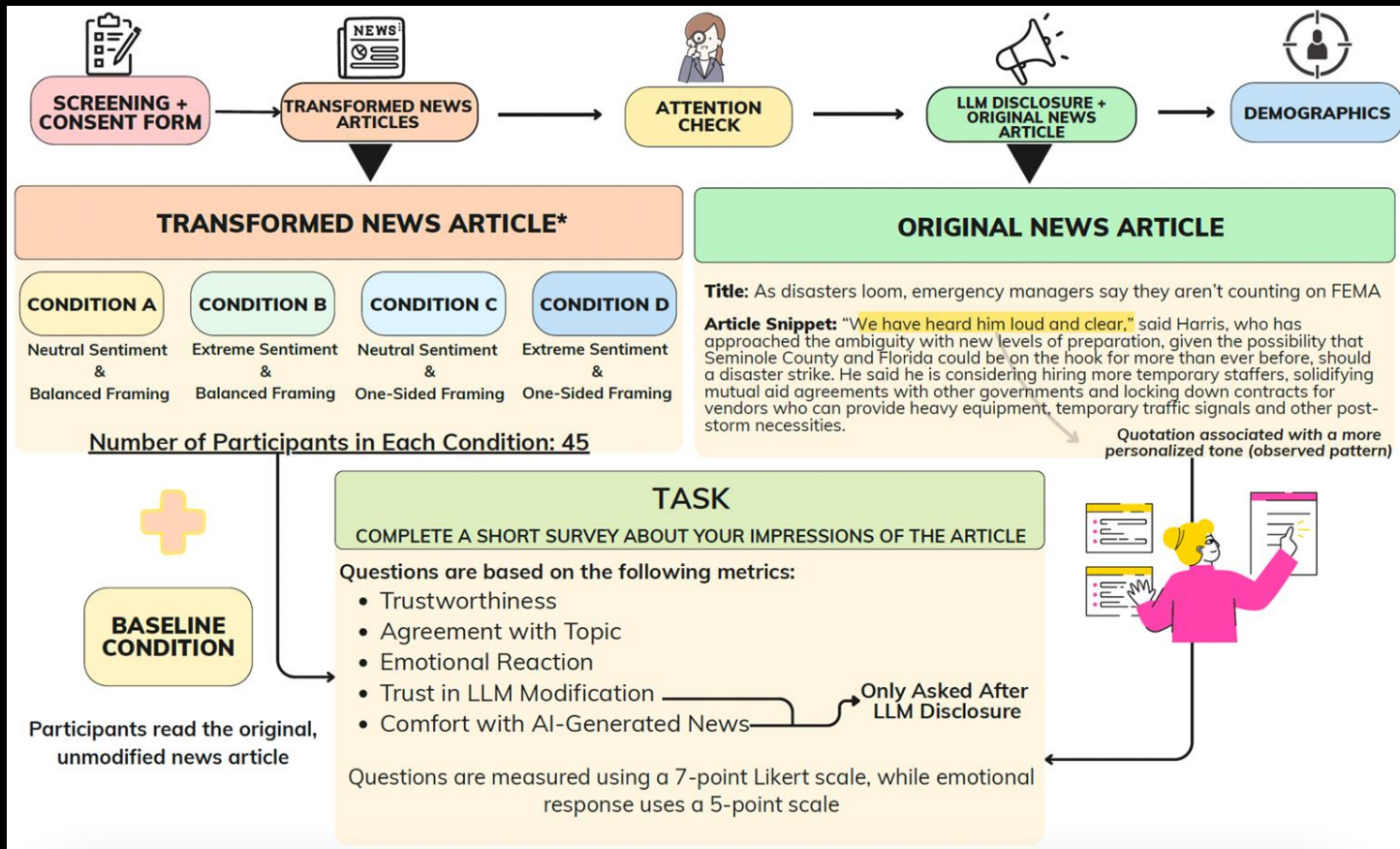
- **3 Coders:** Political Science Experts (Professors & Graduate Students)
- Evaluated **3 original + 4 transformed** news articles (Best article chosen for experimental use)
- **Criteria:** fluency, faithfulness, tone accuracy, framing clarity, bias, conciseness, overall effectiveness

Validation

- **VADER sentiment analysis**
 - confirmed tone alignment
- **Stance entropy**
 - measured framing diversity (higher = balanced)
- **Results consistent with human ratings**

Human results are included in the full paper—please refer to the full paper for detailed results.

How We Transformed Articles



N = 225 total

- 180 in 2x2 design
- 45 baseline
 - original article

Power analysis (G*Power)

- (G*Power): $f = 0.25$, $\alpha = .05$, power = .80
- 180 needed (45 per condition)

Recruitment and Screening

- U.S. MTurk, August 2025
- Weekday business-hour batches
- Qualifications:
 - Age ≥ 18 , regular online news reader, English proficiency

Content

4

Results
Overview



**What Actually Happens When AI
Rewrites the News?**

Key Findings

Perceptions were surprisingly stable

Disclosure slightly reduced trust

Sentiment & framing had minimal impact before disclosure

Extreme sentiment increased perceived bias

Framing had little effect

Overall, disclosure did not dramatically change perceptions

Key Findings

Perceptions were surprisingly stable

Disclosure slightly reduced trust

Sentiment & framing had minimal impact before disclosure

Extreme sentiment increased perceived bias

Framing had little effect

Overall, disclosure did not dramatically change perceptions

Key Findings

Perceptions were surprisingly stable

Disclosure slightly reduced trust

Sentiment & framing had minimal impact before disclosure

Extreme sentiment increased perceived bias

Framing had little effect

Overall, disclosure did not dramatically change perceptions

Key Findings

Perceptions were surprisingly stable

Disclosure slightly reduced trust

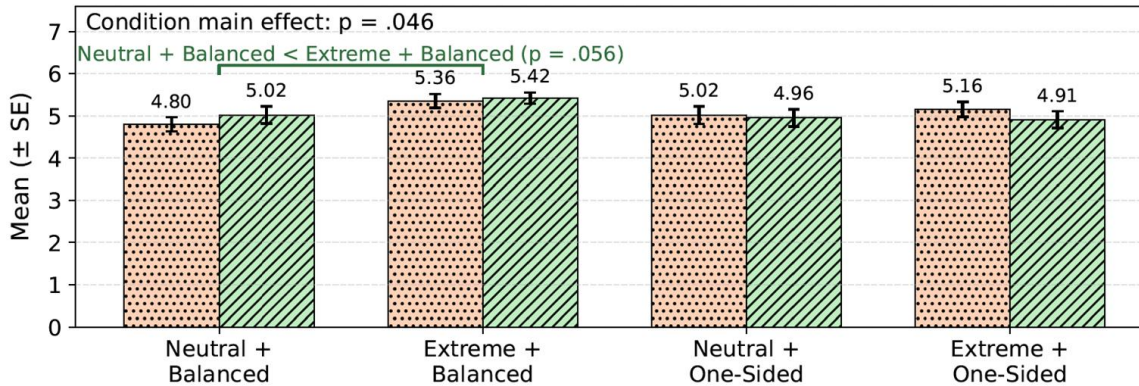
Sentiment & framing had minimal impact before disclosure

Extreme sentiment increased perceived bias

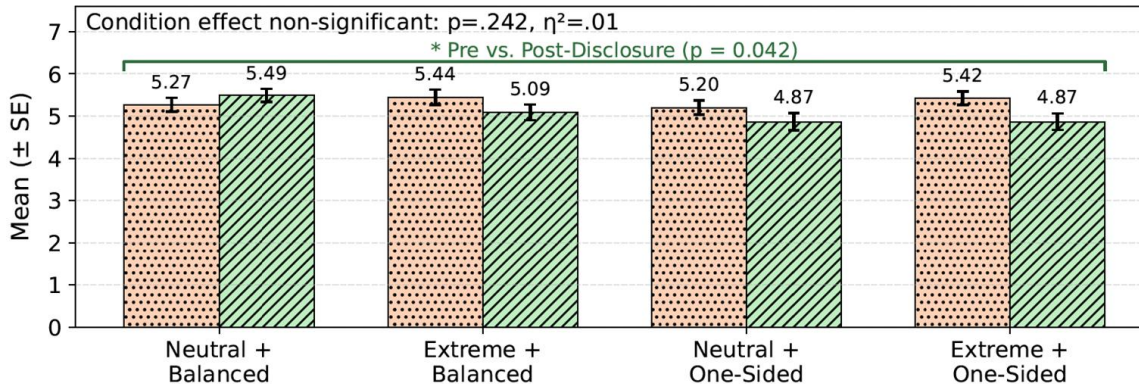
Framing had little effect

Overall, disclosure did not dramatically change perceptions

Perceptions by Condition and Evaluation Phase Perceived Bias of News Article



Perceived Trustworthiness of News Article



Emotional Reaction Results

Extreme sentiment amplified negative emotions

Disgust (+0.74) and anxiety (+0.69) increased compared to original article, suggesting that emotionally intense language heightens negative feelings

Positive emotions remained stable

Happiness changed little across conditions, indicating that **negative emotions were more sensitive to tone changes**

AI disclosure did not strongly change emotional reactions

Most emotions remained similar before and after disclosure

Content style mattered more than disclosure

extreme sentiment with balanced framing triggered the strongest negative emotional responses, while extreme one-sided framing reduced positive reactions.

Emotional Reaction Results

Extreme sentiment amplified negative emotions

Disgust (+0.74) and anxiety (+0.69) increased compared to original article, suggesting that emotionally intense language heightens negative feelings

AI disclosure did not strongly change emotional reactions

Most emotions remained similar before and after disclosure

Positive emotions remained stable

Happiness changed little across conditions, indicating that negative emotions were more sensitive to tone changes

Content style mattered more than disclosure

extreme sentiment with balanced framing triggered the strongest negative emotional responses, while extreme one-sided framing reduced positive reactions.

Emotional Reaction Results

Extreme sentiment amplified negative emotions

Disgust (+0.74) and anxiety (+0.69) increased compared to original article, suggesting that emotionally intense language heightens negative feelings

AI disclosure did not strongly change emotional reactions

Most emotions remained similar before and after disclosure

Positive emotions remained stable

Happiness changed little across conditions, indicating that negative emotions were more sensitive to tone changes

Content style mattered more than disclosure

extreme sentiment with balanced framing triggered the strongest negative emotional responses, while extreme one-sided framing reduced positive reactions.

Perceived Article Modification (Post-Survey)

**Participants generally believed
AI modifications altered the
article
(M = 4.53)**

**Across all conditions, participants
suspected AI edits could
influence readers' opinions**

**Articles were perceived as
moderately partisan overall
(M = 4.73)**

**Omission of important
information was the strongest
concern
(M = 4.67)**

**Subtle modifications sometimes
raised more suspicion**

Content

5

Insights, Limitations, Future
Work



Design Implications

Contextualized AI Disclosure

Move beyond simple “**AI-generated**” labels by explaining **what was modified and whether human editors verified the content**, helping maintain transparency without unnecessarily reducing trust.

Emotion-Aware Highlighting

Use interface cues to **highlight emotionally charged or one-sided language**, helping readers recognize persuasive tone and engage more critically with the content.

Risk-Aware Moderation

AI systems should **detect and soften extreme or inflammatory language while preserving facts**, with stronger transparency signals in highly polarized or sensitive topics.

Distortion and Intent Auditing

Provide newsroom tools that **flag potential omissions, exaggerations, or agenda-driven language** in AI-modified articles, allowing editors to review and maintain credibility before publication.

Limitations

Sample bias: MTurk participants were relatively homogeneous, which may limit generalizability to broader populations.

Limited context: The study focused only on sentiment and framing, while real-world news involves other factors like source credibility and media context.

Single article: Using one news article may limit generalization across different topics.

Disclosure scope: Only one form of AI disclosure was tested, leaving other transparency formats unexplored.

Future Work

Broader datasets: Future work can test multiple news topics and larger datasets to improve generalizability.

Real-world interventions: Future systems could use AI to flag extreme sentiment or one-sided framing and provide multi-perspective summaries to support readers.

Acknowledgements

Team Members:



Let's Chat

Prerana Khatiwada
preranak@udel.edu

Varun Pappu
varunp@udel.edu

We thank our volunteer coders from the Department of Political Science and International Relations for their valuable time. We also thank the UD Undergraduate Research Program for supporting Varun Pappu's participation via the Summer Scholars program.



SENSIFY LAB

COMPUTER &
INFORMATION SCIENCES



DEPARTMENT OF
POLITICAL SCIENCE & INTERNATIONAL RELATIONS



Scan For Full

Paper