

# **From Articles to Posts: Sentiment and News Frame Changes in the Transition to Instagram**

**Yeongseo (Alex) Chae**

New York University Abu Dhabi, Class of 2024

Advisor: Professor Minsu Park

Social Research and Public Policy

Division of Social Science

New York University Abu Dhabi

April 26, 2024

## Acknowledgements

I am greatly indebted to my Capstone advisor Professor Minsu Park. You taught me not only how to push forward with a project with perseverance and passion, but also how to think and grow as a researcher. In moments when I was lost, conversations with you were eye-opening for me and truly shaped my academic and personal journey. Thank you for offering constructive feedback and believing in me - I am so grateful to be your mentee. I sincerely thank Professor Kinga Makovi for her insights and advice. Your kind words and invaluable feedback was crucial in finalizing this Capstone. I am only one among many students who truly admire your passion and work ethic. Taking your class and reaching out to you were the best decisions I have taken throughout university. Thank you for your caring and kind words and I hope to see you in-person again in the near future. I would also like to thank my Capstone conveners Professor Stéphane Helleringer and Instructor Marouane Laouina for providing valuable advice and guidance through my capstone journey. Lastly, I extend my gratitude to my family and friends for all boba, hotpot and matcha latte memories. Special thanks to my sister El, who always brings me laughter. Thank you for making my four years at New York University Abu Dhabi unforgettable.

## Abstract

As social media becomes an increasingly popular news source, particularly among younger audiences, understanding how news content is adapted for platforms such as Instagram is critical, but understudied. This study analyzes a large-scale dataset comprising Instagram posts and their original articles from six leading news organizations. We observe that while headlines are subtly altered without semantic changes, the sentiment of news content shows significant neutralization, countering the common perception of social media as a breeding ground for emotional and polarizing content. An analysis of news frames using a large language model reveals that this neutralization stems not merely from text summarization, but likely from a systematic removal of certain news frames. Interestingly, articles with strong negative sentiments tend to retain their tone more robustly than those with positive sentiments. Furthermore, content related to human interest and responsibility frames is more likely to be preserved in the transition. These findings suggest a potential editorial bias or selection strategy in the way news content is modified during the transition from article to post, which may influence public perception of news and lead to unintended effects on user engagement and discourse.

*Keywords* social media, Instagram, news frames, sentiment analysis, platformization

# **Contents**

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>2</b>	<b>Background</b>	<b>6</b>
2.1	Platform-parting and platformization . . . . .	6
2.2	News on social media platforms . . . . .	7
2.3	News on Instagram . . . . .	9
2.4	News framing . . . . .	9
<b>3</b>	<b>Research Questions</b>	<b>11</b>
<b>4</b>	<b>Data Collection</b>	<b>11</b>
4.1	Selection of news organizations . . . . .	11
4.2	Collecting data . . . . .	11
4.3	Checking robustness of data . . . . .	12
4.4	Summary of the data . . . . .	12
<b>5</b>	<b>Methodology</b>	<b>13</b>
5.1	Fuzzy string matching . . . . .	13
5.2	Semantic similarity . . . . .	13
5.3	Sentiment analysis . . . . .	13
5.4	News framing analysis . . . . .	14
<b>6</b>	<b>Results</b>	<b>14</b>
6.1	Headline modification . . . . .	14
6.2	Content modification . . . . .	18
<b>7</b>	<b>Conclusion</b>	<b>25</b>
7.1	Main findings . . . . .	25
7.2	Discussion . . . . .	26
<b>References</b>		<b>28</b>
<b>Appendix</b>		<b>33</b>
A.	Definition and examples of five news frames used in the study . . . . .	33

B. Robustness check of dataset . . . . .	34
C. Questions and prompts for GPT 3.5 Turbo used in framing analysis . . . . .	37
D. Validation of GPT 3.5 Turbo Performance . . . . .	39
E. Headline sentiment analysis by news media . . . . .	40
F. Conversion rates in sentiment changes from articles to posts . . . . .	41
G. Content sentiment analysis by news media . . . . .	42
H. Distribution of sentiment of article body text and post captions, separated by presence of headlines in post . . . . .	43
I. Presence of news frames in articles and posts within each political leaning . . . . .	44

## 1 Introduction

As social media gains popularity, many people choose to read news on social media, instead of traditional platforms such as television and newspapers (Liedke & Wang, 2023a; Newman et al., 2021; Vorhaus, 2020). This change in news reading behavior is especially pronounced among the younger generation. In fact, 42% of adults between 18 and 29 in the United States report to often consume news from social media and 63% of news consumption among 16 to 24-year olds in the United Kingdom took place on social media (Shearer, 2021; Fletcher, 2023; Ofcom, 2023). Thus, for a full understanding of news consumption habits, we need to examine social media as a news platform. Meanwhile, literature on social media news is not comprehensive, as social media is relatively new compared to traditional news media platforms and is constantly changing with introduction of new platforms and features.

News organizations manage social media accounts that attract millions of followers to publish posts with short texts and images about a news item. The majority of these posts are modified from an article. In this process, news organizations may influence what news stories reach the audience by systematically selecting articles to be published on social media (Lischka, 2021; Tsuriel, Dvir Gvirsman, Ziv, Afriat-Aviv, & Ivan, 2021). Then, social media news readers interact with a pre-selected pool of news that has been catered to the platform. Furthermore, there is an incentive for organizations to employ “clickbait” strategies on social media. Unlike news articles, social media news face intense competition not only against other news content, but also content from a user’s social circle, online interest groups and other forms of entertainment, all of which aim to catch the user’s attention (Webster, 2014). This discrepancy in the environment of news production necessitates that the social media news content differs from news articles. With an increased focus on other modalities such as images and audio, social media offers limited space for text. To hook the audience within a short text, organizations may be motivated to use clickbait strategies such as emotional and personal stories that are associated with high audience engagement. Therefore, how news is presented on social media is distinct from other news reading platforms.

This project aims to map social media news production by comprehensively investigating the differences between how prominent news organizations present news on Instagram and on articles. In particular, we use a dataset of news articles and their adaptations into Instagram posts from six highly followed news accounts. We find that while headlines do not undergo significant changes in semantic meaning nor in sentiment, news content is drastically modified as sentiment is neutralized and less news frames are used to report a story. We also present evidence that clickbait strategies may be used in the modification process.

## 2 Background

As social media news gains traction, there have been efforts to understand social media news production. We present a review of studies on the platformization of news, how news content is modified from online articles to social media posts, specifically on Instagram, and a discussion on news framing. We highlight the gaps in literature that the study aims to fill.

### 2.1 Platform-parting and platformization

News organizations today are not confined to a single medium for publishing their stories. The proliferation of digital platforms has significantly expanded their publishing capabilities, allowing the distribution of content across a wide array of channels. Hanusch (2017) refers to this practice as **platform-parting** — republishing a news story across various platforms to maximize audience reach. During platform-parting, a news story is modified to fit the designs and demands

of the platform. We adopt the discussion of **platformization** by Nieborg and Poell (2018) to explore this process. Platformization is when business models, power structure and infrastructure of digital platforms motivate a change in how cultural artifacts are produced and circulated (Nieborg & Poell, 2018). We explore platformization specifically in how news is impacted when it is published on social media.

A discussion of social media begins with acknowledging that social media is a general-purpose platform. Unlike print newspapers or news websites that are curated for distributing news, social media platforms are designed for “user-generated content,” which includes personal stories, music, and games (Carr & Hayes, 2015; Howard & Parks, 2012). Thus, the relative importance of news organizations have greatly shrunk. Before social media emerged as a news source, news organizations were the sole mediator between news producers and readers and controlled what news is shared (Nieborg & Poell, 2018). Contrastingly, social media operates in a “multi-sided market” where social media companies, advertisers, users, and content producers interact (Nieborg & Poell, 2018). Here, the role of news organizations is greatly diminished to one among many content producers. This change in the power dynamics between platform and news organizations compels news organizations to develop new strategies to profit and attract audiences.

A similar reduction of autonomy is visible in the news content itself. When news organizations produce content for social media, they are constantly under the scrutiny of two features of social media platforms: algorithmic content curation and content moderation (Nieborg & Poell, 2018). Algorithms provide features such as likes, shares and views, which are used to rank popularity of content. Based on this metric, algorithms recommend content to users. Meanwhile, through interacting with the platform, users can signal to the algorithm whether they enjoy the recommended content or not, thereby influencing the popularity of content (Dijck & Poell, 2013). In addition, platforms employ content moderation that detects and removes damaging content, often due to pressures from the government and the public (Morrow, Swire-Thompson, Polny, Kopec, & Wihbey, 2022). However, this moderation generally happens under “nebulous rules and procedural opacity” (Roberts, 2018), making it difficult for content producers and users to anticipate what content will be flagged (Lischka, 2021). Thus, news organizations are under pressure to create content that will gain popularity in ranking algorithms, while avoiding content moderation. Consequently, they need to modify news stories to cater to the requirements of social media. For such reasons, during platform-parting of online news articles to social media posts, news is platformized with numerous changes in content.

## 2.2 News on social media platforms

Owing to its increasing prominence among news readers, there have been many attempts to explore news production in social media. The first branch of studies focuses on the systematic selection of stories by news organizations. News organizations have an incentive to choose the kinds of stories that they expect to thrive on social media, which may not necessarily reflect the diversity of news items as news websites. However, studies find mixed results. Lischka and Werning (2017) and Steiner (2020) found that Facebook pages of German news outlets have a significantly greater proportion of light news stories on entertainment, celebrity and lifestyle, compared to newspapers and website teasers (as cited by (Lamot, 2022)). Through interviews with Finnish and Swiss journalists, Lischka (2021) finds that journalists prefer to publish stories with emotional and surprising elements on social media. However, Hase, Boczek, and Scharkow (2023) finds that news topics on Facebook and Instagram are not significantly different from that of German news websites. Opinion pieces and personalized articles appeared more frequently only on Twitter (Hase et al., 2023). While we do not focus exclusively on how articles are selected to be published on social media, we do find a discrepancy between prevalent news frames on Instagram and that of online news articles found in interview-based previous studies, which hints

at a preference for less polarizing content on Instagram.

Another branch of literature that is directly related to this study compares social media news to news on other platforms. Hase et al. (2023) compiled a dataset of articles from five German news outlets and posts on their Facebook, Instagram, Twitter, and TikTok pages. They found that headlines on posts were not written in more engaging language than articles and contained less personalization elements. Some studies use a dataset of each article and their direct modifications into social media posts. Lamot, Kreutz, and Opgenhaffen (2022) found that approximately half of headlines are rewritten on Facebook. Headlines that already contain question words and emotion words were more likely to be chosen to be posted on Facebook. Park, Kwak, An, and Chawla (2021) conducted a similar analysis on Twitter and found that modification of headlines is an editorial choice with no consistent pattern. Verstappen and Opgenhaffen (2023) focused exclusively on how science-related news are modified to Facebook and Instagram posts and show that 79% and 59% of headlines remained unchanged on Facebook and Instagram posts, respectively. Also, the authors find that 48% and 92% of body texts on Facebook and Instagram posts, respectively, are different from article body texts, but do not present more in-depth analysis on the discrepancies.

We extend this literature in several ways. The dataset compiled for this study consists of pairs of an article and its adaptation into an Instagram post, using links provided by the news organizations themselves. Previous studies such as Lamot et al. (2022) and Verstappen and Opgenhaffen (2023) use CrowdTangle, a database offered by Meta to access social media content, to collect articles and social media posts. However, CrowdTangle does not differentiate between posts that are explicitly determined by news organizations to lead readers to original articles and those that are meant to encourage on-platform activities. Our dataset is compiled using webpages by each news organization to provide links to original articles for most, but not all, of their posts. As Instagram discourages hyperlinks to lead users off-platform, (Hase et al., 2023; Hermida & Mellado, 2020), it is important to confirm any differences between posts that news organizations provide an external link to and those where they do not. Hence, we first identify characteristics of posts with links to articles, then confirm that the findings are consistent in other posts without links to articles (see Appendix B).

Furthermore, we extend the scope of analysis to investigate how body texts of a news piece is modified. Past studies have focused their in-depth analysis on the portion of news headlines that were modified on social media (Lamot et al., 2022; Verstappen & Opgenhaffen, 2023), while the other portion of news whose headlines were kept has been left largely unstudied. We attempt to expand this literature by investigating potential modifications in the news body text, along with headlines. In addition, while studies such as Guo and Sun (2023) attempted to analyze the captions of Instagram posts, they were based on small-scale data on geographically and linguistically specific regions. These studies often focus on the linguistic features of texts such as use of punctuation and emotive words. We attempt to produce more generalizable results by using a large dataset spanning three years from news organizations with global reach. Also, we not only analyze how a news story is written via sentiment analysis, but also what is written, through framing analysis.

Lastly, there is a branch in literature focusing on the aggregated audience engagement of social media posts. This is highly relevant to “clickbait” strategies to increase the popularity of news content (Dijck & Poell, 2013). As real-time collection of user activity data is readily available nowadays, journalists are sensitive to note content that generates positive audience responses (Lischka, 2021). Then, news organizations have an incentive to publish these kinds of content that are validated by data to perform well on social media. Scholars have used metrics such

as number of likes, comments and shares to find strategies that tend to increase engagement. Piotrkowicz, Dimitrova, Otterbacher, and Markert (2017) demonstrate that sentiment-charged and biased headlines have higher popularity on Facebook and Twitter. On Instagram, Guo and Sun (2023) finds that personal stories may increase the engagement rate, while using Instagram features such as videos, hashtags and tagging tend to decrease audience engagement. Overall, these strategies are in line with clickbait strategies in news articles, such as the use of emotionally charged words showing outrage and disapproval (Robertson et al., 2023). Following these findings that demonstrate the importance of text sentiment, we also start the analysis with the headline and content sentiment. We extend this literature by showing an overall neutralization of sentiment with a skew towards negative sentiments as news content is transferred to social media.

### 2.3 News on Instagram

While different social media platforms share commonalities, they are “not homogenous” (Hermida & Mellado, 2020). The specifics of how each one recommends and moderates content may be vastly different (Dvir-Gvirsman & Tsuriel, 2022; Hanusch, 2017; Liedke & Wang, 2023a; Mottola, 2020). Therefore, as Hanusch (2017) finds, news organizations decide on news production based on the characteristics of each platform. Therefore, findings on new production on a particular platform cannot be generalized to all social media platforms. In platform-parting literature, many studies examine Facebook, Twitter or Reddit. One major reason for this may be the availability of an Application Programming Interface (API) for researchers to access large amounts of data. However, increasingly, people consume news through Instagram and TikTok, calling for a need to examine how news organizations operate in these relatively new, visual-focused platforms. In fact, the number of people that regularly consume news on Facebook, X (Twitter), and Reddit have fallen since 2020, while that of Instagram and Tiktok are rising (Liedke & Wang, 2023b; Shearer, 2021; Vorhaus, 2020). Therefore, we focus our analysis of platform-parting of news specifically on Instagram.

Instagram engages in “brand-building news” which present visual content to audiences with a focus on keeping the audiences engaged in content within the platform (Hase et al., 2023). Hence, news organizations make an explicit effort to reiterate their identity through including logos in images and using hashtags with the name of the organization. Similarly, studies have found that Instagram prefers on-platform engagement and discourage hyperlinks to other websites directly on posts (Hase et al., 2023; Hermida & Mellado, 2020). In fact, posts by news organizations examined in this study do not contain a hyperlink to the original article. Rather, to view the article, users need to visit the news account’s profile and find the link in a separate webpage. Compared to Facebook posts or Twitter status messages in which it is not uncommon to find hyperlinks leading to a website out of the platforms, Instagram posts encourage users to consume information on-platform. This further emphasizes the importance of understanding how news is presented on Instagram. For news readers on Instagram, posts are the primary, and likely the only, news that they consume. Therefore, it is imperative to establish an understanding of how Instagram news content differs from their original articles and what impact it has on readers.

### 2.4 News framing

A branch of news literature that has attracted much attention is framing. Frames are “aspects of a perceived reality” that are made “salient in communicating text” (Entman, 1993). Media organizations rely on frames to interpret information and convey it to the audience (Pan & Kosicki, 1993; Semetko & Valkenburg, 2000). In news, frames are presented alongside the core, factual information to make certain aspects of a story memorable to the reader, while obscuring others (Entman, 1993).

Druckman (2001) presents two types of news framing. **Equivalency** framing focuses on the impact of logically same, but differently phrased texts. Two equivalency frames convey the same information in different ways, generating different influences on audience understanding. Meanwhile, **emphasis** framing focuses on which aspects, among many potential relevant considerations, were emphasized by the text. Each emphasis frame supplies additional information, along with the core information of the news. This paper follows the latter approach. The data examined here consists of Instagram posts, which are shorter adaptations of a news article. Thus, we cannot reasonably assume that the two texts are logically equivalent — in fact, we demonstrate through framing analysis that they are not equivalent — and hence cannot apply equivalency framing.

We adopt the five emphasis frames described by Semetko and Valkenburg (2000): conflict, human interest, responsibility, morality, and economic consequences frames, which have been used in multiple studies on news framing (An & Gower, 2009; Gronemeyer & Porath, 2017; Valenzuela, Piña, & Ramírez, 2017). Contrasting to **issue-specific** frames which pertain to specific issues, these five frames are **generic** frames, which allow “comparison between frames, topics, … and framing practices” of different news content (De Vreese, Peter, & Semetko, 2010). Hence, they are applicable to this paper which does not specify issues of focus. Here, we briefly introduce the five frames. Examples of each frame can be found in Appendix A.

**Human interest** frame is when a “human face or an emotional angle” is put to to “personalize the news” (Semetko & Valkenburg, 2000). Studies found that human interest frames are often used in stories on minorities and predict higher emotions of blame among readers (Cho & Gower, 2006; Figenschou & Thorbjørnsrud, 2015; Hong, 2013). **Responsibility** frame attributes responsibility to an individual, group or government (Semetko & Valkenburg, 2000). Studies have found that television news on politics tend to attribute responsibility to the private, rather than the public actors (Iyengar, 1996) and that media may be incentivized to attribute responsibility to individuals, rather than governments, to attract more attention (Kim, 2015). **Economic consequences** frame refers to emphasizing “the consequences [a story] will have economically” (Semetko & Valkenburg, 2000). This frame renders an issue more relevant to the audience (De Vreese et al., 2010; Valenzuela et al., 2017). **Conflict** frame is when disagreement between individuals, groups or institutions is emphasized. It is one of the most frequently used frames in United States news, especially among political and economic news (Valenzuela et al., 2017). Bartholomé, Lecheler, and de Vreese (2015) find that journalists use exaggerating language and emphasize the possible consequences of a conflict to build this frame. Lastly, the **morality** frame is concerned with references to religious tenets and moral responsibilities (Semetko & Valkenburg, 2000) with discussions of what is socially acceptable. While Wasike (2013) finds that Twitter status posts use the morality frame least frequently, when it is used, this frame is heavily used by the audience to make judgments (Druckman, 2001; Semetko & Valkenburg, 2000).

We adopt the methodology of Semetko and Valkenburg (2000) to identify these five frames within a news text. The authors developed a set of 20 questions about a news text for a human coder to answer (see Appendix C). Each question tests whether one of the five frames are present in the text. There are three to five questions testing each frame. While the authors asked human coders to answer the questions, we take an automated approach by using a large language model to generate answers, proposing a more time-efficient and less labor-intensive method of news framing analysis.

### 3 Research Questions

The goal of this research is to understand platform-parting strategies of news organizations in modifying a news article into an Instagram post. By constructing a dataset of online news articles from six news outlets and their adaptations into Instagram posts, we aim to investigate how news headlines and content change according to the platform. Our research questions are:

- **RQ1** How do headlines of news articles change when an online news article is modified into an Instagram post?
  - How does the semantic meaning of headlines change in modification?
  - How does the sentiment of headlines change in modification?
- **RQ2** How does the content of news articles change when an online news article is modified into an Instagram post?
  - How does the sentiment of content change in modification?
  - How do news frames change in modification?

### 4 Data Collection

#### 4.1 Selection of news organizations

Six news organizations were chosen according to the following criteria.

1. Articles are published in English.
2. The news organization has an official Instagram account.
3. Official news website allows automated web scraping of articles for academic purposes.
4. The news organization publishes articles across different topics, rather than focusing on one topic (e.g. entertainment news, financial news).
5. Instagram accounts of the news organizations have a separate webpage that links each post to an original article in the website.

Two organizations that are politically right-wing, two organizations that are politically left-wing, and two organizations that are politically centered with the highest number of Instagram followers as of October, 2023 were chosen. The political bias of each organization was taken from a variety of trustworthy sources (AdFontesMedia, n.d.; AllSides, 2019). NBC News and the New York Times were chosen from the left, BBC News and Reuters were chosen from the center, and Fox News and the Washington Times were chosen from the right. While some news organizations have multiple official Instagram accounts that target specific countries or language groups, we choose one account per organization that runs in English without targeting any specific geographic region. The Instagram handles used are @nbcnews, @nytimes, @bbcnews, @reuters, @washtimes, and @foxnews.

#### 4.2 Collecting data

Instagram posts were downloaded using the Instaloader module in Python. We downloaded images, videos, and captions of all posts of each Instagram handle that were available during data collection between November 2023 and January 2024. According to the native creative style of Instagram, accounts embed news headlines in the first images of each post. Due to difficulties in automated processes of extracting texts from post images, post headlines were extracted via

manual inspection of two professional transcriptionists from Upwork. The transcriptionists were given all images and were asked to transcribe texts that appear on the images. The transcriptionists both had more than 97% job success rate on Upwork. For further quality control, they were first asked to extract texts from 500 sample images. A random sample of 100 images were manually checked and accuracy was 100% for both transcriptionists.

Many, but not all, of Instagram posts are linked to an article. To manage the link between each post and article, each account has a *link in bio* page on its Instagram profile. This page looks similar to each account’s Instagram feed. Each image on this page represents a post and when clicked, leads to its original article. The entire HTML document of the *link in bio* pages of the 6 Instagram accounts as of November 2023 were downloaded to extract links to articles. Invalid links which either 1) led to a third party website or 2) led only to the homepage of the news website or 3) only contained videos or images without a body text were excluded. A custom Python script was used to crawl the published date, headline, and body text from the articles.

After obtaining valid links from each *link in bio* page, we mapped each article to an Instagram post. Each article body text and post caption were embedded using the “multi-qa-MiniLM-L6-cos-v1” pre-trained Sentence Transformer model from Hugging Face. Then each article body text was matched with the post caption with the highest cosine similarity to it. The results were manually inspected to ensure alignment of published date and content.

### 4.3 Checking robustness of data

We check two potential sources of selection bias. First, there may be cases in which an article is adapted to multiple Instagram posts, but news accounts provide links to the article for only a select few of the posts. These posts may have distinct characteristics compared to other posts from the same article. Hence, by including only these posts with links, the dataset may be biased. However, we verify that these cases are extremely rare. We assume that posts that come from the same article discuss the same topic and hence have high semantic similarity. The proportion of consecutive pairs of posts for which semantic similarity of body text is greater than 0.8 is close to zero for all six news media (see Appendix B). Therefore, we assume that any selection bias here is negligible.

Second, there are posts that are modifications of an article, but are not linked to the article on the *link in bio* page. These posts were excluded from the study. However, they may contain inherent characteristics distinct from those that we examine. I provide an estimate of the proportion of such posts for each news media in Appendix B. These cases were rare in NBC News, Reuters, and the Washington Times. While the proportion is greater than 15% of all posts from New York Times, BBC News, and Fox News, we find that these posts behave similarly to posts included in the dataset and thus do not bias the dataset (see Appendix B).

### 4.4 Summary of the data

The dataset consists of 19,134 pairs of an online news article and its modification as an Instagram post (such a pair will be referred to as article and post pair). Table 1 is a breakdown of the number of article and post pairs, the number of images by news outlets and the publication date range of posts included. The date range varies by news outlet, depending on the number of available articles on each link in bio page at the time of data collection.

Political leaning	Left		Center		Right	
News organization	NBC News	New York Times	Reuters	BBC News	Washington Times	Fox News
Number of article and post pairs	4219	4312	462	4647	1787	3707
Earliest date	July 17, 2020	Mar 24, 2020	Jan. 2, 2023	July 17, 2021	Jan. 18, 2023	July 27, 2021
Latest date	Dec. 5, 2023	Dec. 31, 2023	Oct. 31, 2023	Nov. 13, 2023	Dec. 5, 2023	Nov. 13, 2021

Table 1: Description of the dataset

## 5 Methodology

When an article is adapted into an Instagram post, two main modifications take place: the headline and the content. The first part of the analysis examines how headlines are modified and the second part expands the discussion to a comparison of content.

### 5.1 Fuzzy string matching

Fuzzy string matching was used to find whether headlines were modified at all. Fuzzy string matching uses Levenshtein distance to find how similar two strings are. The matching generates a score between 0 and 100, where 0 indicates complete unmatching between two strings and 100 indicates two identical strings. For example, “apple” and “apple” have a score of 100, but “apple” and “orange” have a score of 36.

### 5.2 Semantic similarity

Semantic similarity was used to compare the meaning of headlines in each article and post pair. Embeddings of each headline were generated using a pre-trained Sentence Transformer model from Hugging Face, “multi-qa-MiniLM6-cos-v1” model, which is specifically trained for semantic search (OpenAI, n.d.). Then, the cosine similarity score was calculated between embeddings of an article and that of the corresponding post to estimate how semantically similar the two are. A score of 0 indicates that the headlines are about entirely different topics, while a score of 1 indicates that they are identical.

### 5.3 Sentiment analysis

Previous studies found that sentiment of news stories is often associated with audience engagement (Piotrkowicz et al., 2017; Robertson et al., 2023) and that social media content often contains emotionally charged content (Tsugawa & Ohsaki, 2015). Hence, we start both the headline and content comparison by examining how sentiment changes from an article to a post. We use Valence Aware Dictionary for Sentiment Reasoning (VADER) analysis for its suitability to analyze short texts, such as social media texts, and its ability to detect sentiment in emojis, which occasionally appeared in posts. The sentiment score ranges from -1 to 1, in which values below -0.05 represent negative sentiment, values above 0.05 represent positive sentiment and a value between -0.05 and 0.05 is a neutral sentiment. In the analyses that involves binning sentiment scores into five bins, *extremely negative* sentiment are those with scores below or equal to -0.525, *negative* sentiment are those with scores between -0.525 and -0.05, *neutral* sentiments have scores

between -0.05 and 0.05, *positive* sentiment refers to scores between 0.05 and 0.525, and *extremely positive* sentiment refers to scores above or equal to 0.525.

## 5.4 News framing analysis

We conduct a framing analysis to examine whether the same news story is written differently on articles and posts focusing on five frames: human interest, responsibility, economic consequences, conflict and morality frames. Semetko and Valkenburg (2000) developed twenty *Yes* or *No* questions to determine the presence of each frame in a news text (see Appendix C for the list of questions). We adopt a similar approach to determine whether each frame exists in each article and post. Instead of manually answering the questions, we use Open AI's GPT 3.5 Turbo model (see Appendix C for prompts). We asked the model to read each article and post and return a list of *Yes* or *No* answers to each of the 20 questions. If any of the questions corresponding to a frame contained a *Yes* as an answer, the frame was deemed to be present in the text. To validate the performance of GPT 3.5 Turbo, we randomly selected 10 articles and 10 posts from each of the six news organizations and asked six human coders to read the texts and produce answers to the same questions. Each text was read by two different human coders. The intercoder agreement was 0.968. These human-produced answers were considered to be the “ground truth” and were compared with machine-produced answers. The overall accuracy was 0.813. See Appendix D for the exact accuracy by question and by frame.

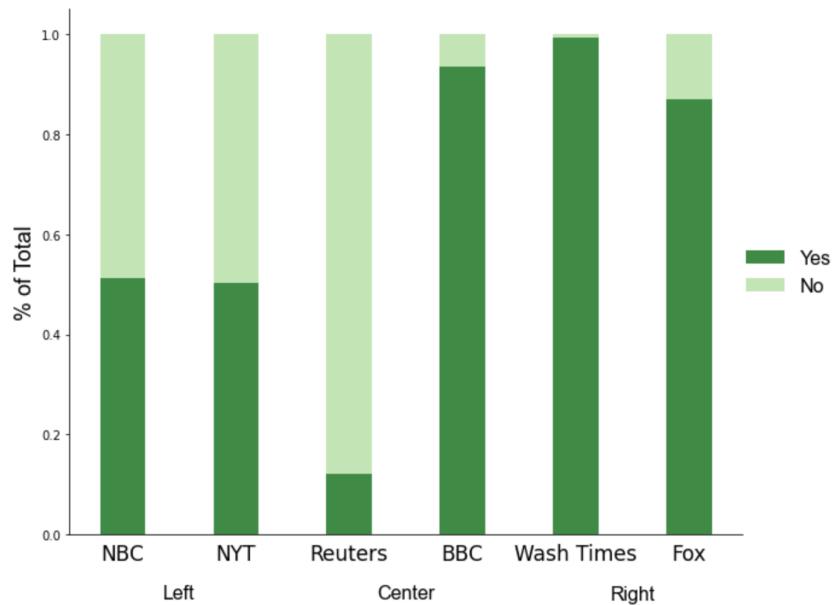
# 6 Results

## 6.1 Headline modification

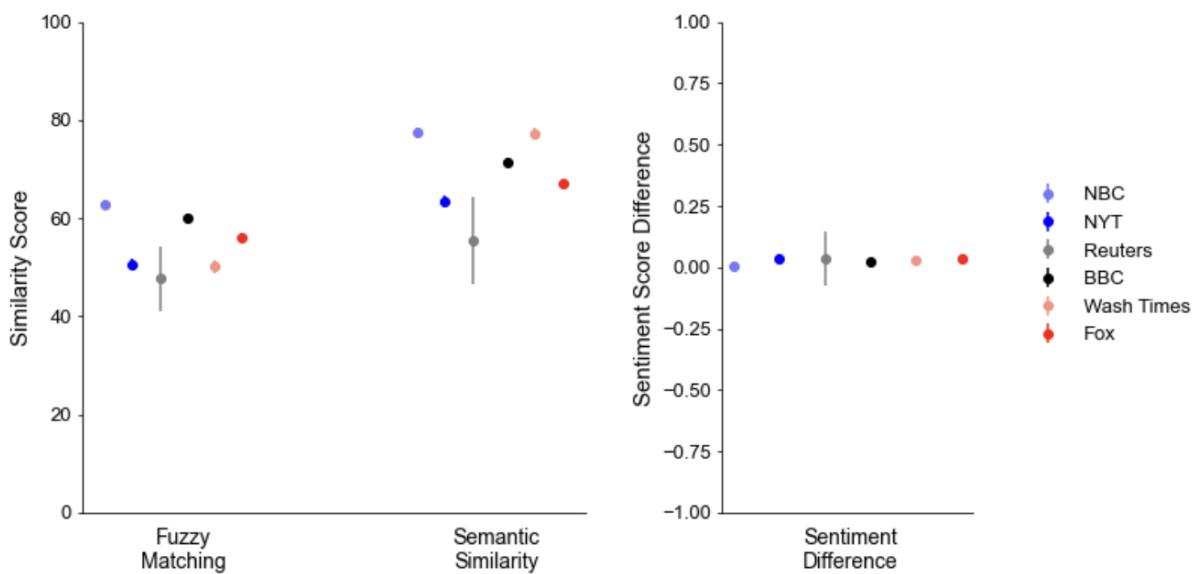
Following the approaches taken by previous studies on platform-parting, we first examined headline modifications. We first examined whether headlines exist in posts. Including headlines in posts seems to be an editorial choice, as no consistent pattern is found across all news organizations (see Figure 1a). Yet, there is a clear pattern by political leaning of each organizations. In left-wing news organizations, around half of the posts contain headlines while the other half does not. The overwhelming majority of posts from right-wing news organizations have headlines, while there are conflicting results in politically centered organizations. This presents suggestive evidence that political leaning heavily influences the social media content strategies of news organizations. The tendency of right-wing organizations to include headlines may imply a desire for an information-dense post, thenceforth including eye-catching headlines.

Using only article and post pairs that have a headline, we examined how headlines were modified. Figure 1b shows that a bulk of headlines were not drastically modified in posts and that their meanings were generally retained. Despite some variations, for all six news media, the mean of the fuzzy matching score was approximately 50%. Similarly, the distribution of semantic similarity shows that there was no significant change in the meaning of headlines, as the value for all news organizations excluding Reuters lie above 60%. Note that Reuters has a noticeably smaller number of data points compared to other organizations, which may have resulted in large error bars. Figure 1b also shows the difference in sentiment score between the article post headlines in each pair. This distribution is centered around 0 for all six news organizations, revealing that there is no noticeable difference in sentiment score between the two headlines in each pair.

Finally, we examine the distribution of sentiment scores of articles and that of posts. Figure 2a shows that in both articles and posts, headlines are concentrated in neutral sentiments with a sentiment score near zero. Figure 2b shows in more detail how the sentiment of headlines changed from article to a post. In each of the five bins representing sentiment scores of article headlines, the bulk of headlines retain the same sentiment level when modified to post headlines



(a) Proportion of Instagram posts with and without a news headline in the first image

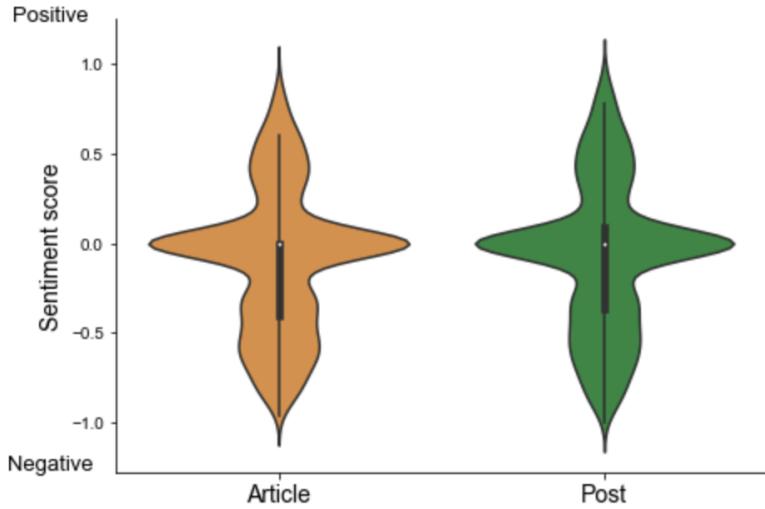


(b) Fuzzy match score, semantic similarity, and sentiment difference (sentiment score of post - sentiment score of articles) of article and post headlines for each article and post pair. Error bars represent 95% confidence intervals

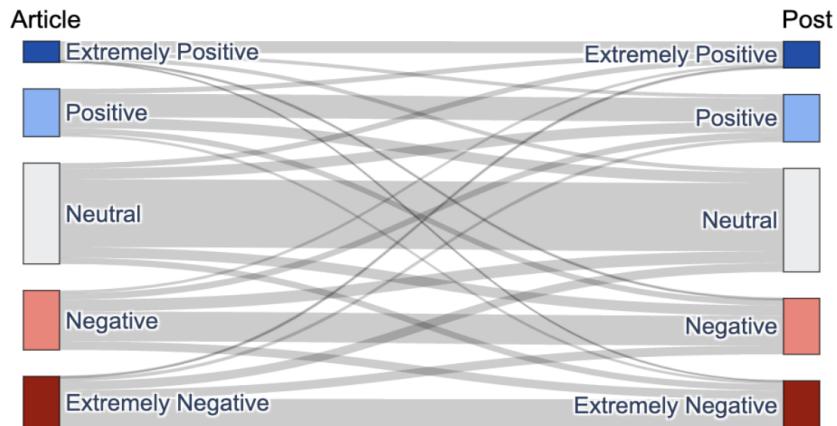
Figure 1: Headline modification in fuzzy matching, sematic similarity, and sentiment difference

(see Appendix F for conversion rates). Most of the cases where the headline sentiment changes to another bin occurs between adjacent bins and drastic modifications are rare. A similar and consistent pattern in sentiment distribution was observed within each news organization (see Appendix E). Figure 2c plots the Gini index values of the distribution of the number of headlines whose sentiment score lies in each of the five bins for articles and posts in each news organization. The two distributions of Gini indices are not significantly different and lie within relatively low values, indicating that the distribution is far from uniform.

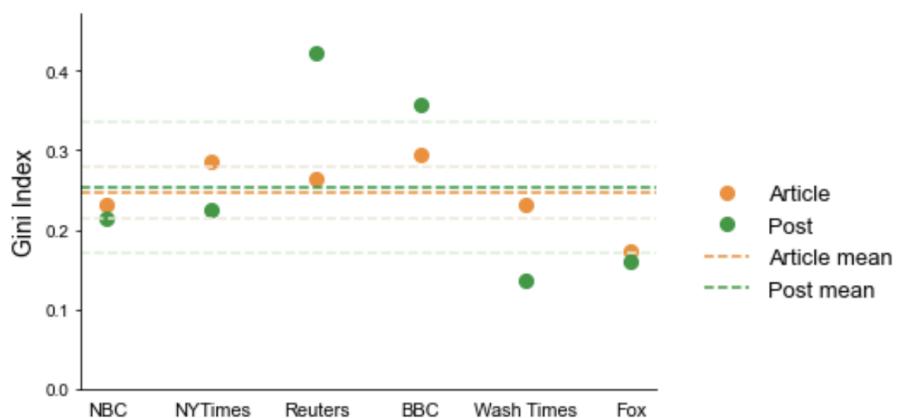
In sum, there is no credible evidence that headlines in posts are systematically changed from articles. Including a headline in a post seems to depend on each news organization and there were no considerable changes in semantics nor sentiment of headlines during platform-parting. This is in line with the findings of (Lamot et al., 2022; Park et al., 2021) that there is no general pattern found in article headline modification to a social media post across all news media. Hence, we continue the analysis to RQ2 on whether content is modified.



(a) Distribution of sentiment score in VADER sentiment analysis of article and post headlines



(b) Changes in the sentiment of headlines when an article headline is modified to a post headline. Sentiment scores are divided into 5 bins, according to the criteria mentioned in Methodology



(c) Gini index of the distribution of sentiment scores of article headlines and post headlines into 5 bins shown in Figure 2b. Higher values indicate inequality in distribution (i.e. distribution is not uniform). Lighter horizontal lines represent 95% confidence intervals

Figure 2: Sentiment score changes in headlines

## 6.2 Content modification

### 6.2.1 Distribution of content sentiment

We examined whether the content of news changes from article to post by comparing article body texts and post captions. Figure 3 shows that articles have longer texts compared to posts. This is plausible, as there is limited space for text on Instagram compared to news websites.

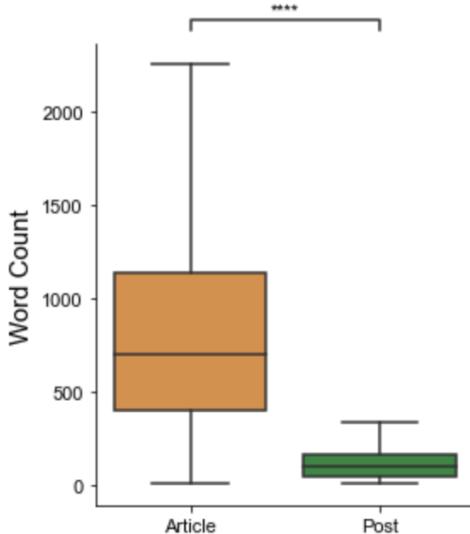
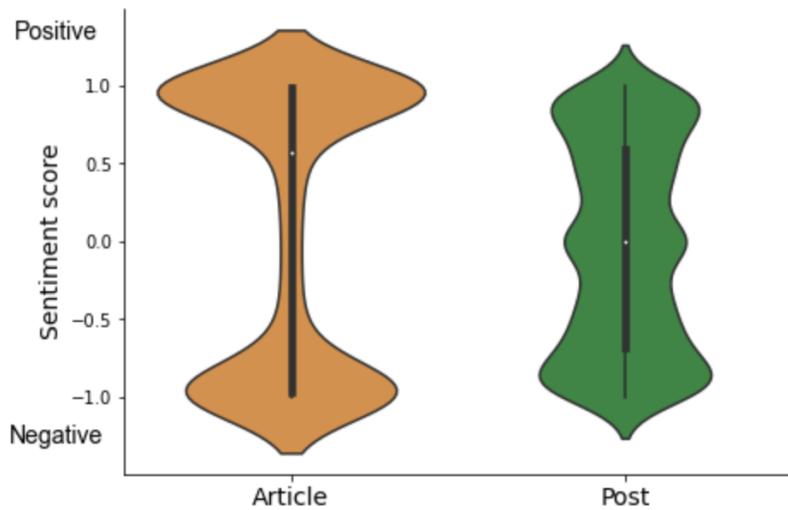


Figure 3: Distribution of the number of words in each article body text and post caption

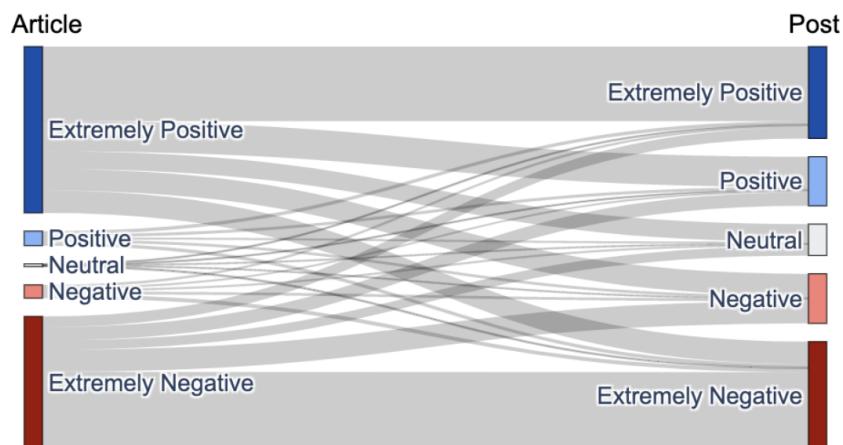
We then compare the distribution of sentiment scores. Figure 4a shows a strikingly different distribution between article and post. Article body texts tend to be polarized — that is, either very positively or very negatively written. Meanwhile, the sentiment scores of posts follow a more uniform distribution. In other words, when articles become posts, intense positive or negative sentiments of the body texts tend to move towards a neutral sentiment. Overall, the mean sentiment score of articles was 0.10, which is a positive sentiment, while the mean sentiment score of posts was -0.03, which is a neutral sentiment.

Figure 4b demonstrates in-depth how article sentiments changed in posts. There is a “toning down” of sentiment, as 16.8% of *extremely negative* article body texts become posts with *negative* sentiment and 18.2% of *extremely positive* article body texts become posts with *positive* sentiment. There are also cases of sentiment “neutralization” where positive or negative sentiments become entirely neutral (see Appendix F for all conversion rates).

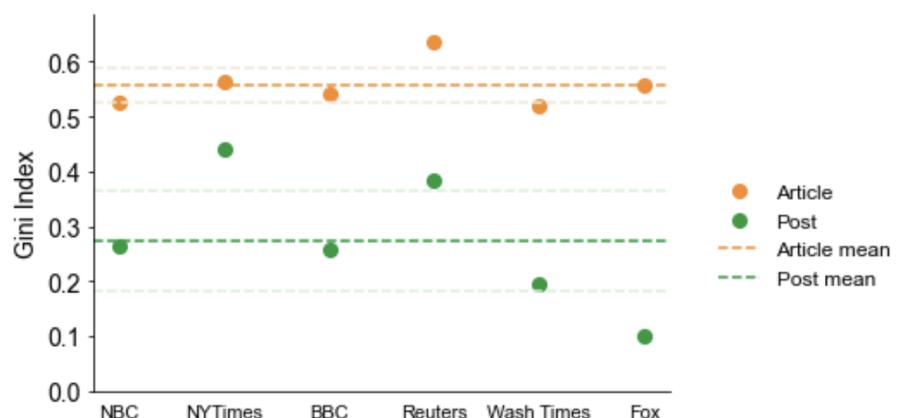
Moreover, negatively written articles are more likely to retain their negative sentiment, while positively written articles are more likely to become neutral. 11% of *extremely positive* articles become *neutral* posts, while only 7% of *extremely negative* articles become *neutral* posts. Similarly, while 26% of *extremely positive* articles were significantly modified to *negative* or *extremely negative* posts, only 18% of *extremely negative* articles become *positive* or *extremely positive* posts. Hence, negative sentiments tend to be retained more frequently than positive sentiments. These findings were consistent across all six news outlets (see Appendix G).



(a) Distribution of sentiment of articles and posts



(b) Changes in sentiment when an article is modified to a post. Sentiment scores are divided into 5 bins, according to the criteria mentioned in *Methodology*



(c) Gini index of the distribution of sentiment scores of articles and posts into 5 bins shown in Figure 4b. Lighter horizontal lines represent 95% confidence intervals

Figure 4: Sentiment score changes in news content

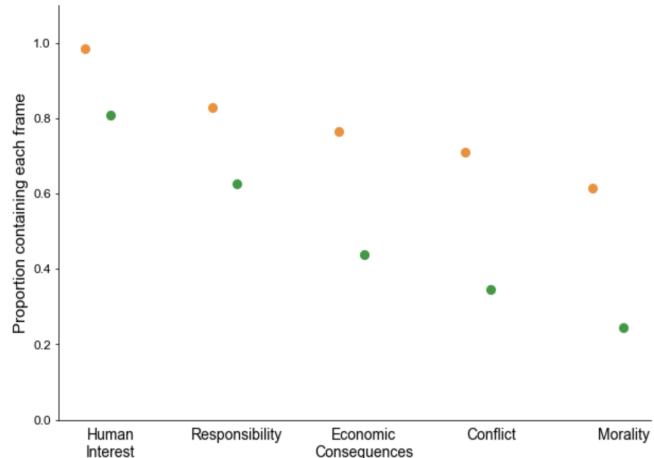
Figure 4c shows the Gini index values of the distribution of the number of texts whose sentiment score lies in each of the five bins. The two distributions of Gini indices are significantly different. Posts are more equally distributed than articles, as demonstrated by the neutralized sentiments on posts. Note that the difference in Gini index is especially bigger among right-wing news organizations. While right-wing news organizations have similar content sentiment distribution in articles as other organizations, their articles become even more neutralized in posts. The finding that right-wing news organizations tend to be less emotionally charged somewhat contradicts previous lines of research in political communication that right-wing politicians on social media heavily use negative emotions such as anger and fear (Gerbaudo et al., 2023; Salmela & von Scheve, 2018; Tuomola & Wahl-Jorgensen, 2023). This opens avenues for further research to examine potential explanations – such as right-wing media’s use of Instagram-specific features or the political climate at the time of data collection.

Overall, the results here are surprising. As social media is often associated with polarization (Campbell, Leister, & Zenou, 2019), it is often assumed to contain intense emotions. Also, studies find that emotionally charged content, especially those with intense negative emotions, tend to attract higher audience engagement, which provides incentive for news accounts to draft emotionally charged content (Brady, Wills, Jost, Tucker, & Van Bavel, 2017; Stieglitz & Dang-Xuan, 2013). However, we found that news organizations publish posts that are more neutral in sentiment than articles, with consistent results observed in subgroups of article and post pairs divided by presence of post headlines (see Appendix H). We attempt to provide a potential explanation through news framing.

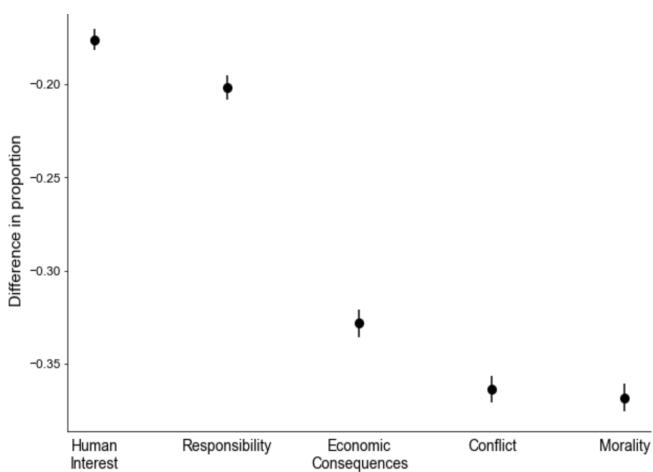
### 6.2.2 Framing analysis of news content

While sentiment analysis provides insights into **how** articles and texts were written, we endeavor to also investigate **what** is written in the body texts. For this, we referred to literature on news framing to understand what aspects of a story are emphasized in articles and posts. We used Open AI’s GPT 3.5 Turbo model to read each article and post and generate *Yes* or *No* answers to 20 questions developed by Semetko and Valkenburg (2000) (see Appendix D) to detect the presence of each frame.

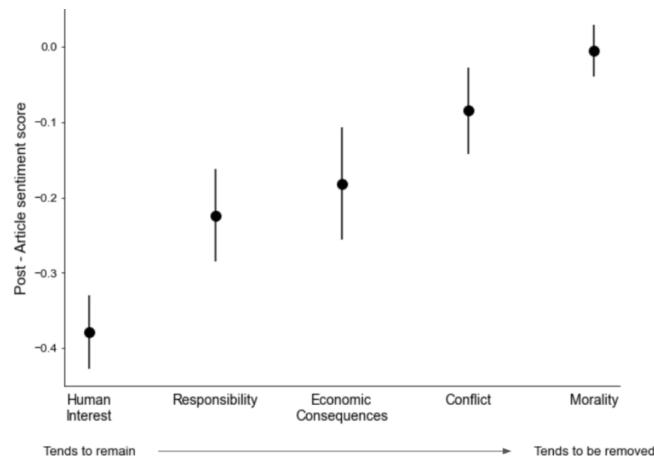
Figure 5a presents the proportion of articles and posts that contain each frame. We found that the human interest frame is the most frequently used frame in both articles and posts, while the conflict frame one of the least frequently used. This finding is unexpected, as literature suggests that the conflict frame is regularly used by journalists to make a news objective (De Vreese et al., 2010; Semetko & Valkenburg, 2000; Valenzuela et al., 2017). A reason behind this discrepancy might be that news organizations intentionally select articles which show human stories, rather than conflicts, to publish on Instagram. This is in line with Lischka and Werning (2017) and Steiner (2020) that social media editors prefer to publish “soft” news that are light and enjoyable, rather than “hard” news such as politics and international affairs which include more elements of conflict. Thus, this result presents suggestive evidence that news stories on Instagram may not be representative of the entire archive of news items available on news websites.



(a) Proportion of articles and headlines containing each frame



(b) Difference in the proportion of posts with each frame and the proportion of articles with the frame (proportion of articles - proportion of posts). Negative values in this difference indicate that a frame exists in a high proportion of articles, but a low proportion of posts and hence were often removed during modification



(c) Difference in sentiment score (sentiment score of posts - sentiment score of articles for each article and post pair) for each frame when only that frame was removed from the body texts, while the presence of other frames remained unchanged

Figure 5: Changes in news framing

The morality frame least frequently appears in both articles and Instagram posts, similar to (Wasike, 2013)'s results that this frame is the least used on Twitter. Instagram is a visual-focused platform. Thus, similar to Twitter, which has a limit on the number of characters on each post, Instagram does not encourage long body texts. Then, this finding presents evidence that news on platforms that prefer shorter texts tend to exclude mentions of morality. As the morality frame involves value judgments of what is right and wrong, (Semetko & Valkenburg, 2000) journalists may consider it to be less important than the core information of news and hence refrain from allocating space for it.

Overall, Figure 5a illustrates that the proportion of each frame is significantly higher in articles than in posts. That is, all frames appear more frequently in articles than in posts, as frames that were present in articles are sometimes removed in posts. Then, information that constitutes a news frame in an article is excluded from the posts, as the frame is omitted. Hence, the process of shortening an article to an Instagram post is not merely a *summarization* of information, but rather a *selection* of particular information at the expense of others. During platform-parting, there seems to be a **selection of articles** to publish, along with a **selection of information** to share about the news stories. Thus, when audiences read news via Instagram, compared to a full news article, they are reading texts where some aspects of a story are omitted. For robustness, we conduct the same analysis among left-wing, center and right-wing organizations and find consistent results (see Appendix I).

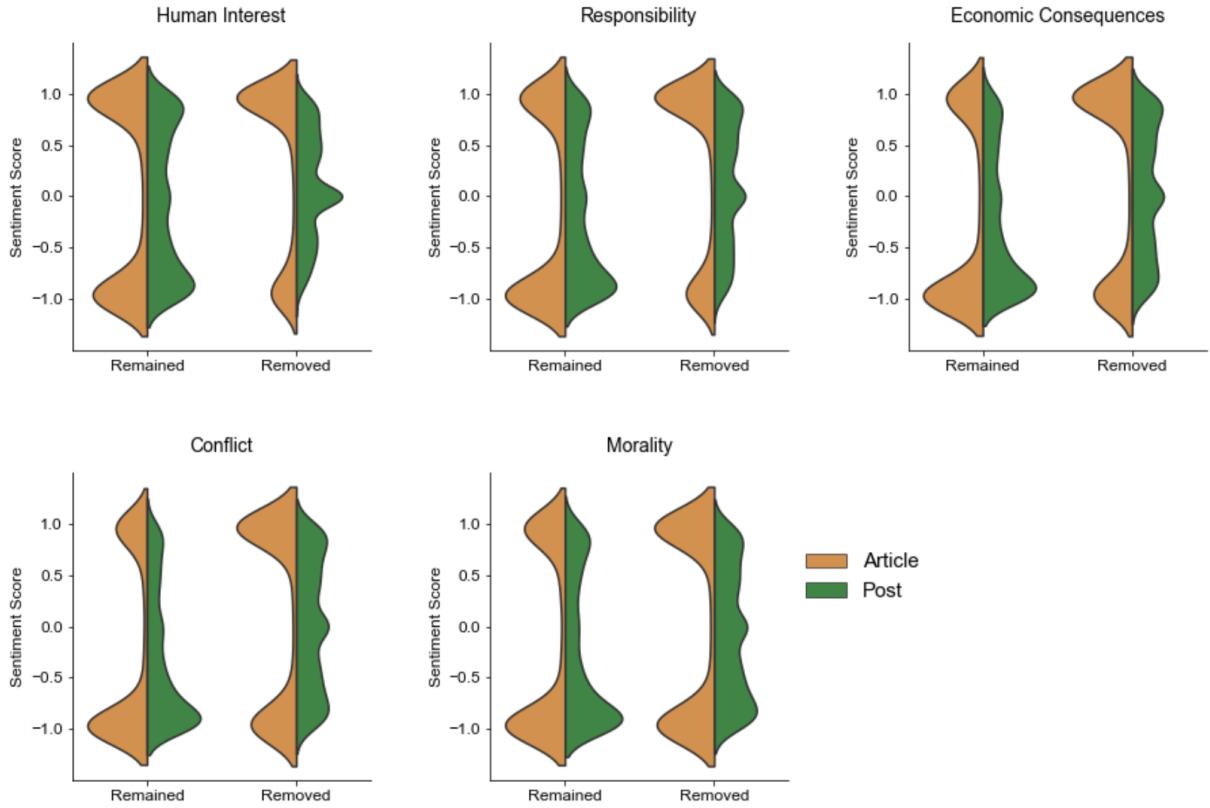
We found that such selection of information occurs with differing degrees for each frame. Figure 5b plots the gap between the proportion of posts and the proportion of articles with each frame. The gap is smallest for the human interest and responsibility frames, as these tend to remain as articles are modified to posts. This is consistent with the findings of (Guo & Sun, 2023) that Instagram captions have less personalization elements than articles. While they measured personalization with the number of names mentioned, personalization level could be interpreted similarly as the human interest frame, which also “personalizes a news” via testimonials or a storytelling of an individual (Semetko & Valkenburg, 2000). Meanwhile, the difference in proportion is greatest for the conflict and morality frames, as these frames tend to be omitted.

To further investigate the implications of this result, there needs to be an understanding of the relative importance of each frame in a news story. We indirectly measure the role that each news frame plays in news through its effect on sentiment changes. Specifically, we measure the change in sentiment score when only one frame is omitted, while other frames are unchanged. Figure 5c plots the difference in sentiment scores of posts and that of articles for such article and post pairs. The mean difference in sentiment score is negative for all frames, indicating that when each frame is removed, content sentiment scores tend to decrease. Moreover, notice that the distribution of points in Figure 5c are almost entirely the reverse of that in Figure 5b. That is, article news frames that generally remain in posts tend to decrease the sentiment by greater magnitudes, while frames that are frequently removed tend to decrease sentiment only by minimal amounts. This suggests that the human interest frame and responsibility frame play an important role in determining the emotional tone of a news story. These frames are likely considered to be necessary and are kept. Meanwhile, the morality frame has a negligible effect on sentiment changes, suggesting that it is not crucial in determining the tone of a news story. This may explain why news organizations often remove this frame in posts. This result opens multiple opportunities for future work. First, one can verify whether the human interest and responsibility frames are also considered to be the most important by news readers. Moreover, through experiments, one can study the effect of removing each frame on news perception, such as attitude towards issue reported and information recall.

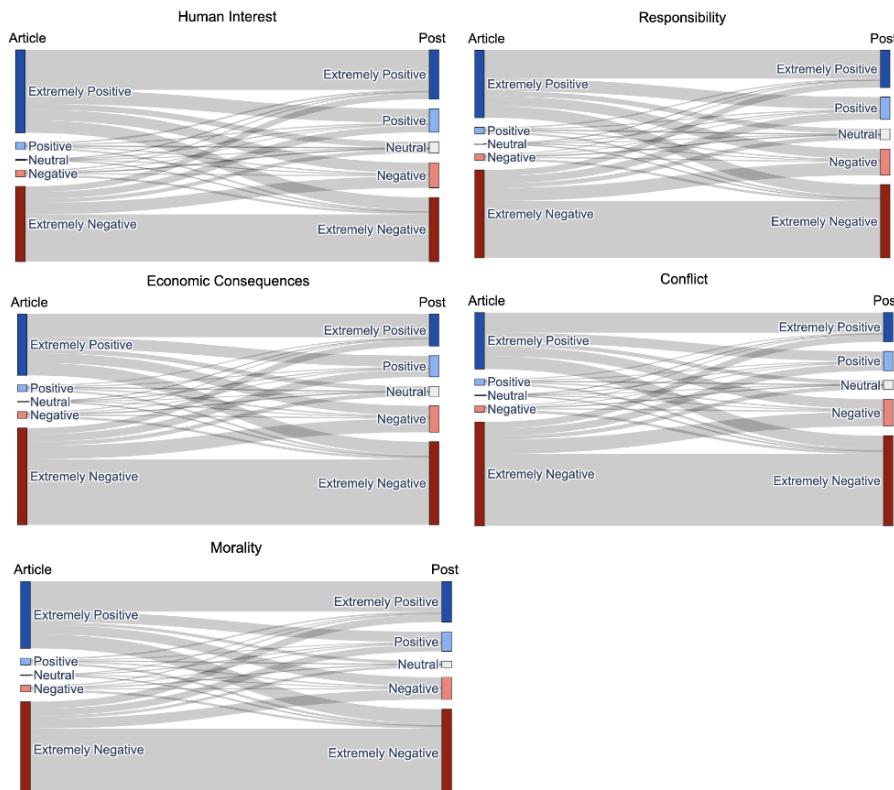
Next, we examined whether the results are consistent when observing pairs for which a frame was kept and removed. Figure 6a shows that sentiment is neutralized more noticeably among article and post pairs for which a frame was removed. The sentiment distribution of articles is bimodal near the two extreme sentiment scores in all cases, as seen previously. However, posts where a frame was removed have a higher concentration of neutral sentiments compared to those that keep each frame. This is to say that the sentiment neutralization may be derived from sentiment changes in articles for which one or more frames were removed. Hence, when news organizations remove some frames from the text, while retaining others, they also neutralize the sentiment.

Figure 6b shows in more detail how content sentiment changes from an article to a post when each frame remains in text. This is the case where the content itself remains the same, as each frame is kept. Overall, the pattern of neutralization is repeated here, as the distribution of content sentiment in posts are uniform than that of articles. Upon closer look, however, we notice that articles written with negative sentiments are generally rewritten into a post also with a negative sentiment, while positively written articles tend to be neutralized into posts. This means that if an article contains a frame that is written in a negative manner, then this negative sentiment is likely stays in the frame in posts. But if an article contains a positively written frame, then the sentiment of the frame likely neutralizes in the corresponding post. In the previous analysis in Figure 4b, we noted that there is a higher proportion of negatively written articles that retain the same negative sentiment in posts than that of positively written articles. Figure 6b presents suggestive evidence that many of such negatively written posts are those that retain their negativity in frames. Again, this contrasts with positively written articles whose frames become toned down as articles are modified to posts, and also with articles whose frames disappear in posts, in which the sentiment was neutralized for all subgroups.

The framing analysis shows that during platform-parting, news frames are frequently removed. Conflict and morality frames contribute least to the emotional tone of a news article and are the most frequently omitted, while responsibility and human interest frames are generally kept. Also, sentiment neutralization is especially strong when a frame is removed, compared to when it is kept. When the presence of a frame is unchanged in a pair, negative sentiments tend to be kept, while positive sentiments are neutralized.



(a) Distribution of sentiment score of articles and posts for which each frame was removed and retained



(b) Changes in sentiment, if any, of content when an article body text is modified to a post caption, only for article and post pairs for which a frame has remained

Figure 6: News framing changes and sentiment

## 7 Conclusion

### 7.1 Main findings

This study endeavors to comprehensively understand platform-parting strategies of news from online news articles into Instagram posts. First, we found no consistent pattern in headline modification across all news organizations examined. Many posts did not include a headline and even those that did not go through drastic alterations. Contrastingly, we identified significant changes in content. When article body texts are rewritten into Instagram post texts, the sentiments become toned down and neutralized. Through framing analysis, we also found that news organizations may be selecting particular articles to disseminate on Instagram along with a selective inclusion of information within a post, presenting the audience with less information overall. When frames are removed, neutralization of sentiments occurs strongly. Meanwhile, when frames remain in texts, there is a tendency for texts to retain negative sentiments, while toning down positive sentiments.

Overall, the results suggest that news organizations use clickbait strategies on social media. First, there seems to be a preference for negative sentiments in Instagram news content. A higher proportion of negatively written articles retain the negative sentiment in posts, compared to positively written articles that remain positive. We also found that when a frame from an article remains in post, negative sentiments tend to be kept as well. In addition, human interest frames, which highlights an individual and emphasizes one's personal story, were often included in posts and were the least likely to be omitted. Previous studies on user engagement have found that sentiment-charged, personal stories are associated with higher engagement rates (Guo & Sun, 2023; Piotrkowicz et al., 2017; Robertson et al., 2023). Therefore, my findings suggest that clickbait strategies of online news and those of social media content in general constitute decision-making in producing Instagram news content.

Also, this study identifies how platform-parting occurs on Instagram specifically. Lack of a consistent pattern of headline modifications align with studies on other social media platforms such as Facebook (Lamot et al., 2022; Park et al., 2021; Verstappen & Opgenhaffen, 2023). Also, the prevalence of human stories align with findings that news outlets prefer "soft" news in Facebook posts (Lischka & Werning, 2017; Steiner, 2020). As such, future studies in the field can conduct a similar content analysis on other social media platforms to identify commonalities and differences in platform parting on each platform.

We found suggestive evidence of patterns that emerge among right-wing news organizations. Washington Times and Fox News tend to include news headlines in their posts and the distribution of the sentiment score of posts is more even, compared to other organizations. This suggests that right-wing organizations tend to provide additional information via headlines and neutral, fact-based captions. Future research can investigate the reasons behind these findings. In this study, data collection was conducted from November 2023 to January 2024, during which the right-wing were not the dominant political parties in the United States. This may have prompted right-wing media to adopt new social media strategies to ensure that their content stands out. Hence, future research can conduct similar analysis across a longer span of time to confirm the robustness of the findings. Also, it may be that right-wing media tend to utilize images to convey emotions, more than texts. Hence, post captions may be neutral overall, but the images used in posts may be more emotionally charged compared to other news organizations. Thus, a multi-modal analysis of images and videos on news posts can further shed light on this observation.

## 7.2 Discussion

This study is the first attempt to collect articles and their direct adaptations into Instagram posts and compare each pair in terms of headlines and content. Essentially, the study suggests that nowadays an increasing number of people are reading a different kind of news that is not only shorter, but void of strong emotions and has less diverse information. Hence, a natural question to ask is whether Instagram news has a different impact on people's perception of news compared to articles. Future research can test the effect of the observed differences between articles and posts through an experimental survey. The experiment can involve two treatment groups – one that receives a text corresponding to an article and another with a text corresponding to the adapted corresponding post. This can measure how audience reception of news, such as emotional reactions, attitude towards reported issue, and accuracy of information recall, changes depending on the platform used to read news.

The findings also have implications for broader social media conversations. A considerable portion of social media content concerns news. For example, Kwak, Lee, Park, and Moon (2010) finds that 85% of status posts on Twitter are news-related. Therefore, the findings advance the literature on understanding what kinds of content are being published on Instagram in general. Moreover, note that Appendix A suggests that social media exclusive content, which were excluded from the dataset, exhibit sentiment neutralization similar to the results of the study. Specifically, for NBC News, Reuters, and Washington Times, the three news outlets with high proportions of social media exclusive content, sentiment scores of posts that we excluded are even more clustered around 0, compared to that of posts that we included. This confirms the robustness of my findings and suggests that content on Instagram may be less emotionally charged than what one may commonly assume.

Related to this, while we exclusively focus on content **published** by news organizations, future studies can expand the scope of analysis to look at how these news content **circulate** on social media. There are accounts on Instagram such as @impact, which are not run by official news organizations, but repost and modify news content published by news organizations. They create social media exclusive content aimed at spreading news to more users. This process is distinct from platform-parting, as sharing of news happens within a platform, but it is nonetheless an adaptation of a news story to reach a wider audience. Studies can examine whether sentiment and news framing are altered during this process, as it is possible that social media polarization originates from the circulation of news.

Moreover, here we focus on the general patterns observed from modification. However, there are rare cases that were not further examined. For instance, 3% of *extremely negative* article headlines became *extremely positive* post headlines and 3% of *extremely positive* article headlines became *extremely negative* post headlines. Also, 89% of articles whose content was *neutral* became posts whose sentiment was either positive or negative, although the absolute number of such articles were small. These rare cases deserve further qualitative examination to identify any underlying reasons behind drastic alterations.

Although we balance political representation and conduct robustness checks of the data, we cannot draw out conclusive evidence or causal links in how articles are modified into Instagram posts, as this study is based on descriptive analysis of six news organizations. Further research can include articles and posts from more organizations encompassing diverse regions and languages. Nonetheless, the organizations chosen in the study are among the news accounts with the highest Instagram followers and hence have wide audience reach. Based on my findings, a qualitative study with in-depth interviews with social media editors of these news outlets could also shed light on underlying motivations behind these modifications.

Lastly, this study proposes an efficient methodology of conducting framing analysis, leveraging advanced large language models (LLMs). While scholars conducting framing analysis have generally relied on human coders to read and identify frames in a news text (De Vreese et al., 2010; Gronemeyer & Porath, 2017; Semetko & Valkenburg, 2000; Wasike, 2013), such methods are labor-intensive and time-consuming. Using the same theoretical framework and criteria for detecting news frames, we adopted an automated approach to finding frames by using the GPT 3.5 Turbo model by Open AI to answer simple text comprehension questions. Future work can utilize state-of-the-art LLMs to increase the speed and accuracy of framing analysis.

Overall, this paper shows that an increasing number of people are now reading news that is written differently than in traditional news platforms. While the sentiment of content is more neutral, which may signal that social media news is less polarizing, texts are written narrowly with fewer frames. These findings provide grounds for numerous future work to unpack the new era of digital journalism in the social media space.

## References

- AdFontesMedia. (n.d.). *Interactive Media Bias Chart*. Retrieved 2024-03-24, from <https://adfontesmedia.com/interactive-media-bias-chart/>
- AllSides. (2019, February). *AllSides Media Bias Chart*. Retrieved 2024-03-24, from <https://www.allsides.com/media-bias/media-bias-chart>
- An, S.-K., & Gower, K. K. (2009, June). How do the news media frame crises? A content analysis of crisis news coverage. *Public Relations Review*, 35(2), 107–112. Retrieved 2024-03-29, from <https://www.sciencedirect.com/science/article/pii/S036381109000113> doi: 10.1016/j.pubrev.2009.01.010
- Bartholomé, G., Lecheler, S., & de Vreese, C. (2015, October). Manufacturing Conflict? How Journalists Intervene in the Conflict Frame Building Process. *The International Journal of Press/Politics*, 20(4), 438–457. Retrieved 2024-03-29, from <https://doi.org/10.1177/1940161215595514> (Publisher: SAGE Publications Inc) doi: 10.1177/1940161215595514
- Brady, W., Wills, J., Jost, J., Tucker, J., & Van Bavel, J. (2017, July). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. Retrieved 2024-03-29, from <https://www.pnas.org/doi/abs/10.1073/pnas.1618923114> (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.1618923114
- Campbell, A., Leister, M., & Zenou, Y. (2019, August). Social Media and Polarization. *CEPR Discussion Paper No. DP13860*. Retrieved 2024-04-21, from [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3428384](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3428384)
- Carr, C. T., & Hayes, R. A. (2015, January). Social Media: Defining, Developing, and Divining. *Atlantic Journal of Communication*, 23(1), 46–65. Retrieved 2024-04-18, from <https://doi.org/10.1080/15456870.2015.972282> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/15456870.2015.972282>) doi: 10.1080/15456870.2015.972282
- Cho, S. H., & Gower, K. K. (2006, November). Framing effect on the public's response to crisis: Human interest frame and crisis type influencing responsibility and blame. *Public Relations Review*, 32(4), 420–422. Retrieved 2024-03-29, from <https://www.sciencedirect.com/science/article/pii/S036381106001056> doi: 10.1016/j.pubrev.2006.09.011
- Deutsch, A., & van den Berg, S. (2023, October). *What war crimes laws apply to the Israel-Palestinian conflict?* / Reuters. Retrieved 2024-04-10, from [https://www.reuters.com/world/middle-east/what-war-crimes-laws-apply-israel-palestinian-conflict-2023-10-11/?utm\\_campaign=trueanthem\\_manual&utm\\_medium=trueAnthem&utm\\_source=instagram](https://www.reuters.com/world/middle-east/what-war-crimes-laws-apply-israel-palestinian-conflict-2023-10-11/?utm_campaign=trueanthem_manual&utm_medium=trueAnthem&utm_source=instagram)
- De Vreese, C., H., Peter, J., & Semetko, H. (2010, November). Framing Politics at the Launch of the Euro: A Cross-National Comparative Study of Frames in the News. *Political Communication*, 18(2), 107–122. Retrieved 2024-03-29, from <https://doi.org/10.1080/105846001750322934> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/105846001750322934>) doi: 10.1080/105846001750322934
- Dijck, J. v., & Poell, T. (2013, August). Understanding Social Media Logic. *Media and Communication*, 1(1), 2–14. Retrieved 2024-03-26, from <https://www.cogitatiopress.com/mediaandcommunication/article/view/70> doi: 10.17645/mac.v1i1.70
- Druckman, J. N. (2001, September). The Implications of Framing Effects for Citizen Competence. *Political Behavior*, 23(3), 225–256. Retrieved 2024-03-29, from <https://doi.org/10.1023/A:1015006907312> doi: 10.1023/A:1015006907312
- Dvir-Gvirsman, S., & Tsuriel, K. (2022, August). In an Open Relationship: Platformization of Relations Between News Practitioners and Their Audiences. *Journalism Studies*, 23(11), 1308–1326. Retrieved 2024-03-26, from <https://doi.org/10.1080/1461670X.2022.2084144> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/1461670X.2022.2084144>) doi: 10.1080/1461670X.2022.2084144

- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. Retrieved 2024-03-29, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-2466.1993.tb01304.x> (\_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-2466.1993.tb01304.x>) doi: 10.1111/j.1460-2466.1993.tb01304.x
- Figenschou, T. U., & Thorbjørnsrud, K. (2015, June). Faces of an Invisible Population: Human Interest Framing of Irregular Immigration News in the United States, France, and Norway. *American Behavioral Scientist*, 59(7), 783–801. Retrieved 2024-03-29, from <https://doi.org/10.1177/0002764215573256> (Publisher: SAGE Publications Inc) doi: 10.1177/0002764215573256
- Fletcher, R. (2023, June). *Attitudes towards algorithms and their impact on news / Reuters Institute for the Study of Journalism*. Retrieved 2024-04-16, from <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/attitudes-towards-algorithms-impact-news>
- Fortin, J. (2023, October). Cherokee Tribe Shares Native American Knowledge in North Carolina - The New York Times. Retrieved 2024-04-10, from [https://www.nytimes.com/2023/10/21/us/cherokee-indians-north-carolina.html?utm\\_source=dash%2520hudson&utm\\_medium=instagram&utm\\_campaign=likeshopme&utm\\_content=ig-nytimes](https://www.nytimes.com/2023/10/21/us/cherokee-indians-north-carolina.html?utm_source=dash%2520hudson&utm_medium=instagram&utm_campaign=likeshopme&utm_content=ig-nytimes)
- Gerbaudo, P., De Falco, C. C., Giorgi, G., Keeling, S., Murolo, A., & Nunziata, F. (2023, January). Angry Posts Mobilize: Emotional Communication and Online Mobilization in the Facebook Pages of Western European Right-Wing Populist Leaders. *Social Media + Society*, 9(1), 20563051231163327. Retrieved 2024-03-29, from <https://doi.org/10.1177/20563051231163327> (Publisher: SAGE Publications Ltd) doi: 10.1177/20563051231163327
- Goldbaum, C., Padshah, S., & Blanchard, E. (2023, October). Driven Out of Pakistan, Afghans Face an Uncertain Future. *The New York Times*. Retrieved 2024-04-10, from <https://www.nytimes.com/2023/10/30/world/asia/afghanistan-pakistan-deportations.html>
- Gronemeyer, M. E., & Porath, W. (2017, July). Framing Political News in the Chilean Press: The Persistence of the Conflict Frame. *International Journal of Communication*, 11(0), 24. Retrieved 2024-03-29, from <https://ijoc.org/index.php/ijoc/article/view/6882> (Number: 0)
- Guo, M., & Sun, F.-S. (2023). Local Television News on Instagram: Exploring the Effects of News Values and Post Features on Audience Engagement. *International Journal on Media Management*, 0(0), 1–22. Retrieved 2024-03-27, from <https://doi.org/10.1080/14241277.2023.2204528> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/14241277.2023.2204528>) doi: 10.1080/14241277.2023.2204528
- Hanusch, F. (2017). Web analytics and the functional differentiation of journalism cultures: individual, organizational and platform-specific influences on newsworthiness. *Information, Communication and Society*, 20(10), 1571–1586. Retrieved 2024-03-26, from <https://eprints.qut.edu.au/102976/> (Number: 10 Publisher: Routledge)
- Hase, V., Boczek, K., & Scharkow, M. (2023, September). Adapting to Affordances and Audiences? A Cross-Platform, Multi-Modal Analysis of the Platformization of News on Facebook, Instagram, TikTok, and Twitter. *Digital Journalism*, 11(8), 1499–1520. Retrieved 2024-03-27, from <https://doi.org/10.1080/21670811.2022.2128389> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/21670811.2022.2128389>) doi: 10.1080/21670811.2022.2128389
- Hermida, A., & Mellado, C. (2020, August). Dimensions of Social Media Logics: Mapping Forms of Journalistic Norms and Practices on Twitter and Instagram. *Digital Journalism*, 8(7), 864–884. Retrieved 2024-03-27, from <https://doi.org/10.1080/21670811.2020.1805779> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/21670811.2020.1805779>) doi: 10.1080/21670811.2020.1805779

- Holligan, A. (2021, December). *Covid: Dutch go into Christmas lockdown over Omicron wave*. Retrieved 2024-04-10, from [https://www.bbc.com/news/world-europe-59715940?utm\\_campaign=later-linkinbio-bbcnews&utm\\_content=later-23229219&utm\\_medium=social&utm\\_source=linkin.bio](https://www.bbc.com/news/world-europe-59715940?utm_campaign=later-linkinbio-bbcnews&utm_content=later-23229219&utm_medium=social&utm_source=linkin.bio)
- Hong, H. (2013, July). The Effects of Human Interest Framing in Television News Coverage of Medical Advances. *Health Communication*, 28(5), 452–460. Retrieved 2024-03-29, from <https://doi.org/10.1080/10410236.2012.693013> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/10410236.2012.693013>) doi: 10.1080/10410236.2012.693013
- Howard, P. N., & Parks, M. R. (2012, April). Social Media and Political Change: Capacity, Constraint, and Consequence. *Journal of Communication*, 62(2), 359–362. Retrieved 2024-04-18, from <https://doi.org/10.1111/j.1460-2466.2012.01626.x> doi: 10.1111/j.1460-2466.2012.01626.x
- Iyengar, S. (1996, July). Framing Responsibility for Political Issues. *The ANNALS of the American Academy of Political and Social Science*, 546(1), 59–70. Retrieved 2024-03-29, from <https://doi.org/10.1177/0002716296546001006> (Publisher: SAGE Publications Inc) doi: 10.1177/0002716296546001006
- Kim, S.-H. (2015, September). Who Is Responsible for a Social Problem? News Framing and Attribution of Responsibility. *Journalism & Mass Communication Quarterly*, 92(3), 554–558. Retrieved 2024-03-29, from <https://doi.org/10.1177/1077699015591956> (Publisher: SAGE Publications Inc) doi: 10.1177/1077699015591956
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010, April). What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web* (pp. 591–600). New York, NY, USA: Association for Computing Machinery. Retrieved 2024-03-26, from <https://dl.acm.org/doi/10.1145/1772690.1772751> doi: 10.1145/1772690.1772751
- Lamot, K. (2022, May). What the Metrics Say. The Softening of News on the Facebook Pages of Mainstream Media Outlets. *Digital Journalism*, 10(4), 517–536. Retrieved 2024-04-19, from <https://doi.org/10.1080/21670811.2021.1974917> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/21670811.2021.1974917>) doi: 10.1080/21670811.2021.1974917
- Lamot, K., Kreutz, T., & Opgenhaffen, M. (2022, July). “We Rewrote This Title”: How News Headlines Are Remediated on Facebook and How This Affects Engagement. *Social Media + Society*, 8, 205630512211148. doi: 10.1177/20563051221114827
- Liedke, J., & Wang, L. (2023a, November). *News Platform Fact Sheet*. Retrieved 2024-03-28, from <https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/>
- Liedke, J., & Wang, L. (2023b, November). *Social Media and News Fact Sheet*. Retrieved 2024-04-16, from <https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/>
- Lischka, J. A. (2021, February). Logics in social media news making: How social media editors marry the Facebook logic with journalistic standards. *Journalism*, 22(2), 430–447. Retrieved 2024-04-16, from <https://doi.org/10.1177/1464884918788472> (Publisher: SAGE Publications) doi: 10.1177/1464884918788472
- Lischka, J. A., & Werning, M. (2017). Wie Facebook den Regionaljournalismus verändert: Publikums- und Algorithmusorientierung bei der Facebook-Themenselektion von Regionalzeitungen. *kommunikation @ gesellschaft*, 18, 25.
- Lynch, S. N., & Stempel, J. (2023). *Trump risked national secrets, prosecutors allege in indictment / Reuters*. Retrieved 2024-04-10, from [https://www.reuters.com/legal/trump-faces-federal-charges-classified-documents-case-adding-legal-woes-2023-06-09/?utm\\_campaign=trueanthem\\_manual&utm\\_medium=trueAnthem&utm\\_source=instagram](https://www.reuters.com/legal/trump-faces-federal-charges-classified-documents-case-adding-legal-woes-2023-06-09/?utm_campaign=trueanthem_manual&utm_medium=trueAnthem&utm_source=instagram)
- Morrow, G., Swire-Thompson, B., Polny, J. M., Kopec, M., & Wihbey, J. P. (2022). The emerging science of content labeling: Contextualizing social media content moderation. *Journal of the Association for Information Science and Technology*, 73(10), 1365–1386. Retrieved 2024-

- 04-18, from <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.24637> (\_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.24637>) doi: 10.1002/asi.24637
- Mottola, M. (2020, July). *The Best Of Social Media: Pros And Cons Of Each Platform*. Retrieved 2024-03-28, from <https://www.forbes.com/sites/theeyec/2020/07/20/the-best-of-social-media-pros-and-cons-of-each-platform/> (Section: Entrepreneurs)
- Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Nielsen, R. K. (2021, June). *Reuters Institute Digital News Report 2021* [SSRN Scholarly Paper]. Rochester, NY. Retrieved 2024-04-19, from <https://papers.ssrn.com/abstract=3873260>
- Nieborg, D. B., & Poell, T. (2018, November). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20. Retrieved 2024-03-26, from <https://dare.uva.nl/search?identifier=4bc08c57-95af-449a-9a1b-d0aef50e96b0> doi: 10.1177/1461444818769694
- Ofcom. (2023, July). *News consumption in the UK: 2023*. Author. Retrieved 2024-04-17, from [https://www.ofcom.org.uk/\\_\\_data/assets/pdf\\_file/0024/264651/news-consumption-2023.pdf](https://www.ofcom.org.uk/__data/assets/pdf_file/0024/264651/news-consumption-2023.pdf)
- OpenAI. (n.d.). *Pretrained Models — Sentence-Transformers documentation*. Retrieved 2024-03-24, from [https://www.sbert.net/docs/pretrained\\_models.html](https://www.sbert.net/docs/pretrained_models.html)
- Pan, Z., & Kosicki, G. M. (1993, January). Framing analysis: An approach to news discourse. *Political Communication*, 10(1), 55–75. Retrieved 2024-03-29, from <https://www.tandfonline.com/doi/abs/10.1080/10584609.1993.9962963> (Publisher: Routledge \_eprint: <https://www.tandfonline.com/doi/pdf/10.1080/10584609.1993.9962963>) doi: 10.1080/10584609.1993.9962963
- Park, K., Kwak, H., An, J., & Chawla, S. (2021, April). *How-to Present News on Social Media: A Causal Analysis of Editing News Headlines for Boosting User Engagement*. arXiv. Retrieved 2024-03-27, from <http://arxiv.org/abs/2009.08100> (arXiv:2009.08100 [cs]) doi: 10.48550/arXiv.2009.08100
- Piotrkowicz, A., Dimitrova, V., Otterbacher, J., & Markert, K. (2017, May). The Impact of News Values and Linguistic Style on the Popularity of Headlines On Twitter and Facebook. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1), 767–774. Retrieved 2024-03-29, from <https://ojs.aaai.org/index.php/ICWSM/article/view/14979> (Number: 1) doi: 10.1609/icwsm.v11i1.14979
- Roberts, S. T. (2018, March). Digital detritus: 'Error' and the logic of opacity in social media content moderation. *First Monday*. Retrieved 2024-04-18, from <https://firstmonday.org/ojs/index.php/fm/article/view/8283> doi: 10.5210/fm.v23i3.8283
- Robertson, C. E., Pröllochs, N., Schwarzenegger, K., Pärnamets, P., Van Bavel, J. J., & Feuerriegel, S. (2023, May). Negativity drives online news consumption. *Nature Human Behaviour*, 7(5), 812–822. doi: 10.1038/s41562-023-01538-4
- Salmela, M., & von Scheve, C. (2018, November). Emotional Dynamics of Right- and Left-wing Political Populism. *Humanity & Society*, 42(4), 434–454. Retrieved 2024-03-29, from <https://doi.org/10.1177/0160597618802521> (Publisher: SAGE Publications Inc) doi: 10.1177/0160597618802521
- Semetko, H., & Valkenburg, P. (2000). Framing European politics: a content analysis of press and television news. *Journal of Communication*, 50(2), 93–109. Retrieved 2024-03-24, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-2466.2000.tb02843.x> (\_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-2466.2000.tb02843.x>) doi: 10.1111/j.1460-2466.2000.tb02843.x
- Shearer, E. (2021, January). *More than eight-in-ten Americans get news from digital devices*. Retrieved 2024-03-26, from <https://www.pewresearch.org/short-reads/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/>
- Steiner, M. (2020, August). Soft Presentation of Hard News? A Content Analysis of Political Facebook Posts. *Media and Communication*, 8(3), 244–257. Retrieved 2024-04-19, from

<https://www.cogitatiopress.com/mediaandcommunication/article/view/3152> doi:  
10.17645/mac.v8i3.3152

- Stieglitz, S., & Dang-Xuan, L. (2013, April). Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior. *Journal of Management Information Systems*, 29(4), 217–248. Retrieved 2024-03-29, from <https://doi.org/10.2753/MIS0742-1222290408> (Publisher: Routledge \_eprint: <https://doi.org/10.2753/MIS0742-1222290408>) doi: 10.2753/MIS0742-1222290408
- Tsugawa, S., & Ohsaki, H. (2015). Negative messages spread rapidly and widely on social media. In *Proceedings of the 2015 acm on conference on online social networks* (pp. 151–160).
- Tsuriel, K., Dvir Gvirsman, S., Ziv, L., Afriat-Aviv, H., & Ivan, L. (2021, August). Servant of two masters: How social media editors balance between mass media logic and social media logic. *Journalism*, 22(8), 1983–2000. Retrieved 2024-04-16, from <https://doi.org/10.1177/1464884919849417> (Publisher: SAGE Publications) doi: 10.1177/1464884919849417
- Tuomola, S., & Wahl-Jorgensen, K. (2023, January). Emotion Mobilisation through the Imagery of People in Finnish-Language Right-Wing Alternative Media. *Digital Journalism*, 11(1), 61–79. Retrieved 2024-03-29, from <https://doi.org/10.1080/21670811.2022.2061551> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/21670811.2022.2061551>) doi: 10.1080/21670811.2022.2061551
- Valenzuela, S., Piña, M., & Ramírez, J. (2017, October). Behavioral Effects of Framing on Social Media Users: How Conflict, Economic, Human Interest, and Morality Frames Drive News Sharing. *Journal of Communication*, 67(5), 803–826. Retrieved 2024-03-29, from <https://doi.org/10.1111/jcom.12325> doi: 10.1111/jcom.12325
- Verstappen, M., & Opgenhaffen, M. (2023). Making it Fit: How Science News Gets Remediated for Facebook and Instagram. *Journalism Studies*, 0(0), 1–19. Retrieved 2024-04-19, from <https://doi.org/10.1080/1461670X.2023.2263799> (Publisher: Routledge \_eprint: <https://doi.org/10.1080/1461670X.2023.2263799>) doi: 10.1080/1461670X.2023.2263799
- Vorhaus, M. (2020, June). *People Increasingly Turn To Social Media For News*. Retrieved 2024-03-28, from <https://www.forbes.com/sites/mikevorhaus/2020/06/24/people-increasingly-turn-to-social-media-for-news/> (Section: Media)
- Wasike, B. S. (2013, June). Framing News in 140 Characters: How Social Media Editors Frame the News and Interact with Audiences via Twitter. *Global Media Journal: Canadian Edition*, 6(1), 5–23. Retrieved 2024-03-29, from [http://www.gmj.uottawa.ca/1301/v6i1\\_wasike.pdf](http://www.gmj.uottawa.ca/1301/v6i1_wasike.pdf) (Publisher: Saint Paul University)
- Webster, J. G. (2014). *The Marketplace of Attention: How Audiences Take Shape in a Digital Age*. MIT Press. (Google-Books-ID: iBNmBAAQBAJ)

## Appendix

### A. Definition and examples of five news frames used in the study

Frame	Definition	Example
Responsibility	Presents an issue or problem in such a way as to attribute responsibility for causing or solving to either the government or to an individual or group	"The Dutch prime minister described this lockdown as an unavoidable response to Omicron. But the government's critics contend the current crisis is partly of the ministers' own making, citing what was perceived by some as a clumsy handling of the Delta strain and relatively slow roll-out of the booster vaccination programme. The over-60s have only just been invited to get their extra shots." (Holligan, 2021)
Human interest	Brings a human face, an individual's story, or an emotional angle to the presentation of an event, issue or problem	"I tried my best in these 40 years to build a life," said Mr. Torjan, 63, the truck parked behind him at the border. "It's difficult. Now I'm starting again from zero" (Goldbaum, Padshah, & Blanchard, 2023)
Conflict	Emphasizes conflict between individuals, groups, institutions or countries	"U.S. prosecutors unsealed a 37-count indictment against Donald Trump on Friday, accusing the former president of risking some of the country's most sensitive security secrets after leaving the White House in 2021" (Lynch & Stempel, 2023)
Morality	Interprets an event or issue in the context of religious tenets or moral prescriptions	"Deliberate killings of civilians, hostage-taking, and collective punishment are heinous crimes that have no justification" (Deutsch & van den Berg, 2023)
Economic consequences	Presents an event, problem or issue in terms of the economic consequences it will have on an individual, group, institution, region or country	"Sales taxes from these shops support the tribe, along with revenues from a casino that generates hundreds of millions of dollars annually" (Fortin, 2023)

Table A1: Definition and examples of each frame. Definitions are from Semetko and Valkenburg (2000). Examples are excerpts from news articles in the dataset, identified by GPT 3.5 Turbo. See prompt in Appendix C

## B. Robustness check of dataset

I check for two possible sources of selection bias arising from data collection. A considerable portion of Instagram posts from all six accounts did not have a link to a news article in the *link in bio* page. Manual inspection showed that some of such posts were social media exclusive content that were not modified from an article and hence irrelevant to this project. However, other posts were modified from an existing news article, but the news account chose not to provide a direct link to the original article. As the reason behind such decisions is unknown, there are two possible sources of bias in the dataset.

First, if there is a series of posts that were modified from the same news article, then the account may have chosen to provide a link to the article on only a few of the posts. Then the dataset only includes the post linked to an article, excluding others that may have different characteristics. To check for this possibility, the semantic similarities between the captions of all consecutive pairs of Instagram posts were found. If the cosine similarity between embeddings of two consecutive captions were above 0.8, then the pair is considered to be adapted from the same article. However, such cases were extremely rare (Table A2). Thus, any selection bias that may arise from these cases were assumed to be negligible.

NBC	New York Times	Reuters	BBC	Washington Times	Fox News
5.30e-6	1.84e-6	6.14e-5	1.82e-6	8.28e-7	3.95e-8

Table A2: Proportion of consecutive pairs of posts for which the semantic similarity of captions is greater than 0.8

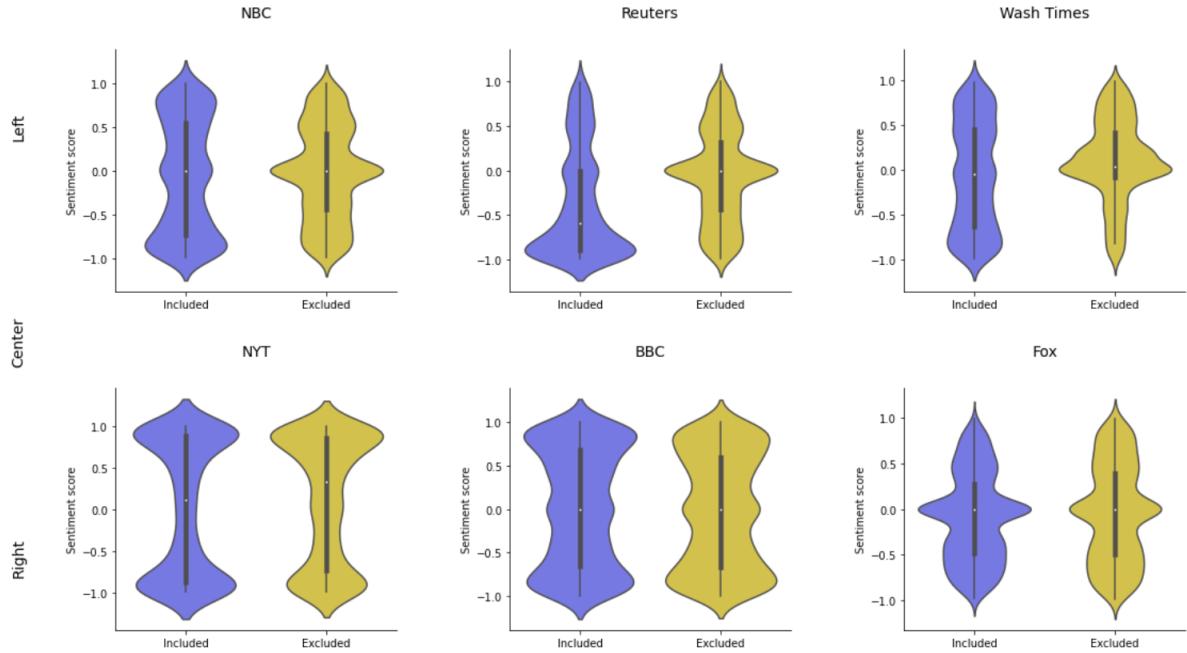
Second, posts for which there was no link provided to an original article in the *link in bio* pages were excluded from the dataset. This may produce a biased sample, as it contains only article and post pairs that news organizations are willing to share. Note that there are two cases for such posts. First, the post may be a social media exclusive content that is not a modification of an article. These posts are irrelevant for this study. Second, the post may be adapted from an article, but the link to the article was not provided for an unknown reason. To estimate the proportion of posts falling in these two cases, I randomly select 100 posts from each news organizations without a link to an article and manually examine them. For each post examined, if an article from the corresponding news organization was found in the first three pages of a search on Google and the article was published within three days before the post, then the post was considered to be adapted from the article. Otherwise, the post was considered to be social media exclusive. The results are summarized in Table A3.

The proportions of posts that are adapted from articles, but were not included in the dataset are negligible for NBC News, Reuters, and Washington Times. Hence, there is limited evidence that selection bias affected the analysis for these three news organizations. However, the proportions of New York Times, BBC News, and Fox News require further examination whether posts that are adapted from an article but were excluded from the dataset introduced bias. As the main finding of the paper involves an analysis of the distribution of the sentiment in posts, in Figure A1a, I examined the sentiment distribution of posts that were included in the dataset and those that were excluded. A clear pattern is visible: for New York Times, BBC News, and Fox News, the sentiment distribution of the two groups of posts are highly similar. For further robustness, I also look at the word count of post captions of the two groups (Figure A1b). For the three news organizations, the distribution of word count is similar between posts that were included in the dataset and those that were not. This suggests that posts that were adapted from news articles behave similarly whether or not news organizations provided a link to the original article. Hence,

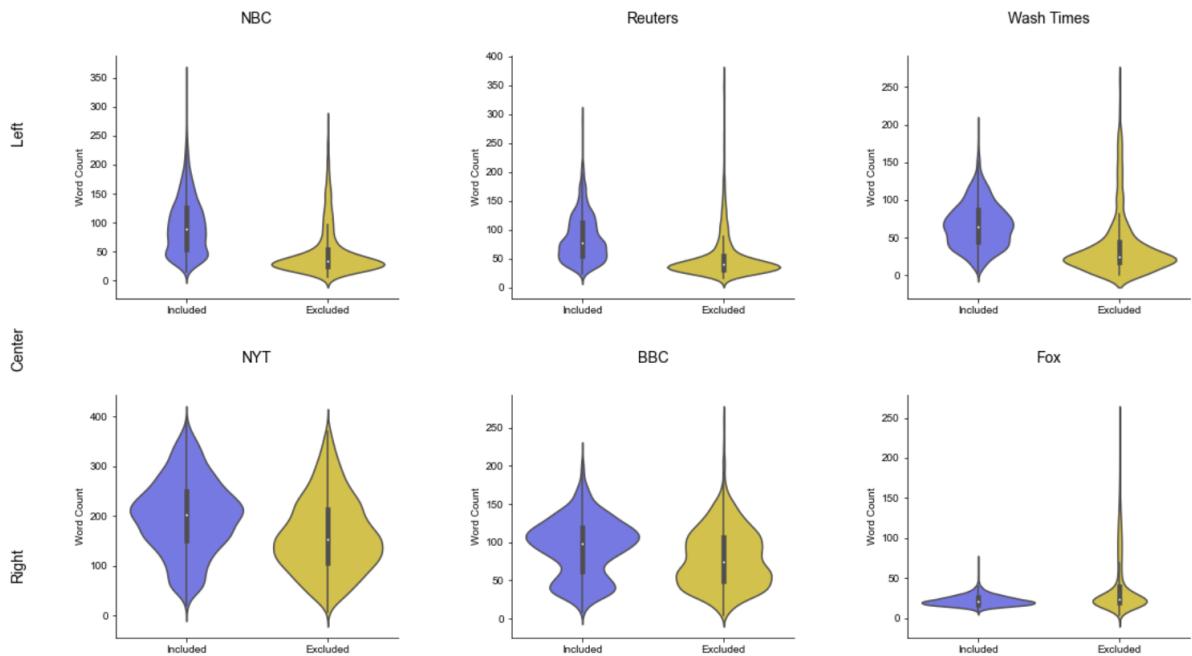
News Outlet	Total number of posts in date range (total)	Proportion of posts with valid link to article ( $v$ )	Proportion of posts with invalid/no link ( $iv$ )	Proportion of $iv$ that are adapted from article ( $p$ )	Proportion of $t$ that are adapted from article, but excluded from dataset ( $p'$ )
NBC News	4955	0.85	0.15	0.34	0.051
New York Times	7233	0.60	0.40	0.48	0.192
Reuters News	542	0.85	0.15	0.41	0.062
BBC News	6372	0.73	0.27	0.55	0.149
Washington News	2199	0.81	0.19	0.36	0.068
Fox News	15907	0.23	0.77	0.3	0.231

Table A3: **total:** Total number of Instagram posts uploaded by each news organization within each respective date range specified in Table 1.  $v$ : Proportion of all posts that have a valid link to a news article and were included in the dataset.  $iv$ : Proportion of all posts that do not have a valid link to an article and were excluded from analysis.  $iv = t - v$ .  $p$ : Estimated proportion of excluded posts  $iv$  that are adapted from a news article, but no link was provided by the outlet.  $p'$ : Estimated proportion of all posts that were adapted from the article, but were excluded from the dataset.  $p' = iv \cdot p$ .

there is limited evidence of selection bias in the data.



(a) Distribution of sentiment scores of post captions that were included and excluded from the dataset



(b) Distribution of number of words of post captions that were included and excluded from the dataset

Figure A1: Robustness check of data

### C. Questions and prompts for GPT 3.5 Turbo used in framing analysis

<b>Frame</b>	<b>Number</b>	<b>Question</b>
Responsibility	1	Does the story suggest that some level of government has the ability to alleviate the problem?
	2	Does the story suggest that some level of the government is responsible for the issue/problem?
	3	Does the story suggest solution(s) to the problem/issue?
	4	Does the story suggest that an individual (or group of people in society) is responsible for the issue/problem?
	5	Does the story suggest the problem requires urgent action?
Human interest	6	Does the story provide a human example or “human face” on the issue?
	7	Does the story employ adjectives or personal vignettes that generate feelings of outrage, empathy-caring, sympathy, or compassion?
	8	Does the story emphasize how individuals and groups are affected by the issue/problem?
	9	Does the story go into the private or personal lives of the actors?
	10	Does the story contain visual information that might generate feelings of outrage, empathy, caring, sympathy, or compassion?
Conflict	11	Does the story reflect disagreement between parties/individuals/groups/countries?
	12	Does one party-individual-group-country reproach another?
	13	Does the story refer to two sides or to more than two sides of the problem or issue?
	14	Does the story refer to winners and losers?
Morality	15	Does the story contain any moral message?
	16	Does the story make reference to morality, God, and other religious tenets?
	17	Does the story offer specific social prescriptions about how to behave?
Economic consequences	18	Is there a mention of financial losses or gains now or in the future?
	19	Is there a mention of the costs/degree of expense involved?
	20	Is there a reference to economic consequences of pursuing or not pursuing a course of action?

Table A4: 20 questions developed by Semetko and Valkenburg (2000) to test the presence of each frame in a news text

Task	Prompt
Answer 20 questions	<p>"You are a smart assistant capable of reading a text and answering 20 comprehension questions about the text."</p> <p>Follow these guidelines:  I will provide you with a list of 20 questions.  For each question, you must answer 'Yes' or 'No'.  1. Does the story suggest that some level of government has the ability to alleviate the problem?  2. Does the story suggest that some level of the government is responsible for the issue/problem?  3. Does the story suggest solution(s) to the problem/issue?  4. Does the story suggest that an individual (or group of people in society) is responsible for the issue/problem?  5. Does the story suggest the problem requires urgent action?  6. Does the story provide a human example or "human face" on the issue?  7. Does the story employ adjectives or personal vignettes that generate feelings of outrage, empathy-caring, sympathy, or compassion?  8. Does the story emphasize how individuals and groups are affected by the issue/problem?  9. Does the story go into the private or personal lives of the actors?  10. Does the story contain visual information that might generate feelings of outrage, empathy, caring, sympathy, or compassion?  11. Does the story reflect disagreement between parties-individuals-groups-countries?  12. Does one party-individual-group-country reproach another?  13. Does the story refer to two sides or to more than two sides of the problem or issue?  14. Does the story refer to winners and losers?  15. Does the story contain any moral message?  16. Does the story make reference to morality, God, and other religious tenets?  17. Does the story offer specific social prescriptions about how to behave?  18. Is there a mention of financial losses or gains now or in the future?  19. Is there a mention of the costs/degree of expense involved?  20. Is there a reference to economic consequences of pursuing or not pursuing a course of action?</p> <p>Return your answer in a JSON format.  The key of the JSON should be an integer, which is the number of the question you are answering.  ONLY output the JSON result.</p> <p>Text:  Output:</p>
Extract excerpts for each frame	<p>"You are a smart assistant capable of reading a text and extracting excerpts from the text that correspond to each of the 5 categories that I will define below."</p> <p>Follow these guidelines:  I will provide you with definitions of 5 different categories of information. I will also provide a text.  Your job is to read the text and extract excerpts from the text that correspond to each of the categories.  If there is no excerpt that corresponds to a category, return an empty string.  Below are the definitions of the five categories.  Conflict: emphasizes conflict between individuals, groups, institutions or countries  Human interest: brings a human face, an individual's story, or an emotional angle to the presentation of an event, issue or problem  Responsibility: presents an issue or problem in such a way as to attribute responsibility for causing or solving to either the government or to an individual or group  Morality: interprets an event or issue in the context of religious tenets or moral prescriptions  Economic consequences: Presents an event, problem or issue in terms of the economic consequences it will have on an individual, group, institution, region or country  Return your answer in a JSON format.  The key of the JSON should be the name of each category, as I defined above.  If the excerpt contains a double quotation mark, remove the double quotation mark from the excerpt.  ONLY output the JSON result.</p> <p>Text:  Output:</p>

Table A5: Prompts used for answering 20 questions and extracting excerpts for each news frame

## D. Validation of GPT 3.5 Turbo Performance

To validate the performance of GPT 3.5 Turbo in reading an article or a post and answering the 20 questions in Appendix C, I randomly selected 10 articles and 10 posts from each news organization. Two human coders were asked to read each news text and produce answers to the same questions. Such human-produced manual answers were used by Semetko and Valkenburg (2000), from which I adopt the methodology. The human-produced answers were considered to be the “ground truth.” The overall accuracy of machine-produced output was 0.813. Table A6a shows the accuracy by question. Excluding questions 14, 17, 18, 19 and 20, the automated approach had an accuracy of above 0.75. Note that questions 18, 19 and 20 test the economic consequences frame. The automated approach was less successful in detecting this frame, compared to others.

Next, I test the accuracy of the GPT 3.5 Turbo-produced results in detecting the presence of a frame in a text. As specified in the Methodology section, if at least one of the questions representing a frame receives a *Yes* as an answer, then the frame was deemed to be present. I find whether each frame exists in the randomly selected texts, using the human-produced result and consider this to be the “ground truth” to compare to the machine-produced result. The overall accuracy over all frames was 0.85. Table A6b summarizes the accuracy by frame.

Question	Accuracy	Question	Accuracy
<b>1</b>	0.9	<b>11</b>	0.9
<b>2</b>	0.917	<b>12</b>	0.875
<b>3</b>	0.825	<b>13</b>	0.783
<b>4</b>	0.867	<b>14</b>	0.683
<b>5</b>	0.892	<b>15</b>	0.767
<b>6</b>	0.908	<b>16</b>	0.908
<b>7</b>	0.85	<b>17</b>	0.708
<b>8</b>	0.892	<b>18</b>	0.625
<b>9</b>	0.842	<b>19</b>	0.733
<b>10</b>	0.817	<b>20</b>	0.575

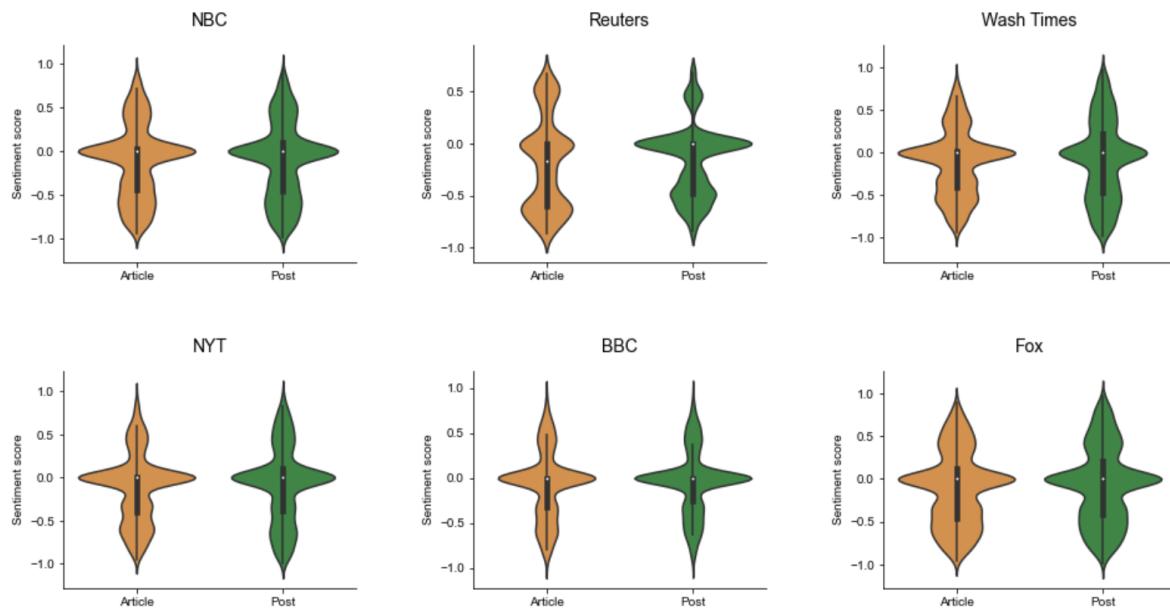
- (a) Accuracy of GPT 3.5 Turbo-produced results for answering 20 questions.

Frame	Accuracy
<b>Responsibility</b>	0.917
<b>Human Interest</b>	0.967
<b>Morality</b>	0.875
<b>Conflict</b>	0.733
<b>Economic Consequences</b>	0.758

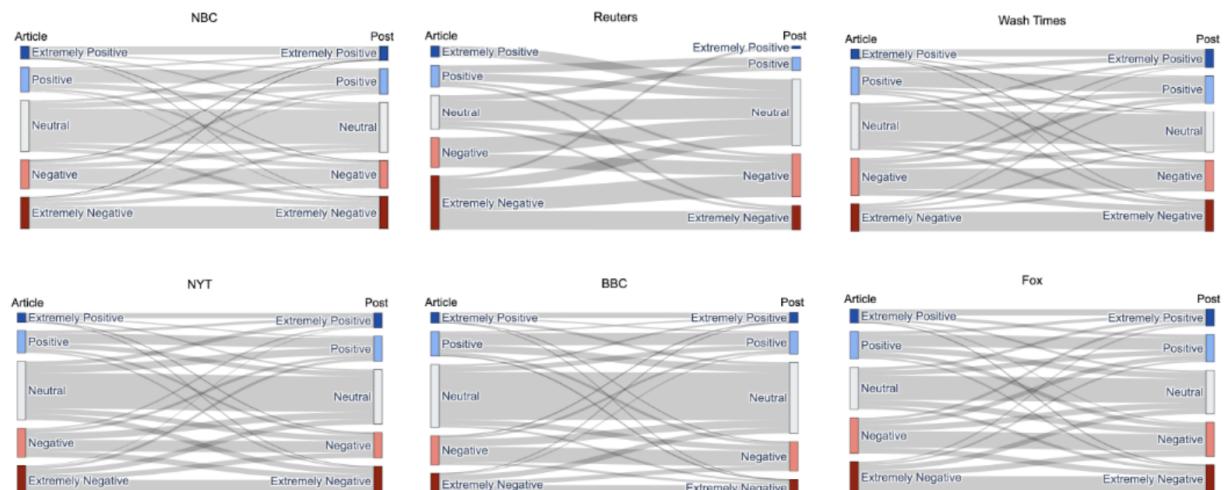
- (b) Accuracy of GPT 3.5 Turbo-produced results detecting the presence of each frame.

Table A6: Accuracy of GPT 3.5 Turbo-produced results

## E. Headline sentiment analysis by news media



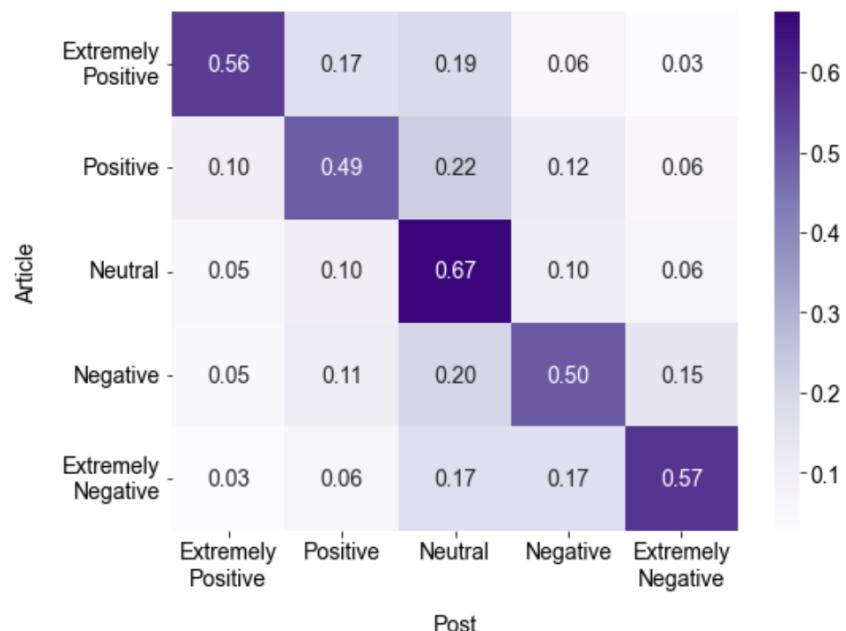
- (a) Distribution of sentiment score from VADER sentiment analysis of the headlines of articles and posts for each of the six news organizations



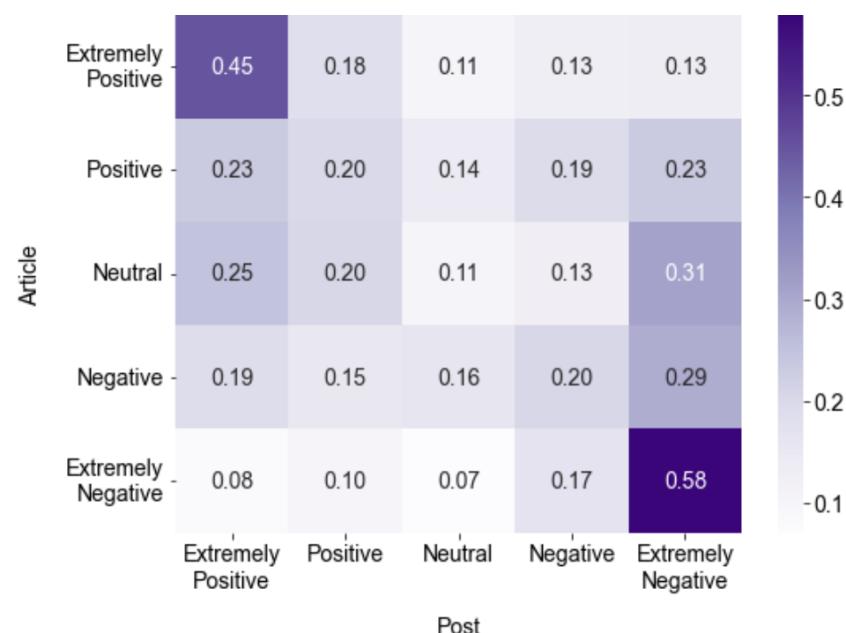
- (b) Changes in sentiment, if any, of content when an article headline is modified to a post headline

Figure A2: Headline sentiment changes by news organization

## F. Conversion rates in sentiment changes from articles to posts



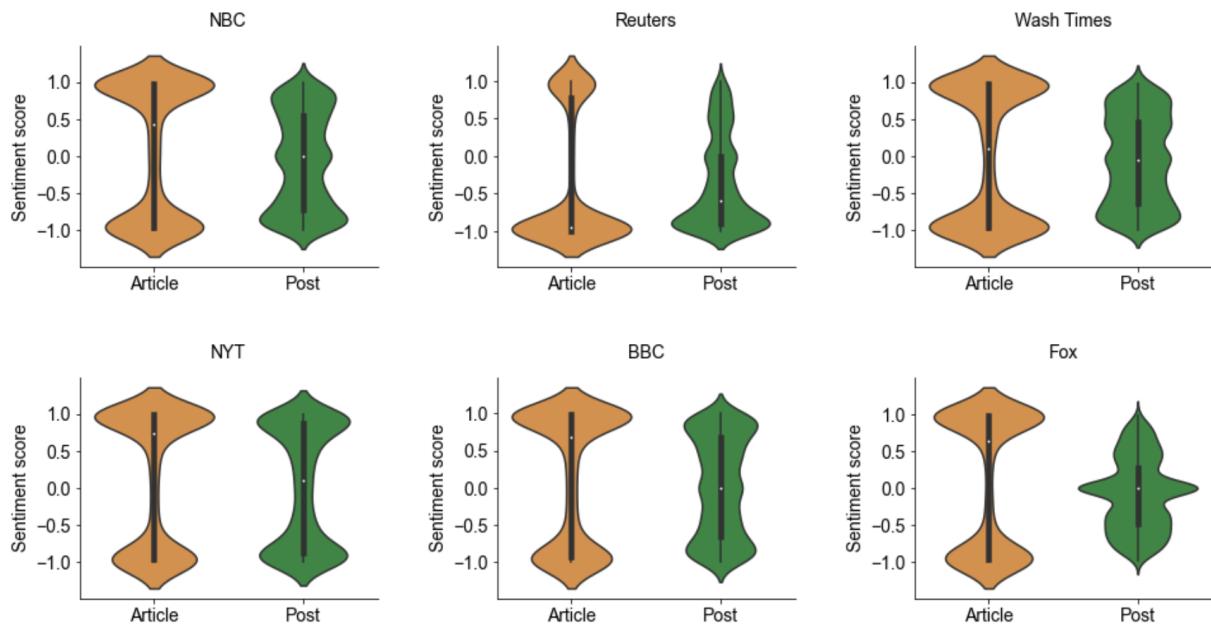
(a) Conversion rates in sentiment bins from article headlines to post headlines



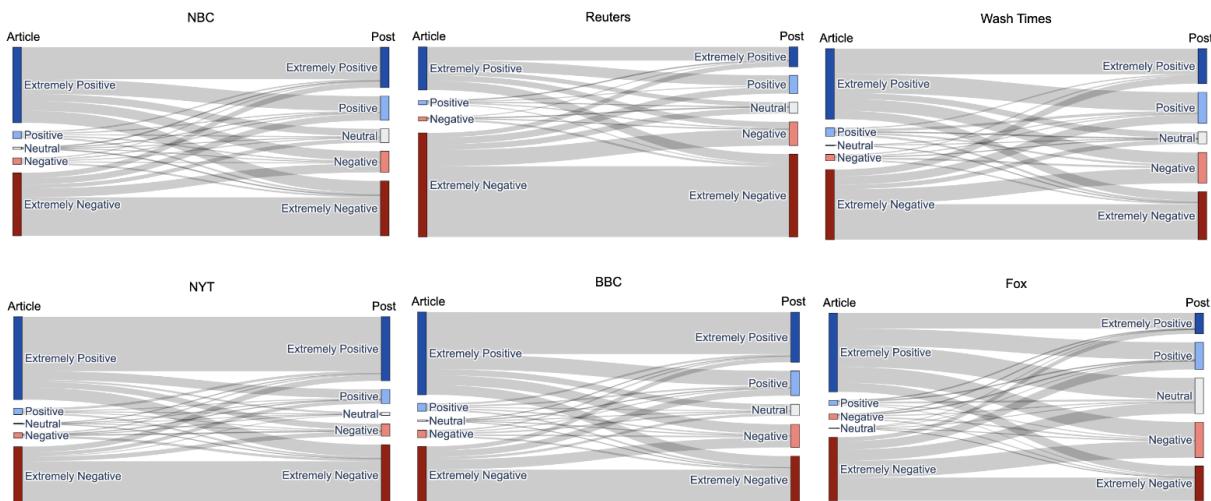
(b) Conversion rates in sentiment bins from article body text to posts

Figure A3: Conversion rates of sentiment changes

## G. Content sentiment analysis by news media



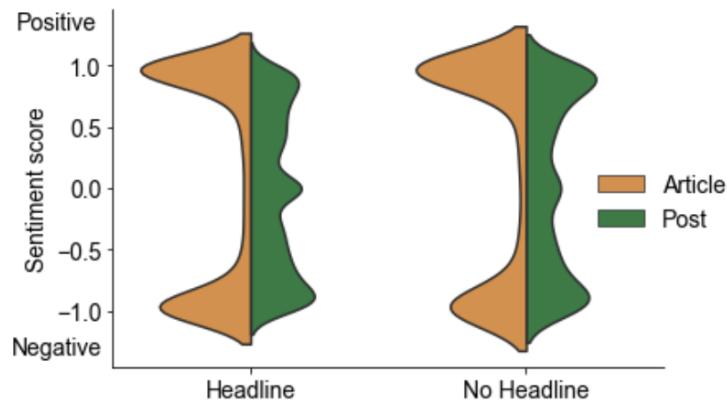
(a) Distribution of sentiment score from VADER sentiment analysis of the body texts of articles and posts for each of the six news media



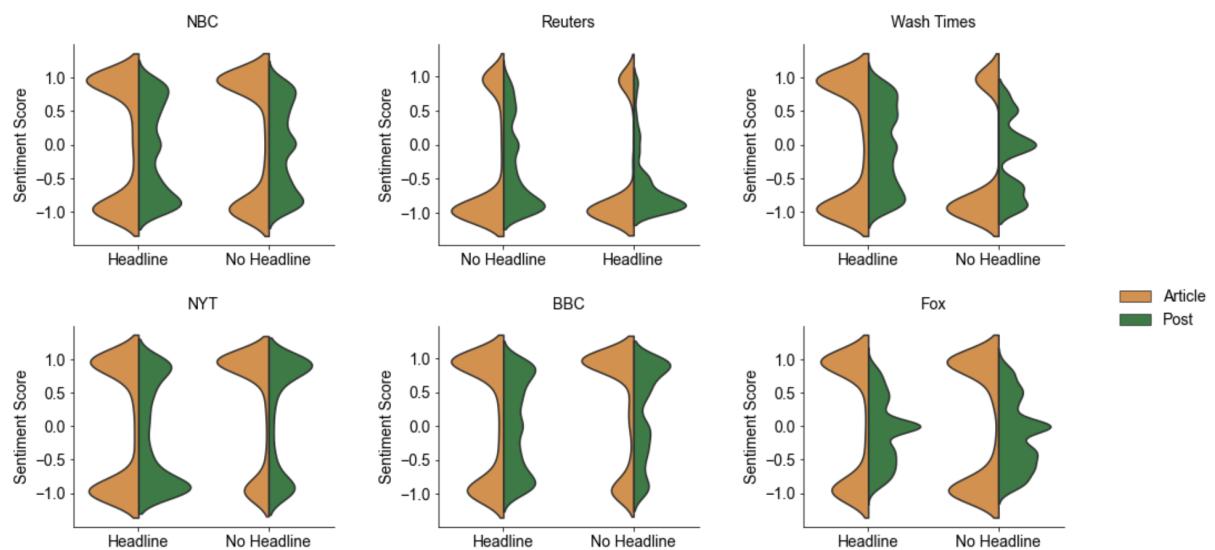
(b) Changes in sentiment, if any, of content when an article body text is modified to a post caption. Sentiment scores are divided into 5 bins, according to the criteria mentioned in *Methodology*

Figure A4: Content sentiment changes for each news organization

**H. Distribution of sentiment of article body text and post captions, separated by presence of headlines in post**



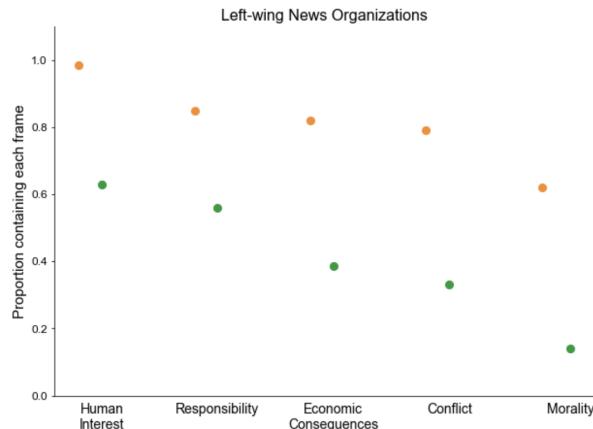
(a) Distribution of sentiment score of content, separated by the presence of headlines in posts



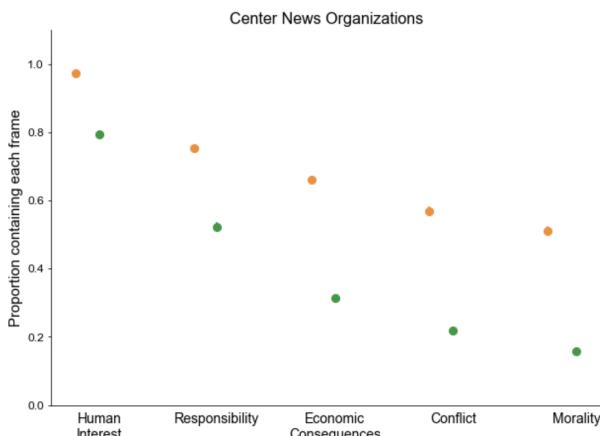
(b) Distribution of sentiment score of content, separated by the presence of headlines in posts for each news organization

Figure A5: Content sentiment changes by headline presence

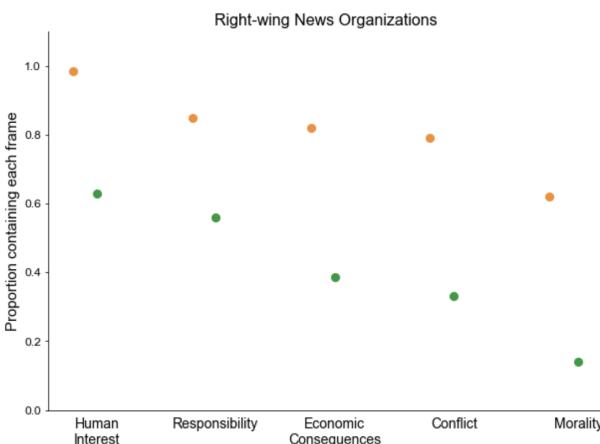
## I. Presence of news frames in articles and posts within each political leaning



- (a) Proportion of articles and posts containing each frame among left-wing news organizations (NBC News, the New York Times)



- (b) Proportion of articles and posts containing each frame among politically centered news organizations (BBC News, Reuters)



- (c) Proportion of articles and posts containing each frame among right-wing news organizations (Fox News, Washington Times)

Figure A6: News framing changes by political leaning