ABSTRACT

Title of Dissertation:	COMPUTATIONAL FRAMING ANALYSIS: PROPOSING AND APPLYING AN UNSUPERVISED ENTITY-CENTRIC SEMANTIC RELATIONS APPROACH
	Mohammad Ali, Ph.D. Candidate
Dissertation directed by:	Professor Naeemul Hassan, College of Information (iSchool), The University of Maryland, College Park.

This dissertation presents a novel computational approach to analyze how the news media construct frames around certain people or groups in their coverage. The approach centers on entity-centric emphasis frames, focusing on the language used to attribute key entities (e.g., shooters and victims in mass shooting incidents). The unsupervised method, named *Semantic Relations-based Unsupervised Framing Analysis* (SUFA), employs computational techniques to identify and analyze framing patterns based on semantic relations, moving beyond existing bag-of-words, co-occurrence, and frequency-based approaches. The dissertation includes three main projects, each building on the previous one to develop, apply, and advance this new approach.

Project 1 (Chapter 2) provides a critical review of existing computational methods used for supervised and unsupervised framing analysis. The review

highlights limitations in traditional unsupervised approaches, which primarily rely on bag-of-words, frequency, and word co-occurrence methods. These often fail to capture the contextual meaning and relationships between words. The survey article recommends integrating semantic relations into unsupervised framing analysis, proposing a more nuanced method for detecting frames.

Project 2 (Chapter 3) builds on the recommendations of Project 1 and introduces *Semantic Relations-based Unsupervised Framing Analysis* (SUFA) as a new computational framing approach to explore entity-centric emphasis frames. This chapter presents a mixed-method study consisting of qualitative textual analysis and computational analysis applied to 100 news reports (600 paragraphs) on the Uvalde school mass shooting from four major U.S. media outlets, *The New York Times*, *Cable News Network (CNN)*, *Wall Street Journal*, and *Fox News*. The qualitative analysis identifies how semantic relations contribute to frame construction, while the computational analysis employs natural language processing techniques, including dependency parsing, to extract and analyze entity-centric frames (e.g., shooter, victims, incident). The study outlines the strengths, limitations, and practical applications of SUFA, demonstrating its potential as a scalable and context-aware framing analysis approach.

Project 3 (Chapter 4) applies SUFA to a large-scale dataset of gun violence coverage in the United States and advances the methodological approach by automating the formation of frames from framing components. This study analyzes one month of news reports (N = 1334) from nine major U.S. news outlets covering the 2022 Uvalde elementary school mass shooting incident in Texas. Three of the outlets were selected from the left-centered bias category: *The New York Times* (n=227), *The Washington Post* (n=228), and *The USA Today* (n=128). Three were selected from the right-centered bias category: *The Wall Street Journal* (n=47), *The New York Post* (n=255), and *The Dallas Morning* News (n=155). And three outlets were chosen from the least-biased category: *The Hill* (n=227), *The Indianapolis Star* (n=39), and *The Des Moines Register* (n=28). The media outlets' biases were determined by scores provided by Media Bias/Fact Check (MBFC), a non-partisan and independent site that provides bias scores for media outlets.

The research under Project 3 goes beyond existing topic modeling approaches, which primarily rely on bag-of-words, co-occurrence, and frequency-based methods that often fail to capture semantic relationships between words. Instead, SUFA leverages advanced NLP techniques such as dependency parsing and coreference resolution to identify how words modify or relate to key entities (e.g., shooter, victims). Furthermore, large language models (LLMs) such as OpenAI's GPT-4o are incorporated to automate the clustering of framing components. This advances SUFA and improves its scalability for unsupervised frame detection. At the same time, it demonstrates how LLMs can be utilized in the coding of textual data and the clustering of framing components. This chapter provides how SUFA's entity-centric emphasis framing focus offers deeper insights into media narratives and bias, particularly in the framing of shooters and victims in a mass shooting incident.

The results revealed that at least six frames were attributed to the shooter, while nine frames were attributed to the victims in various ways. The right-centered news media deployed several frames, including action attribution, younger age, and certainty regarding allegations, significantly more than left-centered ones, to frame the shooter. As provided by the regression analysis, left-centered media are significantly more likely to employ personalized victim framing (our victims, your victims) and to emphasize older victims. In contrast, right-centered media more frequently utilize dehumanization framing. Also, there is a significant difference in how right-leaning and left-leaning news outlets employ individual framing components. These results highlight the ideological differences in how news media deploy framing components and frames when attributing to victims and the shooter.

Guided by framing and attribution theory, exploring frames offers theoretical insights into how media frames allocate responsibility for mass shootings through internal or external attributions. For instance, one prominent frame, action attribution, emphasizes the shooter's agency and responsibility, aligning with internal attribution and frequently utilized by right-leaning media.

This dissertation makes several important contributions to computational framing research. It advances the SUFA approach by applying it to a large-scale dataset and enhancing its methodological rigor by integrating semantic relations, dependency parsing, coreference resolution, and large language models (LLMs) for automated clustering. The study applies SUFA to mass shooting coverage, offering one of the first large-scale, unsupervised analyses of shooter and victim frames across media bias groups. It bridges attribution theory with computational methods, demonstrating how internal and external attributions of responsibility are reflected in media frames. Additionally, two annotated mass shooting datasets, validated through human and GPT comparison, provide a valuable resource for future research. This

dissertation also shows SUFA's interdisciplinary potential, showing how it can support framing analysis across fields such as social science and computer science.

In terms of implications, SUFA has the potential to serve as a powerful tool for real-time media monitoring and enhanced social media analytics by incorporating framing-based insights that extend beyond surface-level metrics like sentiment analysis and mention frequency. It can aid crisis communication by helping crisis practitioners comprehend how key entities are framed during crises, thus informing more effective response strategies. Computational or data-driven journalists can utilize SUFA to uncover and visualize media bias and framing trends. The general public, educators, and activists can also employ this approach to bolster their media literacy and, as expected, hold media outlets accountable for biased or misleading representations.

Overall, this dissertation develops, applies, and advances a computational framing analysis approach grounded in semantic relations to explore entity-centric emphasis frames. Although SUFA has some methodological limitations, including its primary reliance on textual framing components and its focus on entity-centric frames, its successful application across multiple datasets underscores its potential as a powerful tool for computational media analysis, bridging the gap between social science theories and advanced computational methods.

Keywords: Computational framing analysis, natural language processing, machine learning, dependency parsing, semantic relations, method, attribution theory, framing, gun violence, mass shootings, public health crisis, computational strategic communication, social media analytics, media monitoring

COMPUTATIONAL FRAMING ANALYSIS: PROPOSING AND APPLYING AN UNSUPERVISED ENTITY-CENTRIC SEMANTIC RELATIONS APPROACH

by

Mohammad Ali

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Naeemul Hassan, Chair
Professor Sahar Khamis (Dean's Representative)
Professor Ge Gao
Professor Sarah Oates
Professor Zubin Jelveh

Dedication

To my beloved parents, wife, and son, Abdullah Mohammad

Acknowledgements

I am deeply grateful to my advisor, Dr. Naeemul Hassan, for his unwavering support and guidance throughout this journey of my second doctoral study. His mentorship and intellectual leadership have profoundly shaped the trajectory and substance of this dissertation, ultimately enabling me to succeed.

My sincere thanks go to my dissertation committee members—Drs. Sahar Khamis, Ge Gao, Sarah Oates, and Zubin Jelveh—for their invaluable feedback, thoughtful suggestions, and ongoing support. Their collective wisdom not only helped me refine this project but also inspired me to push its boundaries. I am especially thankful to Dr. Sahar Khamis for serving as the Dean's Representative with such grace and care. I also thank Dr. Babak Fotouhi and Dr. Giovanni Ciampaglia for their guidance during my "integrative paper" (equivalent to comprehensive exams) and dissertation proposal phases, respectively.

I have had the privilege of being surrounded by a community of wonderful people, including my mentors, lab mates, and friends, who encouraged me through challenges and triumphs alike. Thank you for keeping me going. I am grateful to the College of Information (iSchool) for its institutional support, including the Doctoral Student Research Award (DSRA), and to the Computational Journalism Lab (CJLab) at the Philip Merrill College of Journalism, which became a second home during this journey. Special thanks again to Merrill College and Dr. Hassan for making that space so enriching.

This dissertation also benefited from various computational tools and platforms. I acknowledge the support of large language models (LLMs), such as

OpenAI’s GPT, Grammarly, Google Colab, and Python libraries, including spaCy, NeuralCoref, Scikit-learn, Pandas, Transformers, BERT, NumPy, and Matplotlib.

No words are sufficient to honor my family's unparalleled sacrifices and unconditional support. My late father, whose absence I feel deeply, remains a quiet source of strength amid the pain of losing him. I am forever indebted to my beloved mother, in-laws, and siblings for their enduring encouragement. With all my heart, I thank my beloved wife, Samsunnahar, and our son, Abdullah Mohammad. While caring for our son, her incomparable support has carried me through every step of this journey. Now over two years old, Abdullah has endured many moments of my absence, often with tears and patience, waiting alongside his mother. All of their love has been both my anchor and my inspiration. As a first-generation college graduate, first-time father, and second-time Ph.D. student, I complete this journey not only with pride and fulfillment but also with the acknowledgment that it has been sustained by the sacrifices of many.

Above all, I am profoundly thankful to Allah for everything.

Table of Contents

List of Tables

List of Figures

A. CHAPTER 1: Introduction

Framing analysis is a widely used research approach across disciplines from social science to computation research. Traditionally, scholars use qualitative textual analysis and quantitative approaches for understanding frames in a manual way. However, the proliferation of digital content through the internet and social media generates high volumes of data that cannot be analyzed manually. To address the problem, scholars have started using various computational tools aiming to explore frames in large datasets.

Existing computational framing analysis approaches are divided into two: supervised and unsupervised. Supervised framing analysis requires prior manual labor to label a small sample of a dataset for developing a training dataset that can be used for training the machine and developing a supervised model for analyzing frames. For instance, Liu et al. (2019) manually annotated a gun violence dataset and then used it to build classifiers with BERT, a deep learning model for natural language understanding and processing. On the other hand, the unsupervised framing analysis generally does not need such a prior labeled dataset and instead inductively explores frames in a dataset (Kotsiantis et al., 2007). Computational tools utilized existing unsupervised framing analysis approaches include topic modeling (e.g., DiMaggio et al., 2013), structural topic modeling (e.g., Roberts et al., 2014), hierarchical topic modeling (Nguyen et al., 2015), and a combination of topic modeling and network analysis (e.g., Walter & Ophir, 2019).

This dissertation focusses on the part of unsupervised framing analysis as an unsupervised approach is replicable across datasets and domains. One of the limitations of the existing unsupervised framing analysis is that these unsupervised approaches are mainly based on the idea of co-occurrence and frequency of words in a dataset, which resulted in the exploration of topics instead of further in-depth insights needed for understanding frames.

The main purpose of this dissertation is to offer a novel unsupervised computational framing analysis approach and test its performance by applying it to a large dataset. Secondly, this project intends to explore the phenomena of gun violence, specifically mass shootings in the United States. There is a pressing need to explore the issue, especially in the context of the U.S. and beyond.

Following the purpose mentioned above, this dissertation includes three interrelated research projects, as elaborated below.

Chapter 2: The first part of this dissertation, Chapter 2, conducts a survey titled *"A Survey of Computational Framing Analysis"* that reviews existing computational framing analyses and puts them together. The survey is conducted in the background where computational framing analysis approaches are scattered. Despite the scholarship growing, there is a scarcity of a comprehensive understanding and resources of computational framing analysis methods. The survey concludes with a few recommendations, including the utilization of semantic relations among words in exploring frames. Such an approach based on semantic relations remains ignored in this area of computational framing analysis. The survey findings are expected to help scholars and journalists gain a deeper understanding of existing approaches and resources so that they can further proceed on advancing the computational tools for framing analysis. This is a completed work and was published in the *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (NLP)*, a premier platform for NLP research.

Chapter 3: Building on the survey elaborated in Chapter 2, the second part of this dissertation in Chapter 3, titled *"Semantic-based Unsupervised Framing Analysis (SUFA): A Novel Approach for Computational Framing Analysis,"* is a mixed-method approach that conducts two follow-up studies. In study 1, we used a qualitative textual analysis as a case study to investigate the potential of using semantic relations as a way of exploring frames beyond the bag-of-word approaches used in the existing unsupervised framing analysis approaches, like topic modeling (e.g., DiMaggio et al., 2013). To do that, we inductively analyze a sample of news reports published by four major U.S. news media outlets relating to the 2022 Uvalde school mass shooting in Texas. Specifically, we look at how news media use various modifying words, such as adjectives and adverbs, in semantic relations with an entity (e.g., a shooter) aiming to frame the entity.

As evidenced in study 1 regarding the potential of harnessing the semantic relations among words for understanding frames of at least an entity, this second part of the dissertation conducts study 2, a computational framing analysis. In study 2, we aim to see if we can capture the semantic relations automatically that we did in study 1 manually to explore how the news media outlets frame an entity. To compare the results and understand the performance of the computational tools, we analyzed the same dataset of study 1 in this part. Here, we utilized various tools of natural language processing (NLP), including coreference resolution, dependency parsing, clustering, and some manual input. With the analysis, study 2 results provide a set of frames similar to that of study 1, meaning that the computational approach based on semantic relations automatically identifies some frames that we initially did manually. Overall, this mixed-method research comprising Study 1 and Study 2 offers a novel computational approach, namely *"Semantic Relations-based Framing Analysis"* (SUFA), that can be used for computationally analyzing emphasis frames of certain entities in a large dataset. The completed part of this project has been presented as the *Top Method Paper* at the Communication Theory and Methodology Division of the Association for Education in Journalism and Mass Communication (AEJMC) 2023, one of the top conferences in the area of communication and journalism.

Developing on the above-completed SUFA project, the following phases are proposed to test further and explore a few computational tools and steps aimed at advancing the SUFA model, especially its analysis part. The computational tools and steps include dependency parsing, coreference resolution, clustering, and utilization of a large language model (LLM). These proposed works are elaborated on in Chapter 3.

Chapter 4: The last research project of this dissertation is included in Chapter 4, which applies the SUFA model in a large dataset in the context of gun violence, aiming to see how various entities, such as a mass shooter and victims, are framed using the SUFA model.

Finally, this dissertation provides a concluding summary of the dissertation's contributions. Overall, this dissertation addresses a widely used and classic research approach of framing analysis using computational tools, including reviewing the approaches, developing a novel computational framing analysis, and applying and evaluating the newly proposed method. Key contributions from this dissertation to the literature include a review and compilation of existing computational framing analysis approaches and resources. Second, the conceptualization of a framing component as a pair of two words in certain semantic relations (e.g., 18-year-old killer and teenage killer). Third, a novel computational approach combining dependency parsing is offered based on semantic relations for entity-centric emphasis frames. Fourth, frames in the media content about gun violence are explored utilizing a new methodological approach of NLP.

B. CHAPTER 2: A Survey of Computational Framing Analysis

Abstract

Framing analysis is predominantly qualitative and quantitative, examining a small dataset with manual coding. Easy access to digital data in the last two decades prompts scholars in both computation and social sciences to utilize various computational methods to explore frames in large-scale datasets. However, the growing scholarship lacks a comprehensive understanding and shared resources around computational framing analysis methods. Existing studies often employ varied computational definitions and approaches in a fragmented way, with limited cross-referencing or engagement with prior methodological developments. This makes it difficult to map the current landscape of computational framing analysis approaches and ultimately hinders methodological advancement in addressing this classic problem computationally. Aiming to address the gap, this article surveys existing computational framing analysis approaches and puts them together. The research is expected to help scholars and journalists gain a deeper understanding of how frames are being explored computationally, better equip them to analyze frames in large-scale datasets, and, finally, work on advancing methodological approaches.

Keywords: Computational framing analysis, survey, computational social science, natural language processing, machine learning

1. Introduction

Understanding how vaccination is framed in news media offers a compelling example of why framing matters, especially in addressing the public health challenge of vaccine hesitancy despite the strong scientific evidence of vaccine effectiveness (Sallam, 2021). As Entman (1993) noted, a frame influences "how [people] evaluate [a problem] and choose to act upon it" (p. 54). Similarly, exploring how other societal issues, such as gun violence, are framed can offer critical insights of frames, especially when analyzed at scale in this data-rich era.

Traditionally, researchers explore frames using qualitative and quantitative methods that require manual labor and can handle small amounts of data (D'angelo, 2018; Reese et al., 2001). Production of and easy access to large volumes of digital data in the last two decades have prompted scholars to harness the exploration of frames in such big data computationally (Card et al., 2015; Liu et al., 2019; Walter & Ophir, 2019; van Atteveldt & Peng, 2018). Prior studies proposed various computational methods, including topic modeling and network analysis. Topic modeling is an unsupervised machine learning technique that identifies clusters of related words, which are interpreted as "topics," while network analysis explores relationships between entities, such as words, concepts, or actors. We elaborated on these later in the analysis section.

As the scholarship grows, a scarcity has appeared regarding a comprehensive understanding and resources of computational framing analysis methods (Nicholls & Culpepper, 2021; Sanfilippo et al., 2008). Researchers might be confused by multiple approaches to this analysis, raising questions: how many computational framing analysis methods exist, and which one should they apply? To address the problem and help researchers with such questions, we survey existing computational framing analysis approaches and put the methods and relevant resources together. As such, the survey is guided by the following three research questions:

RQ1. What computational methods do researchers use to explore frames in large-scale datasets?

RQ2. How do researchers conceptualize a frame in computational framing analysis studies?

RQ3. How do researchers use computational methods in exploring frames?

Though both RQ1 and RQ3 address computational methods, RQ1 specifically focuses on identifying the types of methods used, while RQ3 examines how those methods are applied in practice. The primary contributions of this article are: a) it provides a comprehensive understanding and resources of existing computational framing analysis methods and puts them together for interested scholars to gain deeper knowledge and start building on that, and b) it adds new thoughts to the ongoing discussion on advancing the computational methods of framing analysis.

2. What is Frame or Framing?

This section provides a conceptual understanding of framing. A classic example of framing concerns a debate over whether to permit the Ku Klux Klan to hold a public rally. One news story with the headline "Ku Klux Klan Tests OSU's Commitment to Free Speech" reported the rally as a free speech issue, while another one with the headline "Possible Ku Klux Klan Rally Raises Safety Concerns" reported it as a disruption of public order. As reflected in the headlines, the two stories used different frames. People who read the free speech news story expressed higher tolerance toward the KKK's rally compared to those who read the public order news story (Nelson et al., 1997, p. 581). Figure 2.1 shows similar frames deployed in two news headlines on the 2022 Buffalo mass shooting.

Figure 2. 1. Framing components deployed in the headlines of two news reports published by The New York Times and The Guardian on the 2022 Buffalo mass shooting.

Scholars have not agreed upon any unified definition of framing (Hertog & McLeod, 2001; Van Dijk, 2016). However, a prominent definition, widely used in both traditional and computational framing studies, was provided by Entman (1993). He says:

> To frame is to select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described. (p. 52)

As per this definition, a frame is largely determined by its outcome effects, such as four functions: a) defining problems, b) diagnosing causes, c) making judgments, and d) suggesting remedies. The functions depend on how some selected aspects of "perceived" reality are made salient. In 2003, he defined it a bit differently: “Framing entails selecting and highlighting some facets of events or issues, and making connections among them so as to promote a particular interpretation, evaluation, and/or solution” (Entman, 2003, p. 417). This definition seems to have made a few shifts, such as from “causal interpretation” to “interpretation,” from “moral evaluation” to “evaluation,” and from “treatment recommendation” to “solution.” The salient aspects are also interconnected.

While approaching frames as cultural phenomena, Hertog and McLeod (2001) identified a frame as a cultural "[structure] of meaning that includes a set of core concepts and ideas," including “conflicts, metaphors, myths, and narratives” (p. 160). A frame has also been explained as “a central organizing idea. . . for making sense of relevant events, suggesting what is at issue” (Gamson & Modigliani, 1989, p. 3). Reese et al. (2001) defined a frame from the sociological perspective and focused on six aspects (italicize): “Frames are organizing principles that are socially shared and persistent over time, that work symbolically to meaningfully structure the social world" (p. 11). In a recent definition, D'angelo (2018) defined news framing as "how journalists, their sources, and audiences work within conditions that shape the messages they construct as well as the ways they understand and interpret these messages" (p. xxiv).

To describe a frame's aspect highlighting some selected facets of an issue or event, Fairhurst (2005) utilized an analogy that "choosing language to frame people's actions and events is like moving a telescope into position" (p. 125). The selected aspects are then coherently organized in a way to make an argument, which finally promotes a particular interpretation, evaluation, and solution. This organization of selected aspects could even be subtle, as framing also "refers to subtle alterations in the statement or presentation of judgment and choice problems" (Iyengar, 1994, p. 11). Another crucial aspect of framing is "to choose one particular meaning (or set of meanings) over another" (Fairhurst & Sarr, 1996, p. 3) that is also supported by Entman (1993), who says a frame "operates by selecting and highlighting some features of reality while omitting others" (p. 53).

2.1. Contexts in Framing

A frame is considered context-sensitive. It is shaped in four locations: i) communicator, ii) texts, iii) receiver, and iv) culture (Entman, 1993). The culture is the stock of commonly invoked frames and explained as (a part of) contexts. A news report's content is fully comprehensible when its contextual information is at the disposal of readers. They interpret a frame and its meaning following contextual information (Baden & D'Angelo, 2018; Tewksbury & Riles, 2018).

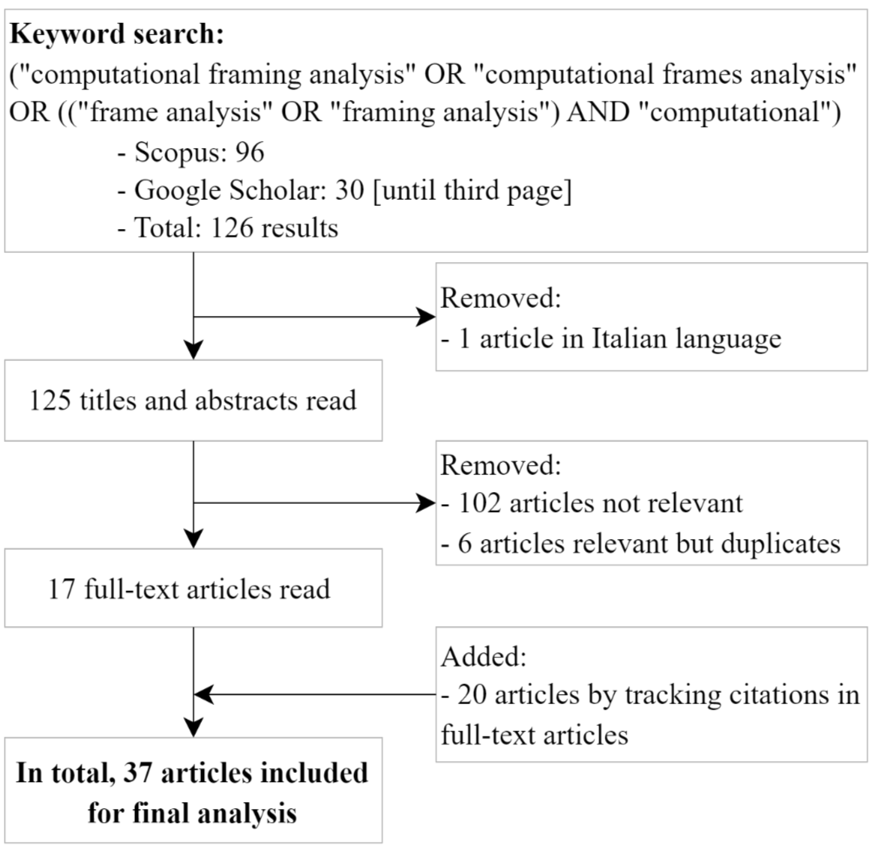
2.2. Framing Components

Framing components can be defined as tools that are used to make a piece of information more salient, which is, in other words, "making a piece of information more noticeable, meaningful, or memorable to audiences" (Entman, 1993, p. 53). While conceptualizing a frame, we accumulated framing components (see Table 2.1 in Appendix 2.1). To make the list concise and convenient, we combined similar devices and put them into four groups: a) content, b) action, c) context, and d) communicator. The devices or tools can be used to provide either higher or lower salience to selected aspects of reality. In some cases, multiple devices can be applied together as a new device. For example, jargon, metaphors, and contrast can together be used to develop a "story" (Fairhurst & Sarr, 1996).

3. Method

We utilized three ways to identify and select relevant articles for a comprehensive understanding of computational framing analysis methods. First, we searched on Scopus, an abstract and citation database of Elsevier, using relevant keywords: ("computational framing analysis" OR "computational frames analysis" OR (("frame analysis" OR "framing analysis") AND "computational"). It provides 95 articles in the English language. We manually read their abstracts and sorted out 13 articles relating to computational framing analysis. In the sorting process, we read the articles' method sections if needed to make the decision. The other 82 articles were excluded due to their irrelevance. The excluded articles were related to "frames" in other fields, such as building structures (e.g., 2D plane frames) and mechanical engineering. Second, we searched on Google Scholar using the exact keywords and included articles until the third page, as no relevant article was found on the third page. This gave us ten relevant articles. Six articles were common in both the Scopus and the Google Scholar searches, resulting in 17 unique articles from both sources. Third, while reading through the 17 selected articles, we tracked down 20 more relevant articles cited in some of those articles. The 20 articles did not appear in the Scopus and Google Scholar searches probably because of the different keywords and phrases used in their titles and abstracts.

Figure 2. 2. *Summary of the Paper Selection Method*



Finally, we got a total of 37 articles selected for this survey (see Figure 2.2). The articles involve journals and conferences in both computation and social science disciplines. Reading through the articles and their supplemental materials (e.g., coding schema guiding the annotation), if any, we utilized an inductive way to scrutinize various aspects, including a) framing conceptualization, b) functions of computational framing analysis approaches, and c) results and their interpretation. We reported available datasets, codes, and other relevant resources, if any.

4. Analysis

This section presents an analysis of the selected articles in two broad parts. The first part answers RQ1, and the second part answers RQ2 and RQ3. Table 2.2 in (Appendix 2.2) summarizes the articles, identified approaches, codebook, corpora, domains, and resources.

Codebook, Corpora, & Approaches (RQ1).

Analysis of the articles identified at least nine approaches and three major coding schema and annotated corpora for computational framing analysis. The approaches are in the categories of supervised, unsupervised, and mixed methods. A supervised method usually needs an annotated subset of data. Here, the model is first trained on a labeled dataset (training data) and then applied to a new similar dataset (test data) to classify or predict each instance (Kotsiantis et al., 2007). In contrast, an unsupervised method does not need any pre-annotated datasets. Instead, it explores all unlabeled data. The approaches, codebook, and corpora are elaborated below while answering RQ2 and RQ3.

Conceptualization & Functions (RQ2 & RQ3)

As a way of answering RQ2 and RQ3, we explore how researchers conceptualize frames and utilize computational methods in analyzing frames in each approach, codebook, and corpora.

4.1. Codebook & Corpora

4.1.1 Policy Frames Codebook

Boydstun et al. (2013) and Boydstun et al. (2014) proposed a codebook named "policy frames codebook" (PFC). The PFC consists of 14 categories of "frame dimensions" and an "other" category. The dimensions include "economic frames," "capacity and resources frames," "morality frames," etc. For example, a news report is labeled as an economic frame if it focuses on "the costs, benefits, or monetary/financial implications of the issue (to an individual, family, community, or to the economy as a whole)" (Boydstun et al., 2014, p. 6).

They developed the codebook through brainstorming and iteration of applying it to random texts. With the codebook, they deployed 3,033 coders to manually code three sets of articles on immigration, tobacco, and same-sex marriage. Using the labeled documents, they finally developed a logistic regression binary text classifier (i.e., present or absent) (Boydstun et al., 2013, 2014).

4.1.2. Media Frames Corpus

Using PFC, Card et al. (2015) offered a manually annotated corpus of news reports named "media frames corpus" (MFC). The news reports were collected from three domains: immigration, smoking, and same-sex marriage. The MFC was applied in other studies (e.g., Field et al., 2018). Card et al. (2015) annotated the three datasets based on PFC's 15 framing dimensions (Boydstun et al., 2013). The authors, however, did not apply the annotations to any new datasets. In 2016, they added four more categories—pro, neutral, anti, and irrelevant.

Conceptualization in PFC & MFC. Boydstun et al. (2013, 2014) conceptualized framing by resorting to the widely used framing definition of Entman (1993). Overall, they put "language" at the center of identifying and analyzing frames. PFC's development is motivated by three framing concepts: a) frame selection varies based on various situations, b) frames evolve over time, and c) frames spread across issues, geographic locations, and institutions or organizations. Card et al. (2015) also used Entman (1993)'s definition in conceptualizing frames. They focused on some framing elements that work coherently as a framing package.

Review. The authors conceptualized frames with existing framing definitions. However, framing aspects they mentioned (e.g., Entman, 1993) were not utilized in developing the 15 "framing dimensions." Considering the development process and broader definitions of each frame, the 15 dimensions seem to be more fit with "topics," not frames, as elaborated below. As per the framing theory, the categorization of these dimensions looks arbitrary and too broad to understand a frame's nuances. For example, a text is identified as an "economic frame" if it focuses on seemingly anything of the whole economy, such as "The costs, benefits, or monetary/financial implications of the issue (to an individual, family, community or to the economy as a whole)." Let's consider the Ku Klux Klan's example mentioned above. As per MFC's 15 dimensions, both KKK news reports could probably be identified as a “law and order, crime and justice frame” under the PFC. The PFC defines a text under this frame if the text is about “specific policies in practice and their enforcement, incentives, and implications. Includes stories about enforcement and interpretation of laws by individuals and law enforcement, breaking laws, loopholes, fines, sentencing, and punishment. Increases or reductions in crime.” This is too broad and over-generalized a framing definition, and it does not answer the “how” question. The dimensions, however, can be considered as topics. The MFC corpus inherited the same limitations as it was developed using the PFC codebook.

4.1.3 Gun Violence Frame Corpus (GVFC)

This article identified another annotated corpus named "Gun Violence Frame Corpus" (GVFC). It was applied in neural network-based models discussed later. In this dataset, the authors manually annotated 1,300 news headlines collected from 21 U.S. news media outlets. Using nine pre-defined codes drawn from the literature, multiple coders annotated the headlines. Finally, they used a BERT model, a state-of-the-art language model that understands the context of words in text, to build a frame prediction classifier. Its overall accuracy is 84.23. In other words, the model accurately predicts 84.23 frames out of 100.

Conceptualization. Liu et al. (2019) used Entman (1993)’s prominent definition to conceptualize framing. They highlighted various ways of constructing frames, such as word choice and labeling by journalists, “to promote a certain side” (p. 504). The authors also focused on generic versus issue-specific frames. In terms of manual codes, they applied a deductive approach— first defining some frames and then manually labeling news articles into those pre-defined frames.

Review. The article briefly conceptualized a frame and included the aspects of a widely-used framing definition (e.g., Entman, 1993). However, not all the framing codes in GVFC were defined following how the framing was conceptualized. For example, a code was defined in the category of politics "... as long as [a] news headline mentions a politician's name," which seems not aligned with the nuances of their conceptualizations.

4.2 Computational Approaches

4.2.1 Topic Modeling

Various prior studies utilized topic modeling (TM) to explore frames (e.g., DiMaggio et al., 2013).

Method. The TM algorithm discovers latent themes in a large collection of documents (Blei, 2012). A topic is a probability distribution over a fixed vocabulary (p. 78). The algorithm produces a number (k) of lists of words based on the words' higher probability of being in a list. Each list of words is considered to be a topic, and each topic has a different probability distribution. The latent Dirichlet allocation (LDA) topic model provides an assignment of each document to the topic(s). As a mixed-membership model, each of its documents may be assigned to multiple topics, considering that a document could have elements of multiple topics. DiMaggio et al. (2013) used the LDA topic modeling to explore frames. They view each topic as a frame, saying that a topic "includes terms that call attention to particular ways" (p. 593).

Conceptualization. In the study of DiMaggio et al. (2013), they conceptualized a frame as “a set of discursive cues (e.g., words, images, and narrative) that suggests a particular interpretation of a person, event, organization, practice, condition, or situation” (p. 593). They cited Gamson et al. (1992)’s definition that a frame is “a central organizing principle that holds together and gives coherence and meaning to a diverse array of symbols.” They considered each topic (aka a group of words) identified by the topic modeling algorithm as a frame.

Review. Here, the conceptualization of a frame looks consistent with the overall framing idea. However, the topic model’s output (i.e., lists of words) and their interpretation do not seem to align with the framing aspects. A list of words in the topic model comes without any connection among them due to its features (e.g., bag-of-words). The interpretation of each word list in DiMaggio et al. (2013) also indicates it as a theme or issue, not a frame. For example, they reported the results by utilizing words like “highlight,” “emphasize,” and “concerned with” (e.g., this topic highlights legislative actions). Framing nuances like a problem and causal interpretation could not be extracted here.

4.2.2 Structural Topic Modeling (STM)

Method. The STM model was also used to explore frames (e.g., Roberts et al., 2014). Compared to LDA topic modeling (Blei, 2012), STM allows including metadata or covariates in the model. With metadata (e.g., political ideology and time) added to the dataset and model, the STM allows researchers to interpret how the topics are associated with that metadata. For example, in terms of political ideology, such as conservatives and liberals, researchers might identify a topic as more aligned with conservatives and another topic with liberals. Metadata can also be used in predicting the prevalence of a topic by metadata (Gilardi et al., 2021; Nicholls & Culpepper, 2021).

In their study exploring topics in a corpus of newspaper texts, Gilardi et al. (2021) used some covariates, including time. Their results show how the topics are distributed over time across various states in the U.S. Since the authors followed DiMaggio et al. (2013)'s argument of considering a topic as a frame, their results' interpretation also focuses on themes or topics, instead of frames.

Conceptualization. Gilardi et al. (2021) conceptualized a frame with Gamson et al.'s (1992) definition that a frame can be understood as a "storyline or unfolding narrative about an issue" (p. 385). In terms of exploring frames by STM, Gilardi et al. (2021) relied on DiMaggio et al.'s (2013) argument that topics identified through TM can be viewed as frames.

Review. Like the topic modeling approach (Gilardi et al., 2021), the STM algorithm is also constrained by considering a topic as a frame. So, the STM contains similar limitations in terms of framing analysis. Compared to topic modeling, the STM offers additional insights into the topics or themes through the analysis of covariates. Both methods are based on the bag-of-words idea, indicating the lack of semantic contextualization needed for exploring frames.

4.2.3 Hierarchical Topic Modeling

Method. Studies also used hierarchical topic modeling (HTM) to explore frames. Nguyen (2015) and Nguyen et al. (2015) introduced an HTM model named "Supervised Hierarchical Latent Dirichlet Allocation (SHLDA)" that aims to analyze frames in a large dataset. As the SHLDA works, each document in the corpus is associated with a continuous level of scores (e.g., conservative vs. liberal ideology). It produces a hierarchy of topics, where the first-level nodes are considered agendas and the second-level nodes as frames. The documents' scores help explain how the topics are framed concerning respective people's positions. Its document generative process combines the hierarchical LDA and hierarchical Dirichlet process (HDP). The authors applied it to three datasets and conducted qualitative and quantitative analyses to validate the models' agenda and frames.

Conceptualization. Nguyen (2015) also used Entman's (1993) framing definition to conceptualize a frame. However, unlike Gilardi et al. (2021), Nguyen distinguished between agenda and frame by treating a topic as an agenda, meaning the issue or subject being discussed (what is talked about). In contrast, a sub-topic was considered a second-level agenda or frame, potentially reflecting the particular perspective and emphasis applied to that topic (how it is talked about).

Review. As elaborated above, the SHLDA is one step ahead of topic modeling. However, a crucial incongruity remains in how they conceptualized a frame (e.g., sub-topics) and interpreted the results. Though there is a lack of unified framing definition, the idea of considering a sub-topic as a frame does not align with traditional framing conceptualization (Entman, 1993; McCombs et al., 1997; Ghanem, 1997). Like many prior framing studies, the SHLDA output might also be considered as simply topics and their relevant attributes, not frames. Moreover, Nguyen (2015)'s qualitative analysis to validate the output as frames is not systematically executed, and the presentation of its results does not illustrate any framing aspects (Entman, 1993)

4.2.4 Cluster Analysis

Method. The k-means clustering algorithm is another unsupervised approach used to explore frames. Burscher et al. (2016) conducted two k-means clusterings in a dataset. One includes all words, and another includes selected words (i.e., nouns, adjectives, and adverbs). After converting the texts from both groups into numerical representations using the TF-IDF method, which emphasizes important words by considering how often they appear in a document (term frequency or TF) and how rare they are across the entire corpus (inverse document frequency or IDF), they used *k*-means clustering to group similar documents together. As a centroid-based clustering approach works, a certain number of clusters (k) is specified in advance, and each cluster is represented by its center. They select the number of clusters (k) using the "elbow method." Each document is assigned to a cluster based on its relatively closer distance to that cluster center (Burscher et al., 2016). Unlike topic modeling, k-means clustering is a single-membership approach where each document generally belongs to one cluster.

Conceptualization. Burscher et al. (2016) conceptualized a frame in terms of "word frequencies" and mentioned words as highly reliable and less biased in producing frames. They "used word frequencies as features [of a frame] in [their] cluster analyses" (p. 533). They utilized the traditional framing definition partially (e.g., the presence or absence of certain keywords and stock phrases) (Entman, 1993).

Review. As Burscher et al. (2016) conceptualized and interpreted frames in terms of word frequencies and co-occurrences, the framing components listed in Table 2.1 (Appendix 2.1) suggest that word(s) are simply one of the many devices to construct a frame. They utilized such conceptualization that does not help explore frames despite their acknowledgment that “based on plain word features, a cluster analysis cannot reveal complex semantic and logical relationships like causality” (Burscher et al., 2016, p. 541). As a single-membership approach, this method is also against one of the core framing ideas that a framing component may belong to multiple frames. The results were presented with words, including “refer to.” For example, “cluster B5 refers to nuclear power... in Iran” (p. 439). The results indicate these as a topic or issue. It does not indicate "how" the “nuclear” issue was discussed and evaluated as a problem. Both conceptualization and output seem to illustrate certain topics, not frames.

4.2.5 Neural Network Model

Method. Some studies utilized the neural network approach to build frame-identifying classifiers and analyzed frames in various text documents (e.g., news reports and tweets). Neural networks are machine learning models designed to detect complex patterns in data, by mimicking the way the human brain processes information. Mainly, two annotated datasets namely, MFC and GVFC, were used in building these models.

MFC was utilized in a number of such studies, including probabilistic soft logic (PSL) (Johnson et al., 2017), LSTM neural network (Naderi and Hirst, 2017), recursive neural network (Ji and Smith, 2017), and transformer-based language models such as BERT and RoBERTa (Khanehzar et al., 2019; Cabot et al., 2020; Mendelsohn et al., 2021). Some studies used MFC's annotated news reports partially and some used the full corpus.

Liu et al. (2019) manually annotated the GVFC dataset and used it to build a classifier using BERT. This technique was later applied in other studies (e.g., Akyürek et al., 2020; Tourni et al., 2021; Bhatia et al., 2021).

Conceptualization. As mentioned above, Liu et al. (2019) used traditional framing definitions (e.g., Entman, 1993) while conceptualizing a frame. The studies applying MFC in building a neural network-based classifier also conceptualize it by drawing works from prior studies in both social and computational science.

Review. In terms of the approach, both groups of studies seem to have applied the state-of-the-art pre-trained models based on transfer learning, which looks promising for advancing computational framing analysis. However, the quality of the annotated training dataset appears not up to the mark, which is reflected in the lack of results interpretation in those studies. As reviewed above, the MFC dataset seems more about categorizing a text into broad topics (e.g., "economic frames"), rather than frames. The way these topics are defined tends to emphasize "what" is simply discussed, rather than "how" it is discussed, which is a crucial distinction for identifying a frame. The subsequent studies applying MFC dataset also did not adequately justify MFC's 15 dimensions as frames. Their results mainly focused on the accuracy of the model built on MFC training dataset, but not whether the results provide framing nuances.

Compared to MFC, GVFC's annotations look more coherent but still lack in capturing framing nuances, as mentioned above in sub-section 4.1.3. For example, based on GVFC's "politics" code, Liu et al. (2019) interpreted its result, saying, "it appears that news media of all types have largely politicized the gun violence issue right after each major mass shooting" (p. 511). Here, the politicization result and its interpretation do not align with how the code is defined. The results might indicate the texts "discussed," "a politician," or politics, which is a simple topic or an issue, not any major framing element like problem definition and its coherent argument.

4.2.6 Parsing Semantic Relations

Another line of computational framing analysis relates to the exploration of semantic relations, going beyond the bag-of-words model.

Method. Sturdza et al. (2018) operationalized Entman (1993)'s four framing elements as their semantic relations in texts. This approach proposed the utilization of a rule-based system that uses existing computational software, such as TurboParser, and implicature rules. Using the parser, the author proposed identifying syntactic structures in texts and then using a set of rules to transform the syntactic structure into semantic networks. The networks determine the semantic roles of each word (e.g., actors, events) through a set of sentiment analysis implicature rules using a sentiment lexicon.

On the other hand, Ziems and Yang (2021) computationally parsed various attributes (e.g., race) of police shooting victims in news reports and explored how differently they are portrayed in news media. They called it "entity-centric framing." A recent study by Yu (2022) looked at iterative adverbs (e.g., again) in the political discourse, considering that the adverbs evoke different attitudinal subtexts. After extracting sentences with relevant adverbs, the author grouped the sentences through k-means clustering and identified the most representative keywords in each cluster by a keyword mining tool.

Conceptualization. In conceptualizing a frame, turdza et al. (2018) relied on four framing elements of Entman (1993, p. 52). However, two other studies lack adequate conceptualization of framing. For instance, Ziems and Yang (2021) mainly explored “entity-centric” frames but did not elaborate on it from existing literature.

Review. Compared to the topic modeling method, this approach looks innovative in terms of understanding semantic relations between words and phrases. However, the idea seems not adequately exploited in understanding the nuances of frames. For example, Sturdza et al. (2018) did not apply the operationalization in a practical dataset. Results of Ziems and Yang (2021) reported frequency and correlations, while Yu (2022)’s results ended up with clustering and keywords, instead of exploring the coherent argument and relations among various framing components. However, by its design, the semantic relations approach holds strong potential for advancing the computational methods in framing analysis. For example, a specific semantic relation, such as an adjective modifier (e.g., “deadliest” is modifying “shooting”), can offer insight into how words are meaningfully connected. This differs from topic modeling or network analysis, where word associations are typically based on co-occurrence rather than explicit linguistic relationships.

4.2.7 Frequency-based Model

Method. This computational model proposed using QDA Miner and its affiliated WordStat program to extract words, and phrases, and examine their repetitions across the corpus (Kang & Yang, 2022). This approach relied more on content analysis and keyword frequency, rather than incorporating machine learning techniques. In this model, Sanderink (2020) proposed little changes, which is to first determine certain frames (e.g., energy security) by reviewing prior scholarship. Researchers then prepare a codebook using QDA Miner. The codebook comprises words, phrases, and rules that capture various elements relating to each of the pre-determined frames. Finally, WordStat was used to calculate the frequency of words and phrases relating to each frame.

Conceptualization. Scholars in this approach defined a frame in terms of word recurrence in a document. They also highlighted the ways of editing, interpreting, organizing, and presenting information for particular news content to be framed. They compared a frame with a theme.

Review. The frame was not appropriately conceptualized here, as per the existing framing definitions (e.g., Entman, 1993). The consideration of only the frequency of words does not compromise the coherent meanings of frames.

4.2.8 FrameAxis

Method. FramAxis model explores "microframes," which is operationalized as a pair of antonyms, such as legal versus illegal and fast versus slow. The antonyms are obtained from WordNet. Then, the authors compute the bias of each microframe (average contribution of all words in a document to the microframe) and the intensity of each microframe (how strongly it is presented in documents). The microframes are analyzed along with the agent-object-action patterns identified by the semantic role labeling (SRL) model in the corpus.

Conceptualization. A frame in this approach was conceptualized utilizing features of existing definitions. For example, they highlighted presenting some selected aspects of an issue and making them more salient, which aims to promote certain values, interpretations, or solutions.

Review. Though the framing conceptualization is derived from prominent framing definitions, the core aspect of FrameAxis is the pair of antonyms, which again limits the coherent argument, problem definition, and other framing elements.

4.2.9 Analysis of Topic Model Networks

Walter and Ophir (2019) proposed this mixed method approach, "Analysis of Topic Model Networks" (ANTMN), that combines topic modeling and semantic network analysis. It was applied in other studies (e.g., Ophir et al., 2021).

Method. ANTMN includes three steps. First, the authors apply LDA topic modeling (Blei, 2012) to the dataset. They label each topic by qualitatively examining three types of information: a) words with the highest loading over each topic, b) prevalent and exclusive words in each topic, and c) full documents that are the most representative of each topic. Second, ANTMN creates a semantic network, where the topics serve as nodes, and topics' similarity relationships serve as edges. The relationship is calculated based on the topics' cooccurrence in the documents. The output provides a fully connected, undirected, and weighted network. Finally, a community detection algorithm was used to cluster the topics into various communities in the network based on the topics' prevalence in similar documents (Walter and Ophir, 2019).

Conceptualization. As the authors noted, ANTMN can analyze emphasis frames (e.g., highlighting one side), not equivalency frames (e.g., gain vs. loss issue). They conceptualize a frame as "a communit[y] in a network of topics" (p. 248), based on linguistic patterns. Borrowing van Atteveldt and Peng (2018)'s idea of arranging various framing components around an overarching idea (e.g., a cluster of relevant framing components), they consider each topic in topic modeling as a framing component. The cluster of topics was named as a frame in ANTMN. They embraced the patterns of a frame that "repeatedly invokes the same objects and traits, using identical or synonymous words and symbols in a series of similar communications that are concentrated in time" (Entman et al., 2009, p. 177).

Review. A few things seem to have restricted ANTMN as a framing analysis model. As per the framing conceptualization, the topics (aka framing components) under each network community need to be coherently connected with each other to render a coherent framing argument. The authors did not explain how the devices are coherently interconnected. This lack is reflected in the interpretation of the results. For instance, they reported a framing result, saying that "the largest community on the right consisted of topics about the cultural and economic consequences…. Articles dominated by these topics portrayed the impact of diseases on the economy at large…. (Walter and Ophir, 2019, p. 259). Here, the authors mentioned topics' names and what these topics portray with words like "consists of" and "portrayed." The results did not provide a coherent argument of the problem or how one aspect is interconnected with another. Though the output demonstrated some topics, the authors' claim of the communities as frames is not supported with adequate evidence.

Despite the authors' claim of this method as unsupervised, manual human labor is still needed in at least two places: a) an examination of words and documents to label topics, and b) an interpretation of findings. However, no systematic method was provided for executing the manual analysis.

5. Discussion and Conclusion

This article synthesizes and critically evaluates existing computational methods, coding schemas, and annotated corpora in framing analysis. By analyzing and drawing insights from 37 empirical studies, the article presents a comprehensive overview of the current state of computational framing analysis approaches. At the same time, it identifies critical gaps and opportunities for future exploration. This consolidated resource is intended to support both new and experienced framing scholars by offering a detailed understanding of existing approaches and laying the groundwork for more integrated analyses and the continued advancement of computational framing analysis approaches in large-scale datasets.

5.1. Algorithmic Functions

As demonstrated above, most algorithms used in computational framing analysis were not originally built for this purpose. For example, LDA topic modeling is basically built to find broader themes in a large corpus (Blei, 2012). The works of Liu et al. (2019) and Walter and Ophir (2019), however, seem to be innovative in terms of their efforts to build a new or modified method to explore comparatively more nuances of frames (Nicholls & Culpepper, 2021). As state-of-the-art models, neural networks appeared promising, but appropriate training datasets need to be developed and used for that.

5.2. Conceptualization of Frames

Though the computational methods mostly conceptualized a frame with prominent definitions (e.g., the definition of Entman, 1993), some of the methods embraced framing aspects partially. Some studies ended up operationalizing a frame in a way that is not supported by the core framing aspects. For instance, Boydstun et al. (2013, 2014) include its main aspects in developing PFC, which defined the 15 dimensions as “topics” in the name of frames. Nguyen (2015) simply equated a frame with second-level agendas or sub-topics without adequate conceptual support. Though Liu et al. (2019) and Walter and Ophir (2019) provided relatively stronger conceptualization, their results suggest that Liu et al. (2019)’s coding schema and Walter and Ophir (2019)’s network communities still lack in providing coherent definition and causal interpretation arguments.

5.3. Interpretation of Results

Even if some studies conceptualized frames in a relatively comprehensive way, their results presentation and interpretation rarely went above describing relevant topics and themes, not frames, as their results lack illustrating the coherent problem, causal evaluations, or potential recommendations. An example mentioned under ANTMN above demonstrated such evidence. Similar gaps in terms of framing conceptualization and presentation of results and interpretations remain in other approaches as well (e.g., topic modeling and cluster analysis). This suggests that beyond the computational challenges of detecting frames, a more fundamental issue lies in how frames are conceptualized and operationalized for computational identification. Limitations in framing conceptualization, particularly within computational studies, make it difficult to translate the concept of a frame into a form that is both computationally tractable and theoretically meaningful.

5.4. Use of Framing Components

The bag-of-words approach automatically excludes from analysis many potential framing components listed in Table 2.1 (Appendix 2.1). The approaches examined in this article mostly utilize only one framing components (i.e., words). Considering the fact that framing analysis is a comprehensive approach involving multiple theoretical and practical aspects (D'angelo, 2018; Golan, 2021), even the qualitative framing analysis through manual labor is challenging work. From that perspective, computational approaches are in the nascent stage in addressing this social science problem of framing analysis. So, the scholarship needs better computational methods and tools that might explore frames as close as possible. For example, computational approaches might want to retrieve the problem definition and causal interpretation by including more framing components (see Table 2.1 in Appendix 2.1) by going beyond the analysis of "words" in future studies. At the same time, advancing the scholarship of computational framing analysis approaches necessitates not only clearer conceptualization of specific framing components, but also careful attention to how these framing components are operationalized for computational detection.

Overall, this survey article contributed to the literature on computational framing analysis in several ways. As the first survey paper, it put together existing computational framing analysis methods and resources in one place, which can benefit future scholars as at least a source of gaining more comprehensive knowledge on computational framing analysis approaches. With this knowledge, they can start further exploring frames in big data and advancing computational framing analysis methods. This article also contributed to the ongoing discussion and scholarly efforts on further improving the computational tools in framing analysis.

5.5. Open Questions

The analysis and discussion offer at least three open questions to be discussed and addressed in future studies: a) How can a computational approach capture all relevant semantic relations, going beyond just words, for better exploration of frames, 2) How can the semantic relations in one text document be connected with or informed by that of other documents for a broader understanding of frames across multiple documents, c) Given the role of many framing components, not only words, in constructing frames (see Table 2.1 in Appendix 2.1), how can we develop a computational model that captures salience deployed through other framing components including sentences, omit, metaphors, size and placement of texts, culture, emotion, sources, catchphrase, exemplars, visual content, etc. A crucial part of framing analysis is to capture "how" a text is presented. Entman (1993)'s definition talks about "perceived reality" that also aligns with people's cognitive thoughts. In texts, the "perceived reality" is usually dissected between what is discussed and how it is framed. Though the "what" part is generally apparent, the main issue is to analyze the "how." In NLP, it appears difficult to automatically distinguish between the “what” and the “how.” So, the framing analysis task in NLP is more complicated than for human analysts.

5.6. Limitations

Selecting articles for this survey was a challenging task as the words “frame” and “framing” are used in studies of other disciplines (e.g., engineering). This prompted us to exploit multiple ways (e.g., Google Scholar and Scopus) to collect relevant articles as comprehensively as possible. Articles not matching the keyword searches might have been left out. So, the list might have some articles missing due to the search constraints. We excluded non-English articles.

Regarding analysis, we mainly focused on methodological design and quality in terms of capturing and examining frames and framing components. We did not focus and report on the accuracy of the models’ performance. For example, we emphasized the quality of the training dataset (e.g., MFC) to explore frames, instead of the models’ accuracies. As this survey article is conducted from a qualitative perspective, our results are constrained by quantitative insights (e.g., the frequency or percentage of applying particular methods in prior studies).

We identified a range of domains and contexts examined in the articles analyzed in this research, including tobacco, immigration, congressional debate, gun violence, and nuclear power, among others (see Table 2.2 in Appendix 2.2). However, given this article’s primary focus on methodological exploration, we limited our critical analysis and did not further examine domain-specific applications of computational framing. Future studies are encouraged to investigate how computational methods vary across different topical contexts and whether certain domains demand unique methodological considerations.

Disclaimer. This article has earlier been presented at the *Empirical Methods in Natural Language Processing Conference 2022*, a premier platform for NLP work, and published in its proceedings in 2022. The published version's citation:

Ali, M., & Hassan, N. (2022, December). A survey of computational framing analysis approaches. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (pp. 9335-9348). https://aclanthology.org/2022.emnlp-main.633/

C. CHAPTER 3: Semantic-based Unsupervised Framing Analysis (SUFA): A Novel Approach for Computational Framing Analysis

Abstract

This research offers a novel approach for computational framing analysis called *Semantic Relations-based Unsupervised Framing Analysis* (SUFA). The approach uses semantic relations and dependency parsing algorithms to analyze frames in news reports. This mixed-method approach is derived from two studies—qualitative and computational— using a dataset on gun violence, demonstrating its potential for analyzing entity-centric emphasis frames. The strengths, limitations, and application procedure of SUFA are discussed. This novel approach can be applied in both social and computation sciences.

Keywords: Computational framing analysis, semantic relations, dependency parsing, natural language processing, communication method, computational method

1. Introduction

Frames are predominantly explored using qualitative methods (e.g., Morin, 2016) and quantitative methods (e.g., McKeever et al., 2022) through manual labor and analysis of small datasets. The recent proliferation of online news reports and social media posts has resulted in the generation of a vast amount of digital data that is difficult to analyze manually. To overcome this challenge, scholars have started using various computational methods, broadly divided into two parts: supervised and unsupervised (Ali & Hassan, 2022). The supervised methods require pre-determined labels and substantial human labor, while the unsupervised methods that this current research focuses on need little human effort and are applicable across domains.

Existing unsupervised methods (e.g., topic modeling) in framing analysis mainly rely on the frequency and co-occurrence of words, leading to the exploration of topics instead of deeper framing insights (Nicholls & Culpepper, 2021; Ali & Hassan, 2022; Entman, 1993). An improved unsupervised computational solution to this longstanding communication challenge is becoming essential in this era of big data. Scholars (e.g., Ali & Hassan, 2022) advocate for methods to capture semantic relationships between words, moving beyond the traditional bag-of-words approach to enhance the methodological framework. In response to these calls, this article examines semantic relationships between words, presenting a novel unsupervised approach for computational framing analysis based on dependency parsing, a natural language processing (NLP) technique largely overlooked in framing analysis.

This mixed-method article involves two studies. Study 1 employs a qualitative textual analysis to inductively examine a sample of news reports published by four major U.S. news media outlets on the 2022 Uvalde school mass shooting in Texas, as a case study. While the political impasse and public debate continue over gun violence, it is important to understand how news media outlets frame the issue, as media framing determines how people “choose to act upon [the problem]” (Entman, 1993, p. 54). Study 1 examines how individual words, such as adjectives and adverbs, convey different meanings related to the shooter, victims, and the shooting event.

This helps us understand how these words and their semantic relationships work together to construct frames. Study 2 employs the computational technique of dependency parsing to analyze the same dataset. Specifically, we investigate dependency parsing, along with word embedding, *k*-means clustering, and manual input, establishing this method as a viable approach for capturing semantic relationships and analyzing the entity-centric emphasis frames.

Integrating qualitative and quantitative approaches in this project provides complementary strengths essential for developing the methodological approach. The qualitative analysis in Study 1 offers interpretive depth by first manually uncovering whether and how specific words and their semantic relationships contribute to frame construction in natural language. This inductive insight helps ground the methodological design. Through quantitative computational techniques, Study 2 validates and extends this insight from Study 1 by systematically extracting these patterns computationally. Together, the two studies demonstrate that semantic structures, captured through dependency parsing, can reliably identify emphasis frames, laying the foundation for a scalable, unsupervised computational framing analysis model.

The outcomes of both studies are discussed. Importantly, this mixed-method project solidified and proposed the semantic relations-based approach for framing analysis, named *"Semantic Relations-based Unsupervised Framing Analysis."* The step-by-step procedure for applying this approach, along with its strengths, limitations, and future research directions, is also discussed.

2. Literature Review

1.1. Framing

Scholars have not reached a consensus on a unified definition of framing (*Goffman, 1974;* Hertog & McLeod, 2001). However, one of the most widely cited definitions in framing studies comes from Entman (1993), who posits:

> To frame is to *select some aspects of a perceived reality and make them more salient in a communicating text in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation* for the item described. (p. 52)

In the news media context, a frame is “a central organizing idea” (Tankard et al., 1991), and it “denotes how journalists, their sources, and audiences work within conditions that shape the messages they construct as well as the ways they understand and interpret these messages” (D'Angelo, 2018, p. xxiv). Going beyond the idea of a simple topic, news framing is “like moving a telescope into position” (Fairhurst, 2005, p. 125), where selected aspects are coherently organized in a way that makes an argument, promoting a particular interpretation, evaluation, and solution (Fairhurst, 2005). Importantly*, a frame “operates by selecting and highlighting some features of reality while omitting others” (Entman, 1993, p. 53). Echoing with this,* Fairhurst and Sarr (1996) notes that a frame is *“to choose one particular meaning (or set of meanings) over another” (p. 3).*

1.1.1. Emphasis vs. Equivalency Framing

The concept of framing revolves around two broad competing aspects: emphasis framing and equivalency framing. Equivalency framing involves presenting two or more alternatives with logically equivalent phrases (e.g., loss versus gain) (Kahneman & Tversky, 1984; Levin et al., 1998). In contrast, emphasis framing refers to the act of repeatedly highlighting or associating certain pieces of information about an issue or topic, while omitting other relevant aspects (D'Angelo, 2017). This article focuses on analyzing emphasis framing with the newly proposed computational approach.

1.2. Words in Constructing Frames

Scholars have long identified words and phrases that construct frames. Prior studies have shown that using certain words helps identify frames (Entman, 1993; Fairhurst & Sarr, 1996; Gamson & Modigliani, 1989; Hertog & McLeod, 2001). For example, "the use of *baby* versus *fetus* signals a very different approach to the topic of abortion" (Hertog & McLeod, 2001, p. 150). Prior framing studies looked at various parts of speeches, including verbs, adverbs, and adjectives, which enhances researchers' ability to identify frame boundaries and relationships (Hertog & McLeod, 2001). The frequent use of verbs such as "falsifying," "forging," and "manipulating" was found to have been utilized in news reports to frame scientists (Boesman & Van Gorp, 2018). News reporters also use various verbs of attribution (e.g., accused, charged, blamed, attacked) to create worth for one person while devaluing another (Dickerson, 2001).

1.3. Conceptualization and Operationalization of the Framing Component

Prior studies provide evidence for using words in constructing frames (Hertog & McLeod, 2001; Miller et al., 1998). When a particular word is selected or coded as part of a frame, this word directly or indirectly operates in relation to other words to express the intended framing meaning. In other words, framing meanings are often produced not by isolated words but through their associative use with surrounding words, particularly when an entity is modified by adjectives, adverbs, or verbs.

For example, Bantimaroudis and Ban (2001) reported how Somali leaders were framed by U.S. news media through the repeated use of the term “warlords” in contrast to their opposition, the United Nations forces. They interpreted the frame by exposing how extensively the word “warlords” was used in the news media coverage. This current research argues that the word “warlords” alone does not sufficiently convey a practical meaning for understanding the frame about Somali leaders. Instead, we better understand the intended frame when the word “warlords” is seen as an adjective modifier to its noun, “Mohammed Siad Barre,” forming a phrase like “warlord Barre.” In this context, the framing component emerges from a meaningful semantic pair, a modifying word and its head noun, which together construct the framing meaning.

Crucially, this pair of words is bound by a meaningful semantic relation. For example, in the dependency parsing output of natural language processing, the noun “Barre” and its modifier “warlord” are linked by an adjectival modifier relation (known as “amod”). Based on this linguistic structure, this current research conceptualizes a framing component as *“a pair of words connected by a meaningful* *semantic relation."* The modifying word may belong to various parts of speech, such as adjectives (e.g., young shooter), verbs (e.g., shooter kills), or even participles and modal verbs (e.g., shooter accused of [killing]).

In qualitative textual and quantitative content analyses that rely on manual labor, scholars might code the keyword "warlord," keeping other parts (e.g., noun and semantic relation) in mind, and consider its semantic context during interpretation to explore meaningful insights. However, for computational analysis, capturing such semantic structures explicitly becomes essential for scaling framing analysis to large datasets.

To this end, this current research operationalizes a framing component as a pair of words connected by a meaningful semantic relation, specifically identified using dependency parsing techniques. For instance, adjective-noun (amod) or verb-subject (nsubj) relationships are used to detect modifier-entity structures, such as "teenage gunman" or "shooter kills." These semantic relations are computationally extracted from the dependency tree of each sentence. By identifying the framing components in the semantic relations-based structure, this approach allows for systematic extraction of entity-modifier pairs in large datasets, ensuring both consistency and scalability. This operationalization is particularly well-suited for analyzing entity-centric frames, as it captures how individuals, organizations, or groups are framed through specific modifying words in large datasets.

1.4. Framing Analysis with Computational Approaches

Traditionally, researchers utilize qualitative and quantitative methods to analyze frames, relying on manual labor and small amounts of data (D'angelo, 2018;

Reese et al., 2001). To tackle the challenge of analyzing frames in large-scale datasets, scholars have begun using computational approaches—both supervised and unsupervised—in the last two decades (e.g., Card et al., 2015; Liu et al., 2019; Walter & Ophir, 2019; van Atteveldt & Peng, 2018).

Supervised. A supervised approach needs pre-labeled datasets. In this approach, a model is first trained on the labeled data and then applied to a new dataset to classify or predict each instance (Kotsiantis et al., 2007). Under the supervised framing analysis approach, Liu et al. (2019) proposed a deep learning-based model developed with manual codes of headlines of news reports relating to gun violence.

Unsupervised. An unsupervised approach does not require any pre-annotated datasets. Instead, it inductively explores all unlabeled data (Kotsiantis et al., 2007). Existing unsupervised approaches used to analyze frames include topic modeling (DiMaggio et al., 2013), structural topic modeling (Gilardi et al., 2021), hierarchical topic modeling (Nguyen, 2015), cluster analysis (Burscher et al., 2016), frequency-based models (Sanderink, 2020), and FrameAxis (Kwak et al., 2021). Compared to supervised models, unsupervised ones demand less time and can be replicated across domains.

Semantic relations. Existing unsupervised computational approaches for framing analysis are mainly based on the ideas of frequency and co-occurrences of words, resulting in the identification of discussion topics or themes, instead of frames (Ali & Hassan, 2022). Such topics do not provide a coherent framing interpretation. As per the framing conceptualization (Entman, 1993; Reese et al., 2001), semantic relations among words are a key to going deeper into frames, compared to the current bag-of-words-based practices, such as topic modeling. This limitation calls for exploring an unsupervised technique to capture semantic relations among words for better identifying frames. This article intends to fill the gap by focusing on unsupervised methods of framing analysis.

Although a few studies attempted to address the task with semantic relations, their approaches are not sufficiently comprehensive or supervised from the data analysis perspective. For example, Sturdza (2018) describes an approach of operationalizing frames using a rule-based system with a software named TurboParser. However, the author did not execute it using a dataset, leaving its usefulness unclear. A recent study by Ziems and Yang (2021) proposes an NLP framework to understand the frames of an entity or issue (e.g., victims in police violence) with relevant attributes (e.g., age, gender, race). However, they pre-determined the attributes and then string-matched relevant tokens as a way of framing particular entities, which is also considered supervised.

Another study by van Atteveldt et al. (2013) presents a computational framing analysis method based on semantic relations. Their approach is also a kind of supervised task, as it first determines and labels particular frames and then identifies occurrences of each pre-determined frame in the dataset. Framing analysis scholars in recent studies (e.g., Nicholles et al., 2021; Ali & Hassan, 2022) call for exploring semantic relations for improved framing nuances.

Therefore, this research seeks to fill the gap by offering and advancing a semantic relations-based unsupervised approach for framing analysis through two studies—qualitative textual analysis and computational analysis. Both studies examine a sample of 100 news reports published by four major U.S. news media outlets on the 2022 Texas school mass shooting.

1.5. Gun Violence and Framing Analysis

Gun violence is a widely studied area in the U.S., as the mass shooting problem has been on the rise for years (El-Bawab, 2022). The body of gun violence research involves various other issues, such as mental illness (McGinty et al., 2014), frames (Morin, 2016), and public health issues (McKeever et al., 2022). Analyzing a sample of news articles on serious mental illness and gun violence, McGinty et al. (2014) found that "dangerous people" with serious mental illness were more likely to be mentioned as a cause of gun violence than “dangerous weapons.” A recent study by McKeever et al. (2022) conducted an online survey (*N*=510) and found gun control and gun rights as the two salience frames. They also revealed that people held individuals responsible for gun violence and identified background checks as the most salient solution.

1.5. Attribution Theory

The root of frames is drawn from the assumptions outlined in attribution theory (AT) (Heider, 1958; Kelley, 1973; Pan & Kosichi, 1993). So, this research analyzes and explains frames through the lens of AT. Originally developed within social psychology, the theory primarily describes how people explain and perceive the causes of an individual’s behavior (Heider, 1958; McLeod, 2010). While defining the theory, Kelley (1973) says:

> Attribution theory is a theory about how people make causal explanations, about how they answer questions beginning with "why?" It deals with the information they use in making causal inferences, and with what they do with this information to answer causal questions. (p. 107).

As naïve psychologists, people tend to make two broad types of causal attributions: a) dispositional attributions and b) situational attributions (Heider, 1958; Kelley, 1973). Dispositional attributions point to an individual's internal factors as being responsible for an incident. For example, in a car crash, labeling people's reckless driving behavior as a cause could be a dispositional attribution. Situational attributions refer to factors that exist outside an individual and are prevalent in specific situations. In the same example, attributing the snowy road as a cause could be considered a situational factor. Two prominent frameworks provide potential factors and insights that shape people's perceptions of dispositional and situational attributions. These are the covariation model (Kelley, 1973) and the correspondent inference (Jones & Davis, 1965).

1.5.1. Covariation model

Kelley's (1973) covariation model identified three potential factors leading to causal perceptions. These are consensus, distinctiveness, and consistency. 1) Consensus is related to a person or entity that explains how many individuals behave in the same way. High consensus indicates a higher level of situational attribution. 2) Distinctiveness is related to the situations that explain how an individual behaves in other similar situations. High distinctiveness indicates a higher level of situational attribution. 3) Consistency is related to time, which explains how frequently an individual’s behavior occurs. High consistency indicates a higher level of dispositional attributions (Kelley, 1973).

1.5.2. Correspondent inference

Jones and Davis (1965) offered three key factors in inferring causal attributions. 1) People’s degree of choice: A freely chosen behavior is considered to infer an individual’s dispositional attributions compared to forced behavior. 2) Social desirability of behavior: An individual’s behavior that is low in social desirability or social expectedness is more likely to make dispositional attributions compared to high social desirability. 3) Intended consequence of behavior: People infer an individual’s behavior as dispositional, especially when the behavior’s intended consequence is negative and harmful to people.

1.6. Case: 2022 Uvalde Elementary School Shooting

This study analyzes media coverage of a mass shooting that occurred on May 24, 2022, in Uvalde, Texas. An 18-year-old former student named Salvador Ramos entered Robb Elementary School with an AR-15-style rifle and opened fire (Peck & Goodman, 2022; Sandoval, 2023). The shooting resulted in the deaths of 19 students and two teachers and the injuries of 17 others (Jacobo & El-Bawab, 2022; Peck & Goodman, 2022; “Mass shootings,” n.d.). The Uvalde school shooting is one of the deadliest shootings in the United States in terms of the number of casualties (“Mass shootings,” n.d.).

The mass shooting incident received extensive coverage in local, national, and international news media (Kellner, 2022), sparking outrage and reigniting longstanding debates over gun control and school safety and calls for action (Livingston, 2022). News media coverage of the Uvalde shooting evolved over time (Kellner, 2022). Soon after the incident, the then-President of the United States, Joe Biden, visited Texas to console the victims and pledged to act (Livingston, 2022; "Remarks by President," 2022). Within a month of the Uvalde school mass shooting that occurred 10 days after another shooting in Buffalo, New York, a gun safety legislation was passed by the Senate and Congress and then signed by the President on June 25, 2022. The gun safety law is reported as the first of its kind in the previous 30 years (Clyde & Miranda, 2022). As the deadliest mass shooting in recent years and drawing widespread media coverage, the Uvalde elementary school shooting has been purposively selected for this study.

3. Study 1: Qualitative Textual Analysis

This study focuses on an in-depth examination of the usage patterns of specific words, such as adjectives and adverbs, and their semantic relations in constructing frames. Typically, computational tools and traditional research methods, such as qualitative and quantitative methods, are broadly pursued as separate lines of inquiry into frames. However, this study seeks to bridge this divide by utilizing the insights of inductive qualitative research to inform computational approaches in framing analysis.

For this analysis, we purposively selected the 2022 mass shooting as a case that took place at Robb Elementary School. Specifically, we looked at how news media outlets in the right-leaning (a.k.a. WSJ and Fox News) and left-leaning categories (a.k.a. NYT and CNN) use selected modifying words (e.g., adjectives and adverbs) structured in a semantic pattern to frame the shooter, victims, and the event.

Therefore, the following research questions are asked for exploration:

RQ1: How do right-leaning and left-leaning news media outlets use words and phrases to construct frames while covering the 2022 mass shooting at Robb Elementary School in Texas?

RQ2: How do the right-leaning and left-leaning news media outlets frame the shooter, victims, and the mass shooting event at the Robb Elementary school in Texas?

RQ3: How do the semantic relations of words in the Texas mass shooting news reports inform the computational analysis of frames?

4. Study 1 Method

To answer the research questions, study 1 used qualitative textual analysis, a widely used approach to analyze frames inductively (Hertog & McLeod, 2001). It fits with the study's purpose of inductively analyzing news reports to gain an in-depth understanding of frames, word usage patterns, and their semantic relations to constructing frames (Entman, 1993). Qualitative textual analysis is "all about language, what it represents and how we use it to make sense of our [social realities]" (Brennen, 2017, p. 203). While exploring "how texts operate to produce meaning" (Browne, 2009, p. 63), the qualitative analysis helps "make an educated guess at some of the most likely interpretations that might be made of that text" (McKee, 2003, p. 1).

4.1. Data Collection

We collected a total of 100 news reports, including 600 news headlines and paragraphs, published by four news media outlets on the 2022 Robb Elementary School shooting in Texas. Each of them includes ten news reports on the shooting that took place on May 24, 2022. Of the news outlets, *The New York Times* (NYT) and *Cable News Network* (CNN) are selected as the left-leaning news media, and *the Wall Street Journal* (WSJ) and *Fox News* as the right-leaning news media (GMedia Bias Fact Check, n.d.). The news media outlets were categorized based on their bias scores provided by Media Bias/Fact Check (MBFC). The MBFC is a non-partisan American independent site that provides bias scores for media outlets (Media Bias Fact Check, n.d.; Odhner, 2024).

It is important to acknowledge that although Fox News and The Wall Street Journal are both considered right-leaning according to MBFC scores, they differ in tone and editorial focus. As noted by Ad Fontes Media (2024), Fox News tends to be more partisan in its opinion content, whereas The Wall Street Journal is comparatively more centrist in its news reporting ("Ad Fontes Media," n.d.). For the purposes of this study, both outlets were categorized on the right side of the media bias spectrum.

Using purposive sampling, we used the keywords "(Texas OR Uvalde) AND ("shoot*")" and searched articles on *Factiva,* a global news database, for these four news media outlets separately from May 24 to 31, 2022. After downloading all news reports identified during the period, we manually scrutinized and removed articles that are not specifically relevant to this Uvalde mass shooting and are other types of content, such as editorials. This continued until 25 news reports were selected for each new media. Collecting news reports continued until the dataset seemed saturated with relevant words and phrases coded in this study. Data saturation is considered a useful guide for sampling data in a qualitative study that deals with a relatively small amount of information (Sandelowski, 1995; Brennen, 2017). With this process, a sample of 100 news articles was finally selected for this analysis. Since the lead and initial paragraphs(s) generally represent the most important messages in a news story (Liu et al., 2019; Van Dijk, 1985), we purposively selected the headlines and first five paragraphs of each news report, totaling 600 headlines and paragraphs, for an in-depth analysis.

4.2. Data Analysis

This study analyzed the news reports with word-by-word coding in three phases—open coding, axial coding, and selective coding (Saldaña, 2016) using NVivo, a qualitative data analysis software. The coding process was guided by the three research questions, framing theory (Entman, 1993), and attribution theory (Heider, 1958; Kelley, 1973). Following the research purpose and questions, the data analysis focused on using words and their semantic relations in constructing frames (Entman, 1993) in the case of the Robb Elementary School shooting. The analysis explores whether and how the news reports used various words and phrases to promote particular interpretations or evaluations relating to the shooter, victims, and the event.

During the open coding phase, we specifically looked at the use of words and phrases that promoted or highlighted four aspects: a) the shooter, Salvador Ramos; b) victims, such as school children and teachers; c) the shooting incident, in certain ways. Each type of word and phrase was coded into a separate code. For instance, the words “kills,” “killing,” and “killed” were coded into a single code. In the axial coding phase, where related codes are grouped into broader categories, we organized initial codes into similar categories based on shared framing purposes. Finally, during the selective coding, where central theme(s) are refined, a few broad themes emerged with adequate exemplars (Table 3.1 and Table 3.2 in Appendices 3.1 and 3.2). To ensure validity, we used two strategies: data triangulation (drawing from multiple data sources) and disconfirming evidence (intentionally seeking and considering both supporting and opposing evidence from data) (Creswell, 2016).

5. Study 1 Findings

RQ1 and RQ2: The analysis identifies distinct sets of specific words and phrases in left-leaning news outlets, such as NYT and CNN (see Table 3.1 in Appendix 3.1), and right-leaning outlets, such as WSJ and Fox News (see Table 3.2 in Appendix 3.2), that frame the shooter, victims, and the shooting event differently.

5.1. Shooter

“Accused” killer. The analysis shows that both right-leaning and left-leaning media outlets use some common verbs (e.g., kills, left dead, opened fire, and shot) while attributing the shooter’s act. Importantly, right-leaning media outlets use weaker verbs and modifiers (e.g., “is accused of shooting,” “claimed lives,” and “allegedly committed by”), which casts doubt on Salvador Ramos' crime and weakens the gravity of killing people. In contrast, left-leaning media use stronger verbs and modifiers (e.g., “shot and killed,” “burst in and killed,” and “horrifically”), which presents the incident with a higher gravity of the mass shooting act.

Differences in identifying the shooter also exist between the two groups of media outlets. In contrast to the left-leaning media outlets, the right-leaning ones use weaker terms like “alleged gunman” and “alleged shooter.” For example, a news report published by *Fox News* on May 27 said, “Salvador Ramos, the alleged gunman accused of shooting his grandmother and then targeting dozens of victims….” This seems to have cast doubts regarding Salvador Ramos’ act of killing people, at least to some extent.

Mental instabilities. In terms of attributing the responsibility or blame to Salvador for the mass shooting, both groups of news media outlets showed their own bias. In contrast to left-leaning news media outlets, the right-leaning ones selected and highlighted Salvador’s mental and family-related instabilities, drawing attention to the social factors while presenting Salvador as responsible for the shooting. For example, the WSJ reported, “Salvador Ramos… came from a broken family and unsettled classmates and co-workers with sometimes aggressive behavior and disturbing social-media posts.”

5.2. Shooting Incident

Low vs. high severity. The shooting incident has been found to have been presented differently in terms of its severity between the right-leaning and left-leaning news media outlets. The former has employed specific words and phrases to portray the incident as a less severe one than the latter. The event modifier columns of Tables 1 and 2 demonstrate that while some words and phrases are common to both groups of news outlets, others are used to attribute the shooting incident differently. For example, the NYT used the word "slaughter" to describe the incident, while Fox News and the WSJ did not. The left-leaning media also used "terrorist attack," which was not used by their right-leaning counterparts. In contrast, right-leaning media outlets identified the shooting as a "senseless crime." The use of words such as "terrorist attack" and "slaughter" might trigger nodes in the human brain related to other deadly terrorist incidents, portraying the shooting as a more severe act (Collins & Loftus, 1975). On the other hand, the use of the phrase "senseless crime" suggests a typical type of crime. Therefore, differences in the use of words have contributed to defining the mass shooting as a problem in terms of its severity between the right-leaning and left-leaning news media outlets.

5.3. Victims

Teenager vs. older. Divergent portrayals of the victims in the news media outlets were observed, with both left-leaning and right-leaning news media using distinct words and phrases, although some terms were commonly employed. The left-leaning media outlets specifically employed phrases such as "school children," "elementary school children," and "kids," whereas the right-leaning outlets used "children" and "students." For instance, the NYT reported on "the killing of at least 19 elementary school children in second, third, and fourth grades." Although subtle, this contrast indicated the left-leaning outlets' emphasis on the word "elementary," framing the shooting incident as an attack on young children of this age group. The word "elementary" distinguishes the age range of 5-10 years from "children" and "students." The word "elementary" emphasizes the victims' age range more clearly, whereas terms like “students” and “children” are more general and can apply across various age groups, including older youth. By promoting the ages of the victims in distinct ways, both the left-leaning and right-leaning news media outlets presented the severity of the shooting incident and drew attention to the shooting problem differently.

RQ3: Answers to RQ1 and RQ2 highlight the use of words in constructing relevant frames. The RQ3 serves the main purpose of Study 1, which is to investigate the semantic patterns or semantic relations of those words in creating frames. As a way of answering it, this analysis provides various groups of words and phrases centering on the shooter, victims, and the event, illustrating semantic relations among the words (see Table 3.1 and Table 3.2 in Appendices 3.1 and 3.2).

5.4. Semantic Relations

The above results and Table 3.1 and Table 3.2 provided in the analysis present two crucial aspects that contribute to the understanding of frames. Firstly, the words used to construct frames are crucial in identifying the framing components utilized by the news media outlets. Secondly, the semantic relations among the words are crucial in establishing the frames’ meaning. Semantic relations indicate how the words are interrelated and which entity the words are attributed to. Reading through only the words might provide some insights into relevant framing components, but the insights are not fully meaningful without the words’ semantic relations. When the words are read with their semantic relations, it renders particular meanings to construct frames. For instance, in the excerpt “An 18-year-old gunman on Tuesday fatally shot 19 children and two adults” from an NYT article, the semantic relation between the phrase “18-year-old” and the “gunman” (Salvador) highlights that the modifier refers to the gunman and not the children. Without considering the semantic relations, it appeared challenging to comprehend the relevant meanings of the words and subsequently construct frames.

6. Study 1 Discussion

6.1. Highlight and Hide

As the findings indicate, both left-leaning and right-leaning news media highlighted some common and different words regarding the Texas mass shooting, conforming to the framing strategy of highlighting and hiding certain aspects of the event (D’Angelo, 2018; Entman, 1993; Greussing & Boomgaarden, 2017). As a frame functions to purvey various judgments about reality (Entman, 1993; D’Angelo, 2018), the frames constructed by left-leaning and right-leaning news outlets may shape how people perceive and understand the causes of the mass shooting and influence their attitudes toward it.

6.2. Attribution of Responsibility

Left-leaning news media outlets attributed more responsibility to Salvador for the mass shooting compared to right-leaning ones. As the attribution theory (Kelley, 1973) and framing theory (Entman, 1993) suggests, with highlighted salience in situational factors (e.g., broken family) in right-leaning outlets, people are more likely to attribute the shooting’s causes to situational factors. This is supported by the phrase “accused of” that right-leaning news used in presenting Salvador’s shooting.

Right-leaning news media highlighted aspects of Salvador's social factors, such as his broken family, which may have made his actions appear more situationally driven. As per the correspondent inference model (Jones & Davis, 1965), such social desirability can reduce attributions of personal responsibility by shifting away from dispositional factors. Overall, the left-leaning news media reports focused on attributing the causal responsibility of the mass shooting more to Salvador, while the right-leaning news media reports went beyond Salvador's individual responsibility to his family factors. Such causal interpretation is supported by McGinty et al.'s (2014) study, which shows "dangerous people" with mental illness were more likely mentioned as a cause of gun violence than "dangerous weapons." The study by McKeever et al. (2022) also extends evidence in support of this current study's findings.

6.3. Semantic Relations for Computational Framing Analysis

Unsupervised computational methods mostly rely on the ideas of frequencies and co-occurrences of words (Blei, 2012; DiMaggio et al., 2013). These bag-of-words-based approaches are not designed to look at the semantic relations of words and end up with identifying topics, instead of frames (Ali & Hassan, 2002). The study 1 findings demonstrate that capturing semantic relations helps discern in-depth nuances in the texts through word relations and, thus, identify relevant frames. For example, in the following excerpt from a *New York Times* article, "An 18-year-old gunman on Tuesday fatally shot 19 children and two adults," the semantic relations show that the phrase "18-year-old" modifies the "gunman" (aka Salvador), not children. Without knowing this semantic relation, relevant meanings of the words and, subsequently, frames do not emerge (see Tables 1 and 2).

Manual data analysis enables the researchers to identify such semantic relations and relevant frames, as presented above. Therefore, semantic relations appeared essential for having relevant meanings and frames in a text. In a computational method, being able to capture the semantic relations seems to be a one-step advancement toward better identification and analysis of frames. As identified in this study 1, the lists of words, their attributes, and semantic relations for the shooter, victims, and the event are so specific that these can be incorporated into an algorithmic model. So, this study suggests incorporating these semantic relations into computational techniques (e.g., dependency parsing) for better automatic framing analysis. As envisioned in Nicholls et al. (2021) and Ali and Hassan (2022), this current study's findings extended additional evidence of how semantic relations among words and phrases, instead of just bag-of-words, can better explain nuances of frames, especially in an unsupervised model.

7. Study 2: Computational Analysis

Study 2 of computational analysis builds on the insights and recommendations from Study 1 of qualitative textual analysis. It focuses on the potential of using dependency parsing, an NLP technique that analyzes the grammatical structure of a sentence by identifying relationships between words, such as which word modifies or depends on another. This approach aims to enhance the identification and analysis of frames computationally. Examining the same dataset of news articles from Study 1, this computational analysis explores how dependency parsing can capture the semantic relations of words and understand relevant frames. We also compare the results of the unsupervised computational model (Study 2) with those obtained through manual data analysis (Study 1) to evaluate the effectiveness of the computational approach. The findings contribute to a better understanding of the role of the semantic relations-based computational approach in analyzing frames and offer insights into the potential of using dependency parsing as a methodological approach for framing analysis.

Since it is one of the first studies to use semantic relations in analyzing frames, we offer similar research questions established in study 1, consistent with the objectives of study 2.

RQ1: How do right-leaning and left-leaning news media outlets use words and phrases to construct frames at the Robb Elementary school in Texas?

RQ2: How do the right-leaning and left-leaning news media outlets frame the shooter, victims, and the mass shooting event at the Robb Elementary school in Texas?

8. Study 2 Method

8.1. Dataset

To answer the research questions, we analyzed the same news report dataset as study 1. Parsed by the spaCy NLP language model, the dataset contains a total of 24604 tokens, with 4768 for CNN, 6282 for Fox, 6759 for NYT, and 6795 for WSJ.

We used the same dataset to compare the frames provided by the computational approach with those of the qualitative study.

8.2. Analysis

The data analysis involved the following seven steps:

1) Coreference resolution: As this study aims to identify modifying words centering three entities, the shooter, victims, and the event, we needed to identify and resolve the coreferences (e.g., “he” or “suspect” for the shooter) to capture all possible modifying words of both “references” and “co-references.” To accomplish this, we applied NeuralCoref, an extension of the spaCy NLP library that provides coreference resolution.

2) Token extraction: We then applied a dependency parser of the spaCy language model that parsed all the news reports and generated a dependency parse tree. This tree provides the syntactic structure of a sentence that includes nodes, such as heads (e.g., gunman) and children (e.g., suspect), representing words, and edges representing the semantic relationships between those heads and children. Each edge is labeled with a specific dependency relation, such as “amod” (adjective modifier).

3) Determining keywords: To capture all possible words that refer and co-refer to each of the three entities, we determined relevant keywords for each entity (e.g., Salvador, gunman, shooter). These keywords were determined based on study 1 insights and then refined through manual checking of some tokens in the output produced in step 2 (see Table 3.3 in Appendix 3.3 for details).

4) Filtering heads and children: Based on the keywords, we filtered out all relevant "heads" and "children" of each entity, all their dependency relations, and associated news outlets.

5) Determining and refining dependency relations: This step determines and refines useful dependency relations based on this study's purpose. We removed some dependency relations (e.g., cc, punc) that were not useful in making meanings in relation to the RQs, by manual checking of the relations grouped in the output produced in step 4 (see Table 3.3 in Appendix 3.3 for details).

6) Filtering "framing components": We consider each pair of head and child with certain dependency relation (e.g., the pair of "shooting" keyword and "deadly" child with "amod" relation) as a framing component that provides a particular attribution to an entity. This step filtered out all framing components for each entity by the news outlets.

7) Framing components to frames: Until the last step, we analyzed the data computationally using spaCy and Pandas, a popular data analysis library for Python. In this step, we followed both computational and manual qualitative explorations. 7a) Computational: We computationally grouped the framing components for each entity by dependency relations. To achieve this goal, we used BERT word embedding and *k*-means clustering of the modifying words (also known as children). 7b) Qualitative: We inductively coded the modifying words and categorized them into groups following the research questions manually. Here, we consider a single framing component as a candidate for being included in multiple groups (Saldaña, 2016), and triangulation and disconfirming evidence were utilized to ensure the validity

(Creswell, 2016). In both parts, each group is considered as a frame. With the process, a number of frames emerged with exemplars.

9. Study 2 Findings

This section reports the findings of the qualitative analysis in step 7, followed by the computational analysis from steps 1 to 6. The results of the computational exploration in step 7 are not reported here, as we found that the findings from manual analysis outperformed them. The clusters revealed through *k*-means clustering were not found to be coherent and adequately insightful for understanding the nuances of frames, as we examined the results manually. The findings of the qualitative analysis reveal that right-leaning and left-leaning news media outlets use different words to construct frames of the shooter, victims, and the mass shooting event at the Robb Elementary school in Texas differently, as presented in Tables 4, 5, and 6, respectively.

9.1. Shooter: "Accused" killer

The shooter was characterized with some words that create doubt over the shooter's killing action. Comparing the attributions used by right-leaning and left-leaning news outlets, it was found that the former used the words "alleged [shooter]" and "suspected [shooter]" more frequently than the latter. Furthermore, a right-leaning news outlet referred to the shooter as "accused [of shooting]", which was not used by the left-leaning outlets. These attribution differences suggest that the two media outlets utilized different priorities in framing the shooter.

9.2. Shooter: Diversity of attributes

As depicted in Table 3.4 (Appendix 3.4), right-leaning news outlets used a greater variety of attributes to highlight various aspects of the shooter than left-leaning outlets. For example, right-leaning outlets used words such as “unhappy,” “deceased,” “civilized,” and “active” to describe the shooter, which left-leaning outlets did not use. However, these words appear scattered and do not seem to form a coherent argument. This may be due to the small dataset used in this study. A larger dataset in future research could reveal more modifying words and categorize them into relevant groups, providing further insights into framing strategies.

9.3. Shooter & Victims: Teenager vs. Older

Right-leaning news media outlets tend to use words depicting the “shooter” as comparatively younger than left-leaning outlets. For instance, words used by the right-leaning outlets to attribute to the shooter include “teenage,” “young,” and “student,” which left-leaning outlets did not mention. Another example is that the shooter was identified as “18-year-old” 26 times in the left-leaning outlets and only 10 times in the right-leaning ones (see Table 3.4 in Appendix 3.4). In contrast, the victims were attributed with the word “young,” an adjective modifier, five times by the left-leaning outlets and zero times by the right-leaning ones. Overall, right-leaning outlets frame the shooter as younger and the victims as older, and the scenario is the opposite in left-leaning outlets (see Table 3.5 in Appendix 3.5).

9.4. Victims: Our Kids vs. Your Kids

There is not much difference between left-leaning and right-leaning news outlets in using personal pronouns to modify the victims (see Table 3.5 in Appendix 3.5). Pronouns addressing victims are important to perceive how the news media outlets stand with them. The left-leaning outlets still used a greater variety of personal words, such as my (2), our (2), and your (2), while the right-leaning ones used two such words, her (2) and our (3).

9.5. Shooting Event: Low vs. High Severity

To describe the shooting, left-leaning news media outlets tend to use more severe and emotionally charged words, such as deadliest (6), deadly (6), horrific (1), horrifying (1), heinous (1), tragic (1), and fatally [shot] (1), which frames the issue as a more significant problem. Such words used by right-leaning outlets include deadly (4), deadliest (3), awful (4), horrific (3), senseless (2), and devastated (1). This shows the right-leaning outlets use less intense words like “senseless” and “awful,” which suggests a less severe framing of the issue (see Table 3.6 in Appendix 3.6). Overall, the framing of the mass shooting as a problem is affected by the language used by news media outlets, and the severity of the framing can differ based on the political leaning of the outlet, which is aligned with framing aspects suggested by Entman (1993).

10. Study 2 Discussion

Study 2 investigates how news outlets frame the shooter, victims, and the Texas school shooting, applying a new computational approach based on semantic relations.

10.1. Attribution of Responsibility

Framing the shooter as "young" or "older" can have significant implications for how people perceive the shooting and the level of responsibility attributed to the shooter (Entman, 1993). The use of the "young" attribute by right-leaning outlets could soften the shooter's image and create a more sympathetic portrayal, thereby reducing the level of responsibility attributed to him (Jones & Davis, 1965). On the other hand, the left-leaning outlets' focus on the victims' youth could create a greater sense of tragedy and urgency and, therefore, a higher level of responsibility attributed to the shooter (Decety et al., 2012). As per attribution theory, people tend to attribute a person's behavior to internal or external factors based on internal and external factors (Heider, 1958; Kelley, 1973). In this case, the framing of the shooter and victims differently by the news outlets might shape how people attribute responsibility for the shooting. The framing differences among news media outlets might have been shaped more by established media routines and practices (Reese & Shoemaker, 2018) than by the specifics of this particular mass shooting event.

10.2. Taking Actions

The news media outlets' different approaches to highlighting selected "severe" words might have significant implications for how the public perceives the incident and “choose[s] to act upon” the problem (Entman, 1993, p. 54). The left-leaning news outlets' higher salience on words like “deadly” and “deadliest” might activate the "amygdala” node in people’s brains, potentially leading them to take actions like protest and advocacy (Phelps, 2006; Barry et al., 2013). At the same time, highlighting more on the words “accused” and “alleged” regarding the shooter’s act, the right-leaning news outlets, compared to left-leaning ones, offered doubt in people’s perception regarding Ramos’s mass shooting. Such higher salience on these words in right-leaning outlets seems to have weakened people’s perception of the shooter’s dispositional factors in committing the offense (Kelley, 1973).

10.3. Highlight and Hide

Taking some meanings or words over others as discussed above conforms the framing technique of highlight and hide, as proposed by Entman (1993) and Fairhurst and Sarr (1996). Both right-leaning and left-leaning news outlets utilized distinct ways of framing the shooter, victims, and the event despite some common depictions between the groups. Overall, the left-leaning outlets attempt to elicit people’s sympathy for “victims” while right-leaning ones sympathize with the shooter, as evidenced above.

11. Integrated Discussion of Both Studies

The primary objective of this research is to introduce and explore a new approach to computational framing analysis. Our initial qualitative inquiry in study 1 revealed in-depth insights into the role of semantic relations in constructing frames and suggested that dependency parsing, a computational method, could potentially serve as a practical unsupervised approach to framing analysis. Based on the findings and recommendations from study 1, study 2 applied dependency parsing to the same dataset as an approach to computationally analyzing frames.

A comparison of findings in study 1 and study 2 demonstrates the potential of this proposed semantic relations-based approach to automate the identification and analysis of frames in large datasets. As the study 2 discussion suggests, its findings on framing the shooter, victims, and the event are well interpretable with relevant theoretical frameworks, and the interpretations are mostly aligned with those of study 1 and prior gun violence research. With that, this article proposes a novel computational framing analysis approach based on dependency parsing named *"Semantic Relations-based Unsupervised Framing Analysis"* (SUFA).

11.1. Semantic Relations-based Unsupervised Framing Analysis (SUFA)

11.1.1. Novelty of SUFA

The SUFA is novel in analyzing frames in several ways. First, it is based on semantic relations that extend beyond the bag-of-words approach utilized by most existing unsupervised computational framing analysis methods, such as topic modeling. Second, as discussed above, a few studies have employed semantic relations in framing analysis (e.g., van Atteveldt et al., 2013; Ziems & Yang, 2021). However, they did not present it as an unsupervised method. In contrast to these studies, our approach demonstrates its distinction as an unsupervised method. Researchers do not need to define a frame in advance to explore frames within a dataset. Third, our approach provides flexibility in using qualitative manual coding or computational tools like word embedding and k-means clustering in step 7 of its data analysis process. In this sense, it is a mixed-method approach that prior studies did not include.

11.1.2. Data Analysis in SUFA.

The procedure for analyzing data in SUFA is outlined in seven steps in the method section of study 2. To effectively apply this method, we recommend following these steps along with a few additional considerations. If news media outlets have specific identities, such as left-leaning or right-leaning, we suggest labeling these identities in the data. Steps 3 and 5 require human intelligence. For example, step 3 involves providing keywords for each entity, which can be informed by domain knowledge, researchers' little manual data exploration, or with the assistance of WordNet (Miller, 1995). In step 5, researchers may need to manually review the output to identify useful semantic relations for analysis. For step 7, either qualitative or computational analyses can be used, depending on the research goals and the number of modifying words derived from the dataset. The computational analysis, such as word embedding and *k*-means clustering, may generally be more appropriate as the SUFA is meant to analyze large datasets. However, if the size of modifying words is small enough to manage manually, a qualitative analysis might be more suitable for step 7, as research suggests that human intelligence often outperforms machines in tasks that require contextual interpretation and subjective judgment (Lazer et al., 2009).

11.1.3. Strengths

The SUFA is an unsupervised approach that does not require any prior labeling or defining of data frames. Instead, it uses an inductive approach to explore and group attributions together to reveal framing components or frames. One advantage of the SUFA is that it allows for the flexibility of utilizing both human intelligence and computational techniques to emerge frames and their interpretations, particularly in cases where the size of modifying words is small enough to be managed manually. Moreover, the SUFA can inductively analyze frames in large-scale datasets in an unsupervised manner.

11.1.3. Weaknesses

The approach requires manual input in determining relevant keywords (step 3) and semantic relations (step 5), which can be time-consuming and subjective. It is limited to analyzing frames centered around entities, such as an individual (e.g., shooter), a group of people or community (e.g., victims), and an incident or phenomenon (e.g., a shooting event). Since this study focuses on exploring emphasis frames through a semantic relations-based approach, it is better suited for analyzing emphasis frames (D’Angelo, 2017) rather than equivalency frames (Kahneman & Tversky, 1984).

Additionally, like other computational framing analysis approaches, this research only considers words and phrases, while other framing components like metaphor, text placement, and visual elements are not analyzed. During the coreference resolution (step 1), some words that are useful as framing components could be replaced with co-references (e.g., replacing the word “gunman” with “Salvador”), which might lead to the loss of some words with important nuances.

12. Conclusion

This research introduces a new computational approach called *Semantic Relations-based Unsupervised Framing Analysis* (SUFA), which utilizes semantic relations to analyze news frames. While the method has some limitations, such as the need for manual input and its focus on emphasis frames, it provides a useful tool for exploring framing components in news media coverage. The mixed-method approach of SUFA offers researchers the flexibility to use entirely computational tools or couple it with qualitative manual coding, where applicable, for data analysis. Overall, SUFA is a valuable addition to the field of computational framing analysis, enabling more comprehensive and nuanced analysis of news media frames.

10.3. Limitations and Future Research

The SUFA was developed and tested on a single dataset of news reports on gun violence in Study 1 and Study 2. However, the approach can be applied to other domains with the provision of relevant keywords and relations. Further research can be conducted to explore the applicability of this method in other domains and to improve its performance. Currently, SUFA only considers words when analyzing frames. However, the computational framing analysis needs to include other framing components such as metaphor, visual content, placement, differences between headline and body texts, and exemplars. Such advancements will provide a more comprehensive understanding of framing effects in news media.

Disclaimer*:* A prior version of this article won the *"Top Method Paper"* Award at the Communication Theory and Methodology Division Competition of the *Association for Education in Journalism and Mass Communication (AEJMC) Conference*, D.C., United States, in 2023.

D. CHAPTER 4: Framing Mass Shooters and Victims in the United States: Applying and Advancing SUFA Approach

1. Introduction

Computational framing analysis appears as an emerging methodological framework in both computation and social science, enabling scholars to identify and analyze frames in large datasets computationally. The computational methods that prior studies utilized in framing analysis are broadly divided into supervised and unsupervised approaches. While supervised methods require predetermined labels and substantial human labor, unsupervised methods are more flexible and require less human effort. Traditional computational methods have relied on a number of unsupervised approaches, such as topic modeling, clustering, and network analysis, to explore frames. However, these techniques fail to capture semantic relationships between words, hindering the deeper insights needed to better understand frames (Ali & Hassan, 2022). A novel computational framing analysis approach titled *"Semantic Relations-based Unsupervised Framing Analysis"* (SUFA) was introduced in Chapter 3 of this dissertation to address this limitation by integrating semantic relations-based techniques with a more contextually informed approach to identifying frames. Unlike the traditional unsupervised approach based on bag-of-words, frequency, and co-occurrence of words, SUFA leverages semantic relations to identify frames, offering a more nuanced approach to computational framing analysis. However, its application remains limited to smaller datasets until now, leaving its potential for large-scale media frame analysis unexplored (Ali & Hassan, 2023).

Therefore, this study seeks to advance the SUFA approach by applying it to a large dataset in the domain of gun violence. Specifically, this research analyzed one month of news media texts (N = 1334) from the nine national news outlets covering a 2022 Texas mass shooting in the United States, focusing on how mass shooters and victims are framed in the news media coverage. Mass shootings in the United States have emerged as a critical public health crisis, sparking widespread media coverage and political debate. With over 40,000 firearm-related deaths and around 76,000 nonfatal injuries annually, disrupting communities, families, and individuals, gun violence poses an economic consequence to the U.S. of $557 billion annually ("The Economic Cost," 2022; "Center for Gun Violence Solutions," n.d.). The mass shooting is one of the key contributors to the overall gun violence.

The framing of shooters and victims in news reports plays a crucial role in shaping public perceptions, influencing policy responses, and guiding discourse on gun control. While the political impasse and public debate over gun violence continue, it is important to understand how news media outlets frame the issues, especially in large datasets, as different frames lead people to "choose to act upon [a problem]" differently (Entman, 1993, p. 54).

Prior studies on mass shootings reveal various frames in news media content relating to different entities (e.g., a shooter and a victim) involved in gun violence. For example, looking into the Virginia Tech shooting's news content published by the *New York Times* (NYT) and *USA Today,* Park et al. (2012) identified the shooter's racialization framed in the news outlets. Around 51% of the articles studied here identified the shooter's race as *Asian*, and approximately 37% of the articles exposed his ethnicity as *Korean.* Many other studies also reported variants of frames regarding the shooter's identities. Importantly, framing analysis traditionally relied on qualitative and quantitative methods involving manual labor and small dataset analysis. However, with the explosion of online news reports and social media posts, there is an unprecedented amount of digital data available that is difficult to analyze manually. Until now, a minimal amount of research addressed the mass shooting issue using large datasets, and these studies used a supervised approach, leaving the knowledge of the unsupervised approach unutilized in this domain (Liu et al., 2019; Bhatia et al., 2021; Guo et al., 2023). Research on gun violence using various methods is minimal, considering the importance of this public health pandemic, requiring further investigation into the national issue (McKeever et al., 2022).

Building on the prior study in Chapter 3, this current research further enhances SUFA, particularly by refining its seventh step—automated clustering of framing components—using advanced computational tools in a large dataset. Prior applications of SUFA in Chapter 3 primarily relied on a small dataset, limiting its effectiveness as a fully unsupervised method. By integrating state-of-the-art natural language processing (NLP) techniques, dependency parsing, coreference resolution, clustering, and a large language model (LLM), such as OpenAI's GPT, this research improves SUFA's capacity for large-scale, entity-centric framing analysis.

Research found that existing approaches like topic modeling and network analysis result in exploring "topics" instead of going deeper into insights into "frames" (Nicholls & Culpepper, 2021; Ali & Hassan, 2022; Entman, 1993). Utilizing a semantic relations-based approach in this project demonstrates an effective approach for exploring deeper insights into frames, especially entity-centric emphasis frames (Ali & Hassan, 2023).

The results are discussed and interpreted through the lens of framing and attribution theory. One key contribution this research made was advancing the gun violence literature from the computational framing analysis perspective. Second, this research advanced the SUFA model by applying it to a large dataset and further improving the seventh step—automatically clustering framing components—in the SUFA analysis process.

2. Literature Review

2.1 Gun Violence as a Public Health Epidemic

Gun violence in the United States represents a significant and pressing public health crisis, with over 40,000 firearm-related deaths and around 76,000 nonfatal injuries annually. This has an economic consequence to the U.S. of $557 billion annually ("The Economic Cost of Gun Violence," 2022; "Center for Gun Violence Solutions," n.d.). Public health experts and professionals have identified gun violence as an epidemic and a unique crisis for the United States (American Public Health Association, n.d.). Considering the importance of addressing gun violence, the government and academicians have been working in different capacities and efforts. This concern received dedicated attention from many universities and institutions. Among others, the University of Maryland took an initiative called *"Prevent Gun Violence: Research, Empowerment, Strategies and Solutions"* (PROGRESS) in 2023 to support reducing gun violence through research. While delivering a speech on the House floor in 2023, Representative Kim Schrier described gun violence as an epidemic that "continues to plague" the United States. As one of the leading causes of death (Wilson, 2023), gun violence underscores the urgent need for academic research aimed at better understanding media frames as a way to contribute to identifying potential solutions for reducing gun violence.

2.2. Gun Violence Research with Framing Analysis

News media deploy different frames attributing shooters and victims in mass shootings, potentially shaping public perception of the individuals (aka entities). These frames might affect audience engagement and the overall understanding of gun violence among people. Prior studies attempted to understand various aspects relating to mass shootings, including racial and ethnic characterization of shooters and victims (e.g., Park et al., 2012), mental health (e.g., McGinty et al., 2014), various frames (e.g., Chyi & McCombs, 2004; Morin, 2016), personal background and experiences (e.g., Emelu, 2023), and other public health issues (e.g., McKeever et al., 2022).

Of the body of gun violence research, the characterization of various entities in the context of a mass shooting deserves further attention, especially utilizing large datasets. As per the prior studies, news media outlets from different spectrums of biases depict a shooter differently (e.g., highlighting or hiding the mental state of the shooter).

In terms of using particular methods, existing gun violence research mainly used qualitative textual analysis (e.g., Morin, 2016) and quantitative research methods, including surveys (e.g., McKeever et al., 2022). The emergence of a large amount of digital texts in recent decades, such as tweets, news reports, and publicly available government documents, poses a challenge to understanding frames using traditional qualitative and quantitative research methods that can reasonably handle a small amount of data. In this context, scholars from both computation and social sciences started using various computational tools to explore frames in large datasets (e.g., Liu et al., 2019; Tourni et al., 2024).

In the area of gun violence research, only a few studies so far have explored mass shooting frames by analyzing large datasets, such as Liu et al. (2019), Guo et al. (2023), and Tourni et al. (2024). However, these scholars utilized supervised models like topic modeling and deep learning as a way of analyzing and identifying frames.

Prior studies indicate that unsupervised computational framing analysis, especially the semantic relations-based approach, has not yet been extensively applied to exploring mass shooting frames in large datasets. As one of the first few studies, this project utilizes the newly proposed semantic relations-based entity framing approach, *Semantic Relations-based Framing Analysis* (SUFA), to explore frames in the context of gun violence (Ali & Hassan, 2023).

2.3 News Spectrum in the U.S.

The news media landscape in the United States is diverse. The news media industry includes media houses, including a wide range of biases. The Media Bias Fact Check (Media Bias Fact Check. (n.d.) assess and document the news media's biases in seven categories: extreme left, left, left-center, least biased, right-center, right, and extreme right. For this current research project, we collected news media content from multiple news media outlets in different bias categories measured by the

MBFC's scale: left-center, least bias, and right-center. The purpose behind selecting news media outlets in the selected MBFC bias categories is to see how different news media outlets in certain bias categories frame various entities in the context of gun violence.

2.4. Framing Analysis

This subsection briefly discusses framing and framing components, considering that we elaborated on them in Chapters 2 and 3. Many scholars defined a frame or framing, but they do not have a consensus on defining a frame in the framing scholarship. One of the widely cited and popular framing definitions is provided by Entman (1993), who says:

> To frame is to *select some aspects of a perceived reality and make them more salient in a communicating text in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation* for the item described. (p. 52)

From the communication perspective, one of the key aspects in this definition is the phrase "in such a way," which indicates that a text could be utilized in many different ways to select some aspects and make them salient. As per prior studies, these ways of presenting a text or message include simple words, phrases, metaphors, catchphrases, stories, hashtags, the size of the texts, the color of the texts, and visuals.

Tankard et al. (1991) defined it from the news media context. They mentioned a frame is "a central organizing idea," and it "denotes how journalists, their sources, and audiences work within conditions that shape the messages they construct as well as the ways they understand and interpret these messages" (D'Angelo, 2018, p. xxiv).

Therefore, a frame does not have a simple topic or theme that shares a simple descriptive discussion. Instead, a bit more in-depth insight is needed to construct a frame. Such insights beyond a simple topic could be offered by various “ways” of presenting, highlighting, and hiding certain texts, aiming to enhance the salience of some selected texts.

As mentioned above, there are many framing components, including words, metaphors, size, color, story, and text placement. For this research, the SUFA model is utilized as a computational framing analysis approach that mainly considers words, unlike other unsupervised framing analysis methods.

2.5. Conceptualization and Operationalization of the Framing Component

As previously elaborated in Project 2 (Chapter 3), the SUFA approach conceptualizes a framing component as a pair of words connected by a meaningful semantic relation, such as an adjective modifying a noun or a verb describing an action taken by an entity. These pairs are systematically extracted using dependency parsing, allowing for consistent identification of how entities are described across a large dataset. This operationalization provides the foundation for analyzing entity-centric frames and enables the SUFA approach to scale framing analysis beyond traditional manual methods.

2.6. Attribution Theory as Theoretical Framework

In previous Chapters 2 and 3, the attribution theory has been elaborated in the context of its application in the computational framing analysis of this dissertation.

Following research in prior chapters, this current project in this Chapter also utilizes attribution theory as a lens to interpret and discuss the results.

This social psychological farmwork has long been used in academic research in understanding and describing people's causal explanations. People's causal explanation or causal behavior is considered a crucial component to understanding how certain entities are attributed and framed across the fields, including strategic communication, strategic decision-making, journalism, politics, and business. Notably, a frame is rooted in the attribution theory's assumptions (Heider, 1958; Kelley, 1973; Pan & Kosichi, 1993). Kelley (1973) defines the attribution theory by saying that "attribution theory is a theory about how people make causal explanations, about how they answer questions beginning with "why?" (p. 107).

As explained by Heider (1958) and Kelley (1973), people tend to make two opposite causal attributions: internal factors and external factors. Kelley's covariation model provides three potential factors for causal attributions. These are: consensus, distinctiveness, and consistency (Kelley, 1973). Additionally, Jones and Davis (1965) provide a few key factors that trigger causal attributions. These are people's degree of choice, social desirability, and intended consequence of behavior.

Following prior research exploring frames of various entities (e.g., the shooter), this study proposes to investigate the framing of two key entities—the shooter and the victims— in the context of the 2022 Uvalde elementary school mass shooting. Grounded in attribution theory, this research discusses causal explanations and responsibility attributions shaped by media frames surrounding these entities. Attribution theory suggests that individuals assign blame or responsibility based on internal factors (e.g., personality, intent) or external factors (e.g., situational influences, systemic failures) (Heider, 1958; Kelley, 1973). By analyzing how the shooter and victims are framed through these attributions and frames, this study aims to uncover underlying patterns in media framing that might influence public perceptions, policy debates, and broader societal responses to mass shootings.

Overall, this research employs and advances the SUFA model, an entity-centric computational framing analysis approach. The strengths of this methodological framework align well with this study, as it represents one of the initial applications of semantic relations-based framing analysis. Accordingly, the following research questions are proposed to examine entity-centric emphasis frames in news media coverage of the 2022 Uvalde mass shooting.

RQ1: What types of frames do U.S. news media outlets deploy to attribute the perpetrator and victims in the 2022 Uvalde school mass shooting?

RQ2: To what extent do U.S. news media outlets deploy framing components (e.g., adjectives, adverbs, and other modifiers) to attribute the perpetrator and victims in the 2022 Uvalde school mass shooting?

RQ3: How do framing components deployed by U.S. news media to attribute the perpetrator and victims in the 2022 Uvalde school mass shooting differ across media bias groups (left-centered, least-biased, right-centered)?

RQ4: How do different framing components predict the bias of U.S. news media outlets in the 2022 Uvalde school mass shooting?

This study seeks to automate the clustering process of framing components, a methodological gap left unexplored in previous research, by leveraging state-of-the-art computational techniques. Specifically, it examines the efficacy of large language models (LLMs), such as OpenAI's GPT-4o, in grouping framing components into frames. LLMs have gained significant attention for their ability to process and analyze textual data based on human-generated prompts, demonstrating promise in semantic understanding and content classification.

Recent research has shown that LLMs can replicate human-level clustering accuracy under specific conditions, reducing cognitive bias while enhancing scalability in classification tasks (Hillston, 2024). Given the increasing interest in prompt-based analysis for text clustering, this study evaluates how effectively GPT-4o groups framing components into frames. By comparing GPT-4o's clustering output with human-labeled data, this study assesses the model's accuracy in detecting entity-centric emphasis frames in media coverage of the 2022 Uvalde school mass shooting.

Accordingly, the following research question is explored:

RQ5: How accurately does a large language model (GPT-4o) cluster framing components used to attribute the perpetrator and victims in the 2022 Uvalde school mass shooting compared to human judgment?

3. Method

3.1. Data

To answer the RQs, this research analyzes media content from nine U.S. news media outlets in the context of a mass shooting in Texas. The shooting left 19 children, and two adults killed at the Robb Elementary School in Uvalde on May 24, 2022 ("Uvalde school shooting," n.d.). Inspired by prior studies (Cassidy et al., 2018;

Chyi & McCombs, 2004), we collected one-month of media content (*N* = 1334) from these purposively-selected nine news media outlets. Three of them were selected from the left-centered bias category: *The New York Times* (n=227), *The Washington Post* (n=228), and *The USA Today* (n=128). Three were selected from the right-centered bias category: *The Wall Street Journal* (n=47), *The New York Post* (n=255), and *The Dallas Morning News* (n=155). And three others were included from the least-biased category: *The Hill* (n=227), *The Indianapolis Star* (n=39), and *The Des Moines Register* (n=28). The media biases were determined by scores provided by Media Bias Fact Check.

After selecting the news media outlets, this study conducted a keyword search—"(Texas OR Uvalde) AND (shoot OR shooting)"—specifying the news outlets' websites (e.g., site:nytimes.com) on Google to identify relevant news reports and their links. Using these links, articles were collected through a combination of automated news scraping via a customized Python script and manual retrieval when subscription barriers applied.

The collected dataset includes various article types, such as news reports, editorials, and opinion pieces. Editorials and opinions were purposively included alongside news reports, considering that news media outlets routinely publish content aligned with their editorial stance. Including these pieces in the framing analysis is crucial as they often reflect the explicit perspectives and ideological leanings of media organizations, shaping audience interpretations. Editorials and opinion articles also play a key role in agenda-setting. After collecting the texts, the documents that contained no text or did not mention "Uvalde," the location of the mass shooting under study, were removed. Finally, the dataset was pre-processed through relevant text-cleaning steps, resulting in a final dataset of 1,334 news media documents for this analysis.

3.2 Analysis

To analyze the data, this research applied the *Semantic Relations-based Unsupervised Framing Analysis* (SUFA) approach. As elaborately explained in Chapter 3, the SUFA approach considers and conceptualizes *"a pair of words in a semantic relation"* as a framing component. After pre-processing the data set, the analysis involved the following key steps that were applied and proposed in Chapter 3 of the research project offering the SUFA approach:

1) Coreference resolution: Given the study's aim to identify modifying words centering around an entity, namely the shooter, coreferences (e.g., "he" or "suspect" for the shooter) needed resolution to capture all possible modifying words of both "references" and "co-references." We accomplished this using NeuralCoref, an extension of the spaCy NLP library, which provides coreference resolution.

2) Token extraction: Next, we employed the dependency parser of spaCy to parse all the news reports and generate a dependency parse tree. This tree outlines the syntactic structure of a sentence, including nodes such as heads (e.g., gunman) and children (e.g., suspect), representing words, and edges representing the semantic relationships between those heads and children. Each edge is labeled with a specific dependency relation, such as "amod" (adjective modifier).

3) Determining keywords: To encompass all possible words referring and co-referring to the shooter and victim entity, we identified relevant keywords (e.g., Salvador, gunman, shooter). These keywords were initially determined based on insights from previous studies and subsequently refined through manual checking of some tokens in the output produced in step 2.

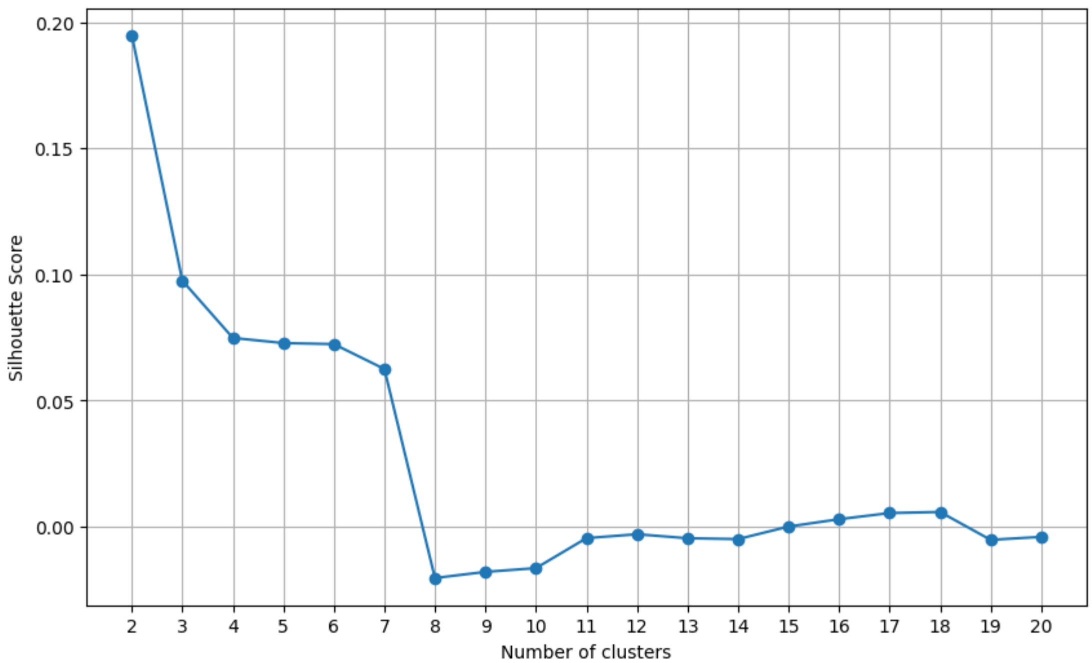
4) Filtering heads and children: Based on the keywords, we filtered out all relevant “heads” and “children” of the shooter and victim entity, along with their dependency relations and associated news outlets.

5) Determining and refining dependency relations: In this step, we determined and refined dependency relations that are pertinent to the study’s objectives. We excluded some dependency relations (e.g., cc, punc) that appeared irrelevant to the research questions through manual checking of the relations grouped in the output produced in step 4.

6) Filtering *framing components*: As elaborated earlier, we considered each pair of words consisting of an entity (e.g., shooter, victims) and a modifying word (e.g., alleged) within a specific dependency relation (e.g., amod) as a framing component providing a particular attribution to an entity. Also, a single framing component was considered a candidate for inclusion in multiple groups (Saldaña, 2021). In this step, we filtered out all such framing components for the shooter entity using pandas. We retained the modifying word that appeared at least more than once, resulting in the modifying tokens (aka framing components) that modify or are associated with the shooter and victims with varied semantic relationships (see Table 4.8 in Appendix 4.3 and Table 4.9 in Appendix 4.4).

7) Clustering *framing components* into frames: The final step includes grouping the framing components into relevant clusters, revealing meaningful framing insights regarding the perpetrator. As mentioned above, automatic clustering is intended as the key purpose of this third project to advance the SUFA approach. It was very challenging to automate the grouping of framing components, which provides meaningful and coherent frames. We tried different computational tools, including *k*-means clustering and DBSscan with BERT embedding vectors and prompt-based LLM, such as OpenAI's GPT-4o.

Figure 4.1. Silhouette Scores for different numbers of clusters



First, we applied the *k*-means clustering algorithm using BERT-based word embeddings to group framing components. To determine the optimal number of clusters, we calculated silhouette scores across a range of cluster counts (see Figure 4.1). The highest silhouette score was observed at $k = 2$, indicating the most distinct separation between clusters at 2. However, when manually examining the cluster outputs across the wider range, including $k = 2$ to 5, the resulting groupings did not reveal meaningful or coherent frame categories. Many clusters mixed semantically unrelated words, and dominant clusters appeared overly broad while others were too sparse, which undermines their framing interpretability. This might stem from the lack of contextual information for each framing component during clustering, a limitation we later attempted to address through prompt-based clustering approaches that explicitly included context.

Second, we applied DBScan clustering algorithm to the BERT-embedded word vectors to identify potential frame groupings. We used an epsilon value of 5 and set the minimum number of samples to approximately 3 to allow for flexible cluster formation. However, the output again failed to reveal coherent or interpretable frame clusters. Visualizing with PCA, the DBScan algorithm produced scattered and poorly differentiated clusters, with seemingly one single cluster of data points. Similar to *k*-means, DBScan also did not yield satisfactory results for clustering framing components into meaningful frames.

Finally, we used the API of OpenAI's GPT-4o, the recently released large language model (LLM). An advantage of LLMs is their ability to follow human instructions and contexts, making them suitable for complex classification tasks like clustering of framing components. This flexibility addresses a key limitation of k-means clustering and DBScan.

Initially, we passed a prompt through OpenAI's GPT-4o API, simply asking the LLM to cluster the shooter-related framing components into an appropriate number of frames. The prompt was entirely unsupervised in a way it did not provide any specific names of potential clusters. However, this GPT returns more than 100 clusters. Framing components of those groups were very scattered. Many clusters include only a few words. Overall, the output largely lacks meaningful and coherent framing insights.

We then customized the prompt by including a comprehensive list of potential clusters for both the shooter category (16 clusters) and the victims category (11 clusters). The lists of clusters were prepared with insights from Project 2 (Chapter 3) and relevant prior studies, including Park et al. (2012) and Holody et al. (2017). The 16 categories for shooter framing include younger age, race/ethnicity, mental health, terrorism/terror, allegation certainty/doubt, and action attribution (see Appendix 4.1 for details). Also, the 11 categories for victim framing include our victims, your victims, innocence and vulnerability, humanization, heroism, and dehumanization (see Appendix 4.2 for details).

7a) Prompts: Through OpenAI's API, we instructed GPT-4o to assign each framing component (i.e., the word modifying the shooter or victim) to one or more relevant clusters. Each input included the token in question along with 40 surrounding tokens from the original sentence, giving GPT-4o contextual information to inform its decisions. These 40 tokens (20 before and 20 after) ensure that the model has some contexts to cluster the framing components in a more informed way. Separate examples of prompts for both the shooter and the victims are in Appendix 4.1 and Appendix 4.2. Two example tokens, along with their contexts, are:

> "City Council member . Everardo Zamora , who represents the district that includes Robb Elementary sister . Ramos , the **ALLEGED** gunman , had

attended Uvalde High School , said Santos Valdez Jr. , 18 , who has known Ramos since”

“Before massacre , **UVALDE** gunman frequently threatened teen girls online . Young people who met Uvalde gunman online said Uvalde gunman had threatened to”

Key observations. GPT-4o performed substantially better than earlier unsupervised clustering attempts applied. Unlike the visually and conceptually inconsistent clusters produced by *k*-means and DBSCAN, the GPT-driven clusters aligned with meaningful and coherent frames observed in prior qualitative and quantitative research. So, this study eventually continued with applying GPT-4o for clustering framing components into frames.

Statistical Analysis

After extracting the framing components (aka modifying words in certain semantic relations with the entity) for both shooter and victims, this study conducted a descriptive analysis of the framing components, chi-square tests, and linear regression analysis using Python to answer the research questions, including the exploration of different types of frames, how significantly the framing components differ across three media groups and how the framing components predict the news media bias.

We conducted regressions twice for each of shooter and victim groups. One regression is conducted with a framing component (e.g., modifying token) as dependent variable (DV) and another one was confuted with a cluster of framing components as DV. For the framing component as DV, we considered only those appears in at least 1% of the articles. The clusters of framing components are already determined above.

In both cases of regression with framing components and clusters as DVs, the study used the media bias score as the independent variable (IV). As mentioned above, with the labels from MBFC, we used the three bias labels of media groups and their bias score as a continuous variable, -1 for left-centered, 0 for least-biased, and 1 for right-centered, for the regressions. Before conducting the analysis, we checked their assumptions. The values (aka frequencies) of framing components and clusters were standardized. We also checked the multicollinearity and removed one variable from the victim group as it was found to be highly correlated.

Evaluation. We evaluated GPT-4o's clustering performance by comparing its output to human-coded clusters of the same set of modifying words. Two graduate students from the communication and information studies disciplines independently labeled each modifying word with one or more clusters. They were given the same prompts provided to the GPT. Intercoder reliability between two coders for each subset was measured using Cohen's Kappa (κ) (Cohen, 1960), with varying levels of agreement observed across different clusters (McHugh, 2012) for both shooter and victim clusters (see Table 4.1 and Table 4.2 below).

Table 4. 1. *Intercoder Reliability of Shooter Framing Components*

Clusters	**Cohen's Kappa**
Mental health	1
Threat	1
Terrorism/Terror	1
Bullying	1
Allegation certainty	1
Allegation doubt	1
Action attribution	1
Younger age	0.935
Older age	0.658
Race/Ethnicity	NA
Religion and culture	NA
Humanization	NA
Immigration	NA
Family background	NA
Stereotypes	NA

Table 4. 2. *Intercoder Reliability of Victim Framing Components*

Clusters	**Cohen’s Kappa (Contextual)**
Innocence and Vulnerability	1
Humanization	1
Younger Age	0.961
Tragic Loss	0.958
Our [Victims]	0.919
Your [Victims]	0.857
Older Age	0.796
Dehumanization	0
Heroism	-0.011
Bullying	NA

Notably, some clusters, such as terrorism and immigration, had missing Kappa values, indicating insufficient overlapping data between the coders to calculate reliability for these categories. Any disagreements were resolved through discussion with the coders (Bracken et al., 2012).

Using the human-labeled clusters as the ground truth, we assessed GPT-4o's performance in terms of precision, recall, and F-score, evaluating how well the model clustered framing components against human coders. This evaluation directly addresses RQ5 by quantifying GPT-4o's clustering accuracy in replicating human annotations.

4. Results

This section addresses both entities: the shooter and the victims. A descriptive analysis of the framing component and subsequent clustering of those framing components by OpenAI GPT-4o into frames revealed different types of frames and their relevant framing components deployed by the nine news media outlets attributing the shooter and victims in the first one-month period following the Uvalde elementary school mass shooting incident.

4.1. RQ1: Types of Frames

The research question 1 focuses on the types of frames that U.S. news media outlets use to attribute the perpetrator in the 2022 Uvalde school mass shooting. The descriptive analysis reveals that the news outlets employed various framing components (e.g., adjectives and adverbs) to attribute the shooter. Around 59% (788) of the articles published by all nine news media outlets used one or more framing components attributing to the shooter. Table 4.8 in Appendix 4.3 lists 164 different words that appeared in the news outlets at least once. At the same time, around 88% (1177) of the articles deployed at least one framing component attributing the victim (see Table 4.9 in Appendix 4.4). This study reports on the clusters of framing components that constitute a cumulative frequency of more than 50 across all three media groups, aiming to gain insights into comparatively more salient frames.

Shooter

The descriptive analysis of the shooter clusters, frequencies, and proportion of their framing components demonstrates that the framing of the shooter in the Uvalde shooting varies across different semantic clusters. The media outlets use various framing components to assign responsibility, portray the shooter, and build narratives around the entity. Based on the GPT-4o-clusters, the shooter framing components dominate in all three media groups: left-centered, least-biased, and right-centered. Considering the reporting of cumulative framing components across media groups, which exceed 50, Table 4.3 below indicates that the news media utilized six different types of frames to attribute to the shooter: i) action attribution, ii) younger age, iii) threat, iv) allegation certainty, v) allegation doubt, and vi) mental health.

i. Action Attribution (922 instances). The action attribution is the dominant frame deployed by news media that emphasizes the shooter's active engagement in the attack. Words such as active (207), shooter (145), and gunman (126) reinforce the idea that the perpetrator was in control and directly responsible for the event. Terms like murderer (71), armed (52), and barricaded (24) further establish his criminality and resistance, while killed (9) and arrested (7) introduce a legal perspective on the consequences of his actions. Among others, definitive action verbs and criminal identifiers were utilized to frame the shooter.

ii. Younger Age (794 instances). This frame highlights the shooter's youth, portraying him as an adolescent or young adult. Tokens such as old (485), 18 (68), and young (43) directly reference his age, reinforcing the idea that he was at the threshold of adulthood. Words like teen (17) and teenage (14) further emphasize youthfulness, subtly positioning the shooter as an immature or impulsive actor rather than a fully accountable adult. The frequent mention of school (148) associates the shooter with an educational environment, reinforcing the common narrative of school shooters as former or current students. Additionally, descriptors like faced (4) and man (4) add complexity, suggesting that media narratives may frame him as someone experiencing challenges or transitioning into adulthood.

iii. Threat (683 instances). The Threat frame portrays the shooter as an immediate and ongoing danger, emphasizing his potential for harm beyond the incident itself. Media coverage describes him as an active shooter (233) and gunman (93), reinforcing a crisis-driven narrative that sustains public fear. Terms like crazed (24), madman (20), and deranged (12) frame him as mentally unstable, linking his actions to erratic or uncontrollable behavior. Additionally, words such as potential (28) and future (7) suggest an extended threat beyond the event, shaping perceptions that shootings like these are part of a broader, ongoing risk.

iv. Allegation Certainty (363 instances). The allegation certainty frame establishes the shooter's definitive involvement in the crime, reinforcing his culpability and legal accountability. Terms like shooter (118), murderer (71), and gunman (44) explicitly identify him as the primary perpetrator, leaving little room for doubt. The inclusion of active (42) and shooting (15) emphasizes the ongoing nature of the attack. This frame removes ambiguity, and the media narratives present the shooter as a known and confirmed offender.

v. Allegation Doubt (185 instances). This frame introduces uncertainty about the shooter's responsibility, reflecting media caution in attributing guilt before official confirmation. Terms such as alleged (106) and suspected (61) indicate that some news reports avoid definitive statements. The words suspect (8) and accused (7) also suggest that the shooter is under investigation rather than conclusively guilty. Less frequent terms such as potential (2) and active (1) further reinforce this ambiguity, ensuring that the framing remains neutral instead of conclusively criminalizing the individual.

vi. Mental Health (110 instances). The mental health frame presents the shooter as psychologically unstable. This seems to reinforce the perception that mental illness might play a key role in mass shootings. Framing components such as deranged (33), crazed (27), and madman (20) depict the shooter as irrational and out of control. This aligns with broader media narratives that link mass violence to individual psychological disorders. Framing components like angry (6) and irate (6) further emphasize emotional instability and impulsiveness. The presence of terms like lone (2) and potential (2) may also contribute to the "lone wolf" narrative.

Victims

The descriptive analysis of the victim clusters, frequencies, and proportions of their framing components unveils a range of framing components and frames deployed by news outlets to attribute victims. Based on the GPT-4o-clusters, the victim framing components are utilized across all three media groups: left-centered, least-biased, and right-centered. As noted above, considering cumulative framing components greater than 50, Table 4.4 shows that the news media employed nine different types of frames, including i) dehumanization, ii) younger age, iii) innocence and vulnerability, iv) tragic loss, v) your [victims], vi) our [victims], vii) humanization, viii) older age, and ix) heroism to frame the victims.

i) Dehumanization (2310 instances). The dehumanization identified victims using more statistics than individuals, which has seemingly reduced their identities to numerical figures and collective terms. High-frequency numerical references such as 19 (1,251), two (258), and dead (158) emphasize the scale of the tragedy. Additionally, framing components like killed (30), slaughtered (11), and deceased (4) reinforce the finality of death but strip victims of humanizing details. By framing victims in impersonal and quantitative terms, this cluster may contribute to psychological distancing, shifting media discourse toward systemic issues rather than individual narratives of suffering.

ii) Younger Age (1480 instances). The news outlets emphasize the youth of the victims, reinforcing their vulnerability and innocence. High-frequency framing components such as school (337), grade (175), and elementary (62) situate the victims within an educational setting, underscoring the tragedy of children being targeted in a place meant for learning and safety. References to age, including young (128), old (96), 19 (73), and 18 (12), highlight the victims' early developmental stages. This amplifies the sense of injustice and loss through the lens of youth. At the same time, framing components like little (47), child (7), students (9), and sweet (4) evoke emotional responses by portraying the victims as innocent and defenseless. The presence of familial terms such as daughter (14), parents (7), and mother (4) reinforces the personal dimension of loss. Overall, contrasting with the dehumanization frame that presents them as statistics, this frame humanizes the victims by centering on their youth and relationships.

iii) Innocence and Vulnerability (1287 instances). Victims' humanization was further emphasized by highlighting their vulnerability and innocence. In this context, news outlets focus on the victims' defenselessness and emotional impact. High-frequency framing components such as innocent (85), young (125), terrified (32), and little (47) underscore their fragility and reinforce a sense of injustice. References to school (286), grade (62), and elementary (49) place the victims within a learning environment. This alignment further underscores the severity of the shooting. Words like wounded (19), shot (18), and slaughtered (10) evoke distressing imagery, while daughter (9), children (7), and students (5) strengthen emotional connections by framing victims as loved ones rather than mere anonymous casualties.

iv) Tragic Loss (1099 instances). The victims were portrayed as deeply mourned individuals. Framing components such as killed (304), dead (199), murdered (50), slain (50), and died (50) emphasize the high level of fatality and loss. Modifying components like beloved (11), desperate (11), and beautiful (2) focus on the victims' personal and sentimental dimensions. This further shifts the framing from statistical death counts to individual suffering. Additionally, references to violence (6), massacre (10), and massacred (10) position the event within the broader discourse of gun violence and mass shootings. Overall, this frame blends personal grief with systemic issues that highlight the shooting's potential broader societal impact.

v) Your [Victims] (768 instances). In this frame, the news outlets personalize the tragedy by establishing relational connections between the victims and the audience. Framing components like their (171), your (123), her (108), and his (39) emphasize possession and familiarity, fostering a sense of shared loss and collective mourning. Familial terms like parents (87), mother (8), daughter (7), and family (4) further reinforce an emotional connection by centering the victims within recognizable personal relationships. The inclusion of words such as fellow (22), own (20), and people (4) also suggests that the victims are not just statistics but individuals who belonged to families, communities, and shared social spaces.

vi) Our [Victims] (561 instances). This frame further reinforces collective grief and solidarity, portraying the victims as part of a shared national or communal loss. Frequent pronouns like our (269), my (190), and we (16) establish a sense of belonging, framing the tragedy as one that affects not just individual families but society as a whole. The presence of familial and community-related terms such as parents (5), daughter (5), and fellow (4) further strengthens this connection.

vii) Humanization (504 instances). This frame portrays victims as individuals with personal identities, relationships, and histories. Framing components such as grade (105), school (21), and teacher (12) ground the victims in their daily lives. This emphasizes the devastating impact of the shooting on the learning environment. Familial references like parents (20), mother (12), and daughter (12) also create an emotional connection and depict the personal loss experienced by families. At the same time, modifiers such as beloved (11), beautiful (8), outgoing (8), and sweet (7) highlight the significance of victims' positive personalities. Additionally, references to roles such as teacher (12), substitute (6), retired (6), and hero (6) elevate certain victims as figures of sacrifice and dedication.

viii) Older Age (365 instances). In this frame, some victims were highlighted as relatively mature and more experienced. Framing components such as old (60), older (10), and aged (2) emphasize maturity, while longtime (8), retired (7), and veteran (6) frame them as figures of wisdom and service. References to teacher (6) and taught (2) also frame them as educators.

ix) Heroism (110 instances). Though this frame is comparatively smaller in the number of framing components, its significance surpasses that of some other frames. Here, news media elevate specific victims by portraying them as courageous figures who acted selflessly in the face of tragedy. Framing components such as hero (10), leading (10), and called (18) highlight their heroic and praiseworthy actions. Modifying words like teacher (5), armed (3), and veteran (2) also depict them as actively engaged in protective efforts.

4.2. RQ2: Extent of Frames

The research question 2 aims to address the extent to which the U.S. news media outlets deploy framing components to attribute the shooter and the victims. As noted above, we considered the cumulative frequency of at least 50 for all the framing components in each cluster for reporting the frames. With this consideration, the analysis reveals that the news media outlets, ranging from left- to right-leaning biases, deployed at least six types of frames toward the shooter and at least nine types of frames toward the victims.

The extent of framing varied among media groups. Right-centered outlets focused heavily on action attribution, threat, and certainty of allegations against the shooter, while left-centered outlets engaged more with doubt in allegations and mental health. When covering victims, all media groups deployed dehumanization most frequently, but left-centered media emphasized innocence, tragic loss, and personal connection, while right-centered outlets used humanization and highlighted older age more often, as depicted in Table 4.3 and Table 4.4 and elaborated below.

Shooter Framing Components. The shooter was framed most frequently through action attribution (922 times) (see details in Table 4.3), with right-centered media deploying this frame in the highest proportion of articles (1.18 per article in that media group), followed by left-centered media at 0.48 per article. As visualized in Figure 4.2, the proportion is calculated by dividing the number of framing components in a cluster by the total number of articles across clusters in the respective media group. This suggests a strong emphasis on portraying the shooter as an active agent in the violence, particularly in right-leaning coverage. The younger age frame appeared 794 times, with left-centered media (0.41 per article in that media group) and right-centered media (0.94 per article in that media group) both highlighting the shooter's youth. The Threat frame (683 instances) was also prominent, with right-centered media (0.74 per article in that media group) utilizing this frame more frequently than other groups, reinforcing a narrative of imminent danger.

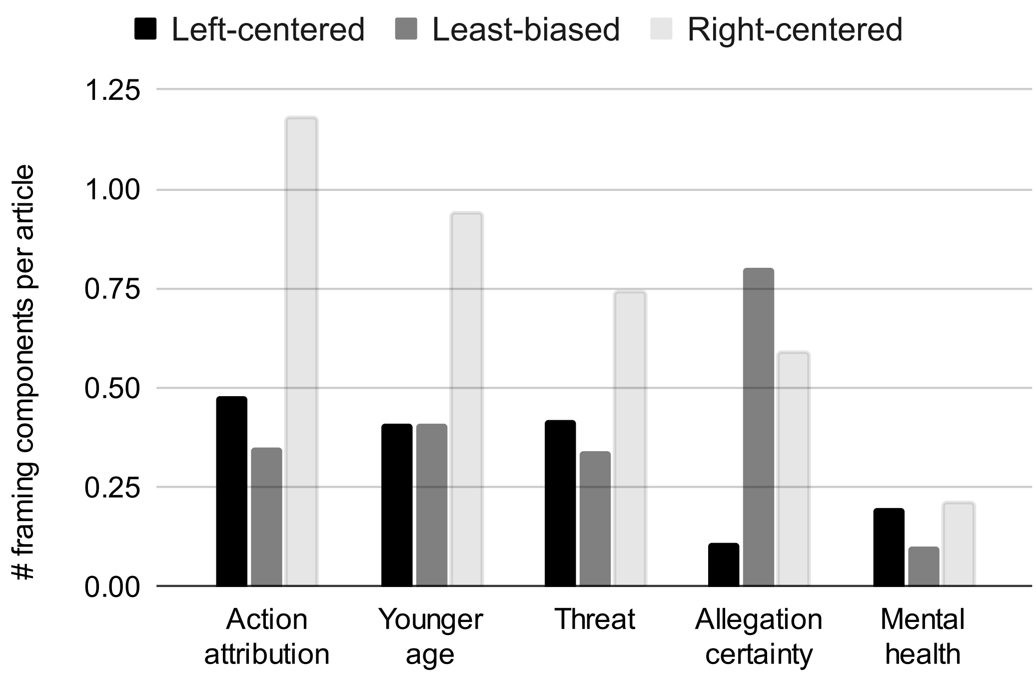
Regarding certainty in allegations, the allegation certainty frame was utilized 363 times, with right-centered media once again leading (0.59 per article within that media group), reinforcing a definitive stance on the shooter's culpability. However, the allegation doubt frame (185 times) exhibited a more cautious approach, with left-centered media employing it more frequently (0.12 per article within that media group) than right-centered outlets (0.15 per article), suggesting differences in reporting styles based on media biases. Mental health framing appeared the least (110 times), yet right-centered media (0.21 per article) applied it more often than other groups, potentially linking the shooter's actions to psychological instability.

Table 4. 3. *Frequency of Framing Components of* <u>*Shooter*</u> *among Three Media Groups—Left-Centered, Least-Biased, and Right-Centered*

Clusters	**Left-centered Freq**	**Least-biased Freq**	**Right-centered Freq**	**Total**
Action attribution	279	102	541	922
Younger age	241	122	431	794
Threat	244	99	340	683
Allegation certainty	66	27	270	363
Allegation doubt	70	45	70	185
Mental health	12	4	94	110

Note: The total frequency of more than 50 is reported.

Figure 4.2. The proportion of framing components of shooter among the three media groups.



Note: The proportion is calculated by dividing the number of framing components in a cluster by the total number of articles across clusters in the respective media group.

Victim Framing Components. The victims were framed most frequently through dehumanization (2310 instances) (see details in Table 4.4), which was used substantially in right-centered media (1.92 per article in that media group) and left-centered media (1.72 per article). As visualized in Figure 4.3, the proportion is calculated by dividing the number of framing components in a cluster by the total number of articles across clusters in the respective media group. This frame, which largely reduces victims to statistics, was more pronounced than any other category. The younger age frame (1480 times) was also highly prevalent, with left-centered media (1.18 per article in that media group) leading in its use. This reinforces the youth and innocence of the victims.

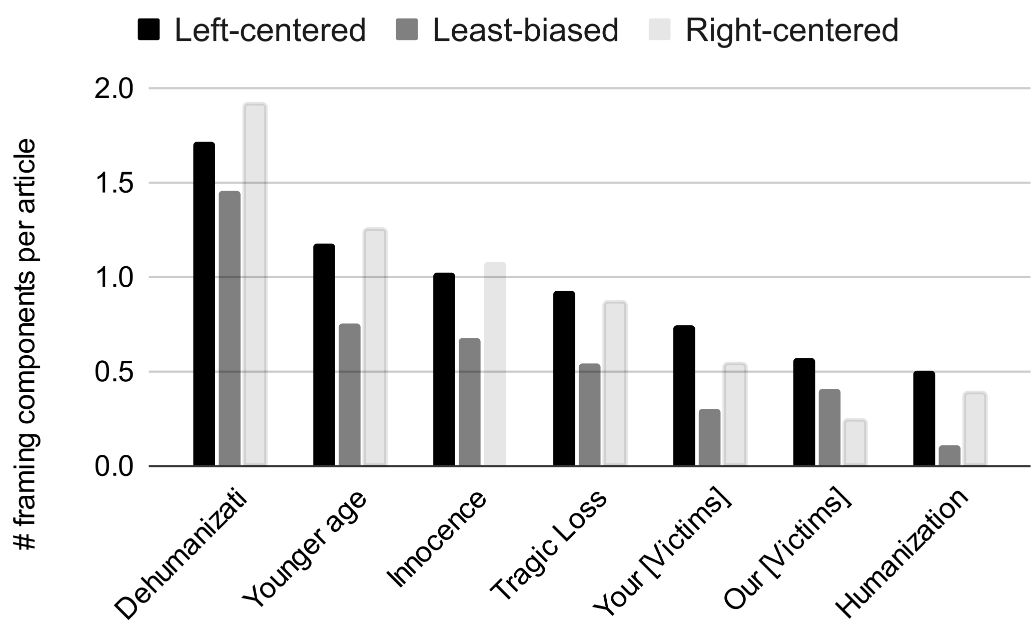
Table 4. 4. *Frequency of Framing Components of* <u>*Victim*</u> *among Three Media Groups—Left-Centered, Least-Biased, and Right-Centered*

Clusters	**Left-centered Freq**	**Least-biased Freq**	**Right-centered Freq**	**Total**
Dehumanization	1005	428	877	2310
Younger age	690	220	570	1480
Innocence and Vulnerability	595	200	492	1287
Tragic Loss	545	158	396	1099
Your [Victims]	431	88	249	768
Our [Victims]	333	120	108	561
Humanization	293	31	180	504
Older Age	192	60	113	365
Heroism	46	19	45	110

Note: The total frequency of more than 50 is reported.

The frames of innocence and vulnerability (1287 instances) and tragic loss (1099 instances) were both widely utilized, with left-centered media leading in these categories (1.02 and 0.93, respectively, per article in the media groups), highlighting emotional and humanizing narratives. Personal connection frames, including “your victims” (768 instances) and “our victims” (561 instances), were more frequently employed by left-centered outlets (0.74 and 0.57, respectively, per article in the media groups), suggesting an effort to foster communal mourning.

Figure 4.3. The proportion of framing components of victims among the three media groups.



Note: The proportion is calculated by dividing the number of framing components in a cluster by the total number of articles across clusters in the respective media group.

4.3. RQ3: Statistical Significance of Framing Differences among Media Groups

RQ1 and RQ2 described differences in framing components and frames across media groups. Now, RQ3 examines whether these differences are statistically significant. To answer this question, this study conducted separate chi-square analyses to determine how frames varied across left-centered, least-biased, and right-centered media outlets. These tests assess whether media groups significantly differ in their use of internal and external attribution frames when covering the shooter and victims.

The chi-square analysis was conducted using proportional data instead of total frequency counts. This approach mitigates potential biases caused by variations in the number of published articles across media groups. The proportion is calculated by dividing the number of framing components in each cluster by the total number of articles containing framing components within the respective media group. This ensures that differences in framing are assessed based on relative emphasis rather than absolute coverage volume.

Shooter

The results reveal statistically significant differences across multiple framing categories, indicating that media bias plays a role in shaping the portrayal of the shooter, as presented below (also see Table 4.5).

The differences in deploying various shooter framing components among the three media groups were found to be statistically significant in most clusters, including action attribution, younger age, threat, allegation certainty, mental health, and race/ethnicity.

Action attribution. The action attribution frame was used most frequently by right-centered media (118%), followed by left-centered media (48%) and least-biased outlets (35%), with a statistically significant difference ($\chi^2 = 59.49$, $p < 0.001$). Pairwise comparisons show significant differences between left-centered and right-centered media ($\chi^2 = 29.52$, $p < 0.001$) as well as between least-biased and right-centered media ($\chi^2 = 45.03$, $p < 0.001$). Therefore, the right-centered media deployed significantly more action attribution frames compared to both least-biased and left-centered media.

Table 4. 5. *Chi-square analysis of <u>shooter</u> framing components across media groups—left-centered (left), least-biased (least), and right-centered (right)*

	Observed				Overall			
Clusters	**Left**	**Least**	**Right**	**Expected**	**chi-square**	**Left-Least**	**Left-Right**	**Least-Right**
Action attribution	48	35	118	67	59.49***	2.04	29.52***	45.03***
Younger age	41	41	94	58.67	31.92***	0	20.81***	20.81***
Threat	42	34	74	50	17.92***	0.84	8.83**	14.81***
Allegation certainty	11	9	59	26.33	60.86***	0.2	32.91***	36.76***
Allegation doubt	12	15	15	14	0.43	0.33	0.33	0
Mental health	2	1	21	8	31.75***	0.33	15.7***	18.18***
Older age	5	0	4	3	4.67	5.0*	0.11	4.0*
Race/ethnicity	6	1	1	2.67	6.25*	3.57	3.57	0
Family background	0	0	3	1	6.0*	nan	3	3
Humanization	0	1	2	1	2	1	2	0.33
Stereotypes	1	0	1	0.67	1	1	0	1
Terrorism/terror	2	0	0	0.67	4	2	2	nan

Bullying	1	0	1	0.67	1	1	0	1

Note: ***p < 0.001, **p < 0.01, *p < 0.05.

The proportion is calculated by dividing the number of framing components in a cluster by the total number of articles across clusters in the respective media group. To clarity, 75% means 75 times per 100 articles.

Younger age. Similarly, the younger age frame appeared more frequently in right-centered coverage (94%) compared to left-centered (41%) and least-biased (41%), showing significant variation ($\chi^2 = 31.92$, $p < 0.001$). Pairwise comparisons indicate significant differences between least-biased and right-centered media ($\chi^2 = 20.81$, $p < 0.001$) as well as between left and right-centered media ($\chi^2 = 20.81$, $p < 0.001$). Thus, this frame was also utilized by right-centered media significantly more than the other two.

Threat. The threat frame was also more prevalent in right-centered media (74%) than in left-centered (42%) and least-biased (34%) outlets ($\chi^2 = 17.92$, $p < 0.001$). Right-centered media differed significantly from least-biased ($\chi^2 = 14.81$, $p < 0.001$) and left-centered ($\chi^2 = 8.83$, $p < 0.01$) media.

Allegation certainty. The allegation certainty frame also showed substantial variation, with right-centered outlets using it at 59%, compared to 11% in left-centered media and 9% in least-biased media ($\chi^2 = 60.86$, $p < 0.001$). Significant differences were revealed between left and right-centered media ($\chi^2 = 32.91$, $p < 0.001$) and between least-biased and right-centered media ($\chi^2 = 36.76$, $p < 0.001$).

Thus, right-centered media utilized a significantly higher allegation certainty frame than the other two groups.

Mental health. The mental health frame was more frequently employed by right-centered media (21%) than by left-centered (2%) or least-biased (1%) media, with significant differences ($\chi^2 = 31.75$, $p < 0.001$). Pairwise comparisons show that right-centered media differed significantly from both left-centered ($\chi^2 = 15.7$, $p < 0.001$) and least-biased ($\chi^2 = 18.18$, $p < 0.001$) outlets.

Older age. The older age frame also varied across groups, with left-centered outlets utilizing it slightly more than right-centered and least-biased media. Although this difference did not demonstrate statistical significance, pairwise comparisons reveal a marginally significant difference between left-centered and least media ($\chi^2 = 5.0$, $p < 0.05$), and between least and right-centered media ($\chi^2 = 4.0$, $p < 0.05$).

Race/ethnicity and Family background. The race/ethnicity and family background frames showed some variation. However, their pairwise comparisons did not yield statistically significant differences. Frames such as allegation doubt, humanization, stereotypes, terrorism/terror, and bullying did not demonstrate statistically significant differences across media bias groups.

Victims

Similar to the shooter chi-square results, the victim framing results show statistically significant differences across most framing categories, as presented below (also see Table 4.6).

The dehumanization frame was the most frequently used, with right-centered media (192%) and left-centered media (172%). The chi-square test shows a significant difference across media groups ($\chi^2 = 6.26$, $p < 0.05$), with a significant difference between least-biased and right-centered outlets ($\chi^2 = 6.26$, $p < 0.05$).

Table 4. 6. *Chi-square analysis of victim framing components across media groups—left-centered (left), least-biased (least), and right-centered (right)*

Clusters	Observed Left	Least	Right	Expected	Overall chi-Square	Left-Least	Left-Right	Least-Right
Dehumanization	172	146	192	170	6.26*	2.13	1.1	6.26*
Younger age	118	75	125	106	13.83***	9.58**	0.2	12.5***
Innocence and vulnerability	102	68	108	92.67	10.04**	6.8**	0.17	9.09**
Tragic loss	93	54	87	78	11.31**	10.35**	0.2	7.72**
Your [victims]	74	30	54	52.67	18.43***	18.62***	3.12	6.86**
Our [victims]	57	41	24	40.67	13.39**	2.61	13.44***	4.45*
Humanization	50	11	39	33.33	24.26***	24.93***	1.36	15.68***
Older age	33	20	25	26	3.31	3.19	1.1	0.56
Heroism	8	6	10	8	1	0.29	0.22	1
Bullying	0	0	1	0.33	2	nan	1	1

Note: ***p < 0.001, **p < 0.01, *p < 0.05.

The proportion is calculated by dividing the number of framing components in a cluster by the total number of articles across clusters in the respective media group. To clarity, 120% means 120 times per 100 articles.

The younger age frame was applied significantly more in right-centered coverage (125%) than in left-centered (118%) and least-biased (75%) outlets ($\chi^2 = 13.83$, $p < 0.001$). Pairwise comparisons find significant differences between least-biased and right-centered outlets ($\chi^2 = 12.59$, $p < 0.001$) and between left-centered and least-biased outlets ($\chi^2 = 9.58$, $p < 0.01$). Left-centered and right-centered media did not significantly differ. So, there is no difference in using this frame between left-centered and right-centered outlets.

Like younger age, the innocence and vulnerability frame varied across media bias groups, with left-centered media (102%) and right-centered media (108%) using this frame more frequently than least-biased outlets (68%), yielding a significant chi-square result ($\chi^2 = 10.04$, $p < 0.01$). Also, significant differences were found between least-biased and right-centered media ($\chi^2 = 9.09$, $p < 0.01$) and least-biased and left-centered outlets ($\chi^2 = 6.98$, $p < 0.01$).

The tragic loss frame was more prevalent in left-centered (93%) and right-centered (87%) media compared to least-biased (54%), with a significant chi-square result ($\chi^2 = 11.31$, $p < 0.01$). Pairwise comparisons revealed that least-biased media differed significantly from right-centered ($\chi^2 = 7.72$, $p < 0.01$) and left-centered ($\chi^2 = 10.35$, $p < 0.01$) outlets. Furthermore, left-centered and right-centered media were not significantly different in this frame either.

The "your victims" frame was more common in left-centered media (74%) than in least-biased (30%) and right-centered (54%) media ($\chi^2 = 13.43$, $p < 0.001$). Pairwise comparisons show significant differences between left-centered and least-centered media ($\chi^2 = 18.62$, $p < 0.001$) and between least and right-centered outlets ($\chi^2 = 6.86$, $p < 0.01$), while the difference between left and right-centered media was not significant.

The “our victims” frame is also used differently, with left-centered outlets (57%) employing it more frequently than least-biased (41%) and right-centered (24%) media ($\chi^2 = 13.39$, $p < 0.001$). Pairwise comparisons indicate significant differences between left-centered and right-centered media ($\chi^2 = 13.44$, $p < 0.001$) and between right-centered and least-biased outlets ($\chi^2 = 4.5$, $p < 0.05$). In both “your victims” and “our victims,” the left-centered media utilized significantly higher frames compared to right-centered media.

The humanization frame was used most frequently by left-centered media (50%), followed by left-centered (39%) and least-biased (11%), with a significant chi-square result ($\chi^2 = 24.26$, $p < 0.001$). In pairwise comparisons, the use of this frame was significant between left-centered and least-biased media ($\chi^2 = 24.93$, $p < 0.001$) and between least-biased and right-centered outlets ($\chi^2 = 15.68$, $p < 0.001$), while the differences between left-centered and right-centered media were not significant.

The chi-square test did not reveal significant differences in the use of older age and heroism frames across the media groups.

4.4. RQ4: Predicting media bias by framing components and clusters

The research question 4 asks to explore how different framing components predict the media bias score of U.S. news media outlets covering the 2022 Uvalde school mass shooting. As noted above, separate OLS regressions examine how 1) individual framing components and 2) clusters of framing components related to both shooter and victim predict media bias scores. The media bias scores range from -1 (left-centered), 0 (least-biased), to 1 (right-centered). We conducted four regressions—two for the shooter with individual framing components (IV) and then clusters of framing components (IV). Two additional regressions are for the victim with individual framing components (IV) and then clusters of framing components (IV). In all cases, the dependent variable is the media bias score.

Shooter (IV: framing components)

The OLS regression model analyzes how individual framing components associated with the shooter predict media bias scores. As shown in Table 4.10 (Appendix 4.5), this model explains 17.3% of the variance in media bias ($R^2 = 0.173$), indicating a statistically significant overall effect ($F = 4.755$, $p < 0.001$).

Several framing components emerged as significant predictors of media bias scores. Media bias was found to be positively associated with the use of four framing components: “shooter” ($B = 2.803$, $p < 0.001$), gunman ($B = 1.958$, $p < 0.001$), teen ($B = 1.278$, $p = 0.036$), and deranged ($B = 1.053$, $p = 0.017$). In other words, outlets with a more right-leaning bias are more likely to employ these specific framing components in their coverage.

Conversely, media bias scores were found to be negatively associated with the use of framing components: white ($B = -0.916$, $p = 0.01$), active ($B = -1.085$, $p = 0.012$), and shooting ($B = -0.789$, $p = 0.041$). The results suggest that media outlets with a more left-leaning bias are more likely to utilize these three specific framing components.

Table 4. 7. *Regression Analysis: Predicting Media Bias Scores Based on* *Shooter Clusters*

Effect	Estimate (B)	SE	t	p	95% CI LL	95% CI UL
Intercept	-0.0895	0.044	-2.036	0.042	-0.176	-0.003
Younger age	1.8062	0.403	4.482	0	1.015	2.597
Older age	-1.181	0.777	-1.521	0.129	-2.706	0.344
Race/ethnicity	-1.1773	0.347	-3.397	0.001	-1.858	-0.497
Mental health	3.1887	0.966	3.302	0.001	1.293	5.085
Religion and culture	1.0618	0.737	1.441	0.15	-0.385	2.508
Humanization	-0.3439	0.851	-0.404	0.686	-2.015	1.327
Threat	-0.711	0.573	-1.242	0.215	-1.835	0.413
Terrorism/terror	-1.2964	0.523	-2.479	0.013	-2.323	-0.27
Family background	1.1732	0.947	1.239	0.216	-0.686	3.032
Stereotypes	-0.4962	0.692	-0.717	0.474	-1.855	0.863
Bullying	-1.1028	0.624	-1.766	0.078	-2.329	0.123
Allegation doubt	-0.2587	0.45	-0.575	0.566	-1.142	0.625
Allegation certainty	2.3881	0.966	2.472	0.014	0.492	4.285

DV: Media bias score—left (-1), least (0), and right (1)

R-squared: 0.087

Adj. R-squared: 0.070

Df Model: 13

F-statistic: 5.092

Prob (F-statistic): < 0.001

N = 698

Note: CI = Confidence Interval, LL = Lower Limit, UL = Upper Limit. Bias scores range from -1 (left-centered) to 1 (right-centered).

Shooter (IV: clusters)

Another regression examines how the clusters of the shooter framing components predict media bias. The model accounts for 8.7% of the variance in media bias ($R^2 = 0.087$, $p < 0.001$), as presented in Table 4.7.

Several clusters of the framing components related to the shooter significantly predict and shape the level of media bias in the United States. As the results show, the media bias scores were positively associated with and predicted by three clusters of framing components: younger age ($B = 1.806$, $p < 0.001$), mental health ($B = 3.188$, $p = 0.001$), and allegation certainty ($B = 2.388$, $p = 0.014$). In other words, right-leaning news outlets are more likely to use framing components related to younger age, mental health, and allegation certainty.

On the other hand, the media bias score was negatively predicted by race/ethnicity ($B = -1.177$, $p = 0.001$) and terrorism/terror ($B = -1.2964$, $p = 0.013$). In other words, left-leaning news outlets are more likely to associate the shooter with race/ethnicity- and terrorism-related frames.

Victim (IV: framing components)

This regression model examines how individual framing components relating to victims predict media bias scores. As presented in Table 4.11 (Appendix 4.5), the model explains 10.8% of the variance in media bias ($R^2 = 0.108$, $p < 0.001$). Several individual framing components associated with victims significantly predict the level of news media bias.

The use of framing components “our” ($B = -1.155$, $p = 0.025$), “your” ($B = -1.283$, $p = 0.032$), and “his” ($B = -1.186$, $p = 0.026$) when referring to victims negatively predicts media bias, indicating that left-centered media outlets are more likely to employ this personalized framing component. Additionally, references to “parents” framing components related to victims were used more frequently by left-centered media ($B = -1.2698$, $p = 0.012$). In contrast, portraying victims with “terrified” framing components positively predicts media bias ($B = 1.543$, $p = 0.043$). In other words, right-centered media are more likely to use this framing component, portraying victims as afraid and helpless.

Victim (IV: clusters)

Another regression examines how the clusters of victim framing components predict media bias scores. The model explains 2.3% of the variance in media bias ($R^2 = 0.023$, $p < 0.05$), as presented in Table 4.12 (Appendix 4.5). The results show that at least four victim frames significantly predict media bias in the United States.

The level of media bias was negatively predicted by three victim frames: older age ($B = -2.117$, $p = 0.024$), our [victims] ($B = -1.8296$, $p = 0.001$), and your [victims] ($B = -1.499$, $p = 0.012$), and positively predicted by one victim frame, namely dehumanization ($B = 1.8085$, $p = 0.04$). In other words, left-centered media are significantly more likely to use personalized victim framing (our victims, your victims) and emphasize older victims. In contrast, right-centered media more frequently use dehumanization framing. The more left-leaning the news media are, the more they use personalized victim frames; conversely, the more right-leaning the news media are, the more they deploy the dehumanization frame.

4.5. RQ5: GPT's Clustering Performance

To evaluate the clustering accuracy of GPT-4o, the model-generated clusters were compared to human-labeled clusters using precision, recall, and F1-score. The analysis was conducted separately for shooter framing components (with and without context) and victim framing components (without context). The context was considered to be 20 tokens (where available) surrounding the modifying token. Cohen's Kappa scores are included in Table 4.13, Table 4.14, and Table 4.15 in Appendix 4.6 for transparency, reflecting intercoder agreement between two human coders rather than GPT-4o performance. The human-labeled clusters served as the ground truth for evaluating GPT-4o's clustering performance.

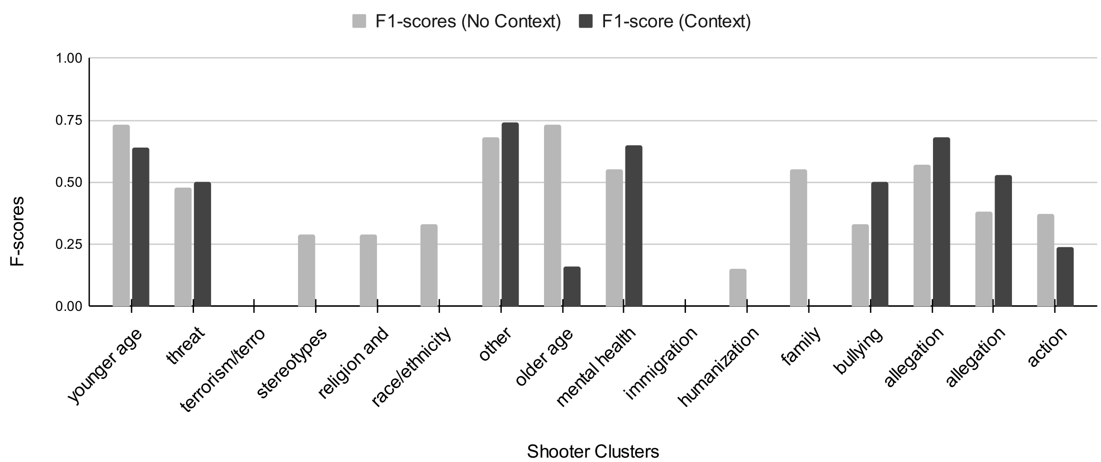
GPT-evaluation for shooter framing components

The GPT-evaluation results reveal that this LLM performed well in capturing certain shooter frames, with F-scores ranging from 0.73 to 0.33, as visualized in Figure 4.4. For example, the framing clusters "older age" and "young age" show high performance with an F1 score of 0.73. Also, the framing clusters "allegation doubt," "family background," and "mental health" performed well, with their F-1 scores of 0.57 and 0.55, respectively.

As the results demonstrated, including context (surrounding words) led to visible changes in GPT-4o's performance. For example, the F-scores of clusters of "younger age" and "older age" were highest at 0.73 when context was not provided.

However, these scores declined to 0.64 and 0.16, respectively. On the other hand, F-scores of clusters like “allegation doubt” improved from 0.57 to 0.68 when context was provided. Overall, the results indicated that context appears to have helped recall but reduced precision in some cases, indicating more misclassifications in the contextual clustering.

Figure 4.4. GPT performance evaluated against judgments for shooter-related framing components (with vs. without context)

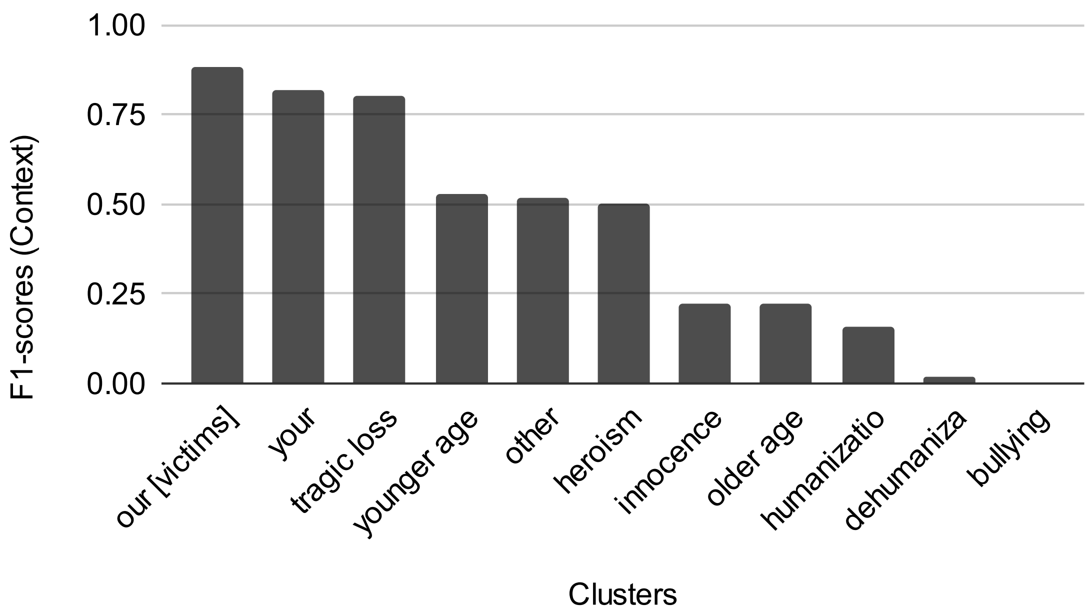


GPT evaluation for victim framing components

For victim-related framing components, GPT-4o demonstrated high performance against human-labeled clusters in certain categories and struggled in others, as visualized in Figure 4.5. For example, clusters like “our [victims]” (F1 = 0.88), “your [victims]” (F1 = 0.82), and “tragic loss” (F1 = 0.80) showed strong performance. The clusters of “younger age” (F1 = 0.53) and heroism (F1 = 0.50) also showed solid performance. Although the cluster of “innocence and vulnerability” has an F1 score of 0.22, its recall (0.74) is higher.

Overall, GPT-4o demonstrates moderate to high accuracy in clustering shooter and victim framing components compared to human judgment. The results indicate that the LLM performs better in seemingly well-defined and frequently occurring clusters such as “shooter,” “young age,” “tragic loss,” and “personalized references to victims.” It likely struggles with more nuanced and abstract categories such as humanization and dehumanization.

Figure 4.5. GPT performance evaluated against judgments for victims-related framing components (with context)



5. Discussion and Conclusion

This third study in Chapter 4 of this dissertation examines how U.S. news media outlets frame the mass shooter and victims in the context of the 2022 Uvalde school mass shooting. Importantly, this chapter applied and advanced a novel computational framing analysis approach, *Semantic Relations-based Unsupervised Framing Analysis (SUFA)*, proposed in the previous Chapter 3, using a large dataset of 1334 articles from nine United States news media outlets. As part of this unsupervised framing analysis approach, the study also assessed the effectiveness of a large language model (LLM), such as OpenAI's GPT-4o, in clustering framing components into frames. The inclusion of the LLM in the automatic clustering process enhanced the SUFA approach, addressing the approach's clustering limitation mentioned in the previous Chapter 3.

The results explored through the advanced SUFA approach reveal deeper insights into different frames and their comparisons among three media groups: left-centered, least-biased, and right-centered. Briefly, the results unveiled at least six frames attributed to the shooter and nine frames attributed to the victims in various ways. The novelty of the semantic relations-based approach, which goes beyond the existing bag-of-words method, helps to retrieve deeper insights into the frames, rather than just the topics. For example, the attribution frame includes 26 unique words that appeared 922 times while attributing various semantic relations to the shooter. Unlike the topic modeling-based framing analysis approach, the insights of semantic relations on how the shooter was associated with these 26 unique words help to capture the informed meanings and relationships between the shooter and modifiers. In this action attribution frame, the shooter was found to be framed as armed and actively engaged in the attack, murdering students. This insight into how the shooter was framed is well-informed by the semantic relations, which are absent in the existing unsupervised framing analysis approach based on topic modeling.

Exploring framing components and frames through the SUFA approach, this chapter further employs chi-square and regression analysis to gain a deeper understanding of how significantly the frames differ across news media groups and how framing components predict the level of news media bias. A few key results identify that right-centered media deployed certain frames, including action attribution, younger age, and allegation certainty, to a significantly higher extent compared to left-centered media. The regression analysis shows that left-centered media are significantly more likely to employ personalized victim framing (our victims, your victims) and to emphasize older victims. In contrast, right-centered media more frequently use dehumanization framing. There is also a significant difference in how right-leaning and left-leaning news outlets utilize various framing components, including younger age, mental health, allegation certainty, humanization, and personalization of victims. These findings highlight the ideological differences in how news media frame victims and the shooter.

Attribution theory suggests that individuals assign causes to behaviors through either internal attribution (e.g., personal traits, intentions, or choices) or external attributions (e.g., environmental or situational influences) (Heider, 1958; Kelley, 1973). Frames that emphasize internal attributions, such as personal responsibility, often elicit moral outrage, demands for punishment, and calls for individual accountability. In contrast, frames that highlight external attributions, such as situational or environmental factors, tend to foster calls for structural change, policy interventions, and prevention efforts.

For example, the “action attribution” frame, the most frequently used frame in media coverage (922 instances), emphasizes the shooter’s agency and control over the violence. Words such as active, gunman, and murderer attribute direct responsibility to the perpetrator, reinforcing the concept of internal attribution (Heider, 1958). Right-centered media deploy this frame more frequently (1.18), suggesting a potential influence on shaping public perceptions and driving calls for action on mass shootings, as theorized by Entman (1993).

By reinforcing the shooter's personal culpability, the action attribution frame in right-centered media (1.18 per article) enhances moral outrage and support for punitive measures. Psychological research on just-world theory (Lerner & Miller, 1978) suggests that when individuals are perceived as having complete control over their actions, the public tends to exhibit less empathy and become more inclined to demand severe retribution. Simultaneously, the allegation certainty frame (murderer, shooter, gunman), commonly used by right-centered media, reinforces absolute culpability, diminishing public uncertainty regarding the shooter's responsibility. As an example of internal attribution (Heider, 1958), this framing can lead to quicker calls for legal action, public demonstrations, and stricter sentencing expectations.

Additionally, right-centered media placed significant emphasis on the shooter's mental instability (deranged, crazed, madman), suggesting that his actions were driven by psychological illness rather than premeditated intent. This framing reduces personal blame, shifting the focus from individual punishment to policy discussions on mental health treatment, early intervention, and institutional support. According to cognitive dissonance theory (Festinger, 1957), when people perceive an act of violence as senseless due to mental illness, they experience less direct anger toward the perpetrator and greater concern for preventing similar incidents in the future.

The dehumanization frame, widely deployed by both right-centered and left-centered media, may act as an external attribution by psychologically distancing audiences from the tragedy. Research on psychological numbing (Slovic, 2007) suggests that when victims are reduced to statistics rather than depicted as individuals with personal stories, audiences feel less empathy and are less likely to take action. From an attribution theory perspective, this could explain why mass shootings often evoke short-term public outrage but limited long-term policy change. Further research from a psychological perspective could provide deeper insights into potential solutions for reducing gun violence.

Conversely, personalized frames such as "your victims" and "our victims", widely used by both left- and right-centered media, establish a personal connection between the audience and the tragedy. According to social identity theory (Tajfel & Turner, 1979), people are more likely to take collective action when they perceive a tragedy as affecting "their" group rather than a distant one. This personalized framing reinforces a sense of collective responsibility and urgency for action.

Ultimately, the varying deployment of these frames across the three media outlets demonstrates their role in shaping public perceptions and reactions to mass shootings. By emphasizing internal or external attributions, media frames potentially influence how the public responds with demands for punitive justice, calls for mental health reform, or mobilization for systemic policy changes.

5.1. Contribution

This study in Chapter 4 makes at least four major contributions, including the i) advancement of the SUFA approach, ii) application of semantic relations-based unsupervised framing analysis in mass shooting coverage, iii) integration of attribution theory in the domain of computational research, and iv) release of two annotated mass shooting datasets for shooter and victims.

i) Advancing SUFA. As a key contribution, Chapter 4 advances the *Semantic Relations-based Unsupervised Framing Analysis* (SUFA) by applying it to a large-scale dataset and enhancing its methodological rigor in framing analysis. Unlike prior applications limited to smaller datasets and requiring manual oversight in clustering framing components, this chapter thoroughly refines and advances Step 7 of SUFA—the automated clustering of framing components—by integrating advanced natural language processing (NLP) techniques, including dependency parsing, coreference resolution, and clustering models. In particular, this study leverages large language models (LLMs), such as OpenAI's GPT-4o, to enhance the contextual understanding of framing components. This marks a key methodological innovation, as previous computational framing analyses primarily relied on frequency-based and co-occurrence methods that failed to capture semantic relationships between words. By incorporating semantic relations and contextual clustering, SUFA demonstrates a more nuanced and scalable approach to identifying media frames, making it a valuable tool for computational social science research. The effectiveness and advancement of SUFA are evidenced through its application in the large dataset and the exploration of insights frames. This advancement not only improves the frame detection goal in large datasets but also broadens SUFA's applicability to various events.

ii) Computational framing analysis of mass shooting coverage. Another key contribution is that this research is one of the limited studies to apply large-scale computational framing analysis to news coverage of mass shootings, specifically focusing on both shooter and victim frames through semantic relations. While previous research has explored media bias and framing in mass shootings, most studies relied on qualitative content analysis or supervised machine learning models, which require pre-labeled data and extensive manual effort. By using unsupervised computational techniques, this study provides a data-driven, scalable approach that identifies entity-centric emphasis frames across different media bias groups. This contribution is particularly valuable in exploring gun violence frames in large datasets, better understanding this public health pandemic, and finding potential interventions to help reduce gun violence.

iii) Integrating attribution theory with computational methods. This research bridges computational framing analysis with attribution theory by empirically demonstrating how media frames shape internal and external attributions of responsibility. While attribution theory has been widely applied in media studies, most research in this domain has relied on traditional qualitative and quantitative methods, with limited application in computational research. By integrating computational techniques, this study systematically analyzes media frames in a large dataset and leverages attribution theory to provide deeper theoretical insights. Specifically, the distinction between internal and external attributions offers a more nuanced interpretation of computationally retrieved framing components, enhancing our understanding of how different frames influence public perception. This approach not only expands the theoretical application of attribution theory in computational research but also highlights the psychological implications of media framing in shaping public reactions to mass shootings.

iv) Annotated mass shooting dataset. As part of this project in Chapter 4, two subsets of mass shooting datasets for the shooter and victims were labeled by both OpenAI's GPT and humans, followed by the evaluation of GPT's performance against human judgment. With the process, separate labeled datasets are now available for future research to use. The datasets can be used as ground truth to computationally analyze other mass shooting incidents, particularly using the SUFA approach.

v) Framing Analysis Advancement Across Disciplines. This dissertation contributes a timely and interdisciplinary advancement to the study of media framing in the age of rapidly evolving artificial intelligence, particularly regarding the emergence of large language models (LLMs). The proposed SUFA approach bridges the gap between the conceptual rigor of the social sciences and the computational capabilities of AI. For disciplines such as information science and computer science, SUFA demonstrates how AI, including dependency parsing and LLM-assisted clustering, can be harnessed to uncover deeper semantic structures in text beyond traditional topic modeling. For social science disciplines, including communication and journalism studies, the approach offers a scalable method to explore frames centered around particular entities, enabling researchers to engage with real-time framing trends. Among others, political science can also benefit from SUFA by exploring frames in public discourse and their influence on public opinion and policy debates. Overall, this work shows how AI, when paired with theory-driven frameworks, can enhance both the interpretability and scalability of framing analysis, enabling scholars and practitioners across domains to extract actionable framing insights from vast, complex text data.

5.2. Implications

This research has implications for both researchers and practitioners, as elaborated below.

First, scholars across academic fields—including media studies, political communication, public health, psychology, computation, and social science—can benefit from the SUFA approach and the annotated datasets. The SUFA can help them analyze frames in large-scale text data without relying on manual coding of textual data or predefined categories. This advancement broadens its application beyond traditional framing research based on qualitative and quantitative methods, as well as existing unsupervised computational framing analyses based on bag-of-words, co-occurrence, and frequency, ultimately supporting studies across disciplines and scholarship.

Second, by extracting semantic relations and context-driven frames through computational methods, SUFA can facilitate real-time media monitoring and longitudinal studies on how frames attribute certain entities, including a political leader, a company official, a brand, a country, an issue (e.g., tariff), and a group of entities (e.g., victims). This could be particularly useful for tracking various public discourses.

Third, SUFA has a specific potential to enhance existing social media analytics platforms, such as Brandwatch or Cision, by adding a deeper, context-aware, theory-driven, and framing-focused layer of insight. While current platforms primarily track mentions, sentiment, and keyword trends, SUFA could enable the extraction of entity-centric emphasis frames through semantic relations and contextual analysis. This would allow analysts to understand not only what is being said but also how key individuals, brands, or issues are being framed—whether as responsible agents, victims, threats, or sympathetic figures. By integrating SUFA, social media analytics platforms could move beyond surface-level metrics to deliver more nuanced framing analysis, thereby offering richer insights into public narratives, framing bias, and discourse dynamics in real time.

Fourth, crisis managers and strategic communication professionals can leverage SUFA to support real-time monitoring of media narratives and public sentiment, specifically regarding how their clients (e.g., organizations, individuals, or brands) are being framed. By offering framing-aware insights, SUFA enables a deeper understanding of public reactions during a crisis and of how responsibility or sympathy is attributed. This, in turn, can inform the development of more effective crisis response strategies based on how entities are framed, moving beyond surface-level metrics like sentiment and mention frequency.

Fifth, SUFA provides an automated method for data-driven journalists and media analysts to uncover framing trends across media outlets. This tool can assist computational journalists in analyzing and visualizing framing patterns, comparing media bias, and enhancing data-driven storytelling from text documents more effectively.

Sixth, the general public can also benefit from SUFA by using the approach to enhance media literacy and critically assess how news outlets frame individuals, groups, or events. While SUFA is not yet available to the public as an open-source platform, there are plans to develop one. In the meantime, people can still apply the approach to analyze how key entities, such as political figures, organizations, or brands, are framed in news coverage. This empowers media consumers, educators, and activists to uncover framing patterns, compare narratives across outlets, and potentially hold media organizations accountable for biased or misleading representations.

5.3. Limitations

Despite its contributions, this study has several limitations. The dataset consists of one month of news media coverage from nine outlets. While it is comprehensive, this may still not fully capture long-term shifts in framing. The analysis focused on English-language sources, limiting insights from other national media covering the shooting in different languages (e.g., Spanish).

From a methodological perspective, as mentioned above, this study aims to capture only entity-centric emphasis frames. Additionally, only the modifying words of an entity have been conceptualized and operationalized as framing components for the purpose of this study. However, there are many other forms of framing components, such as metaphor, placement of news (e.g., lead news), font size and color, and visuals (e.g., images, videos, and graphics) that news media use to construct frames. Therefore, the frames explored through SUFA are not fully comprehensive, leaving room for further advancement of this approach.

The human-labeled clustering used to evaluate GPT-led clusters of framing components may have introduced subjective bias, as such labeling reflects the interpretive judgments of coders. Also, the “older age” prompt instruction in victim framing seems not to have captured the age-related attributions of all victims, as the victims include both adult (teacher) and child (student) victims. Future refinements of SUFA and prompts could address these limitations.

Also, this study stems from the broader debate surrounding gun violence in the U.S. The media outlets analyzed reflect differing ideological orientations and audience expectations, which likely influence their framing practices. As such, the variation in shooter and victim frames across media groups may reflect routine journalistic norms and editorial positioning. These differences might also reflect the context of a polarized media environment.

Finally, while LLMs like GPT-4o provide more meaningful and coherent clusters compared to traditional methods, their initial use without a list of cluster names results in unhelpful groupings, underscoring the importance of prompt engineering and structured human guidance. Additionally, because LLMs operate as black boxes, their decision-making processes are not fully transparent. This poses challenges for reproducibility and interpretability. The evolving nature of these models also means that future outputs may differ even when using identical prompts, complicating long-term validation in computational social science.

5.4. Future Research Directions

This study opens several promising avenues for future research.

Gun violence framings. Longitudinal studies could expand the current work by examining how media framing patterns surrounding mass shootings evolve over time. Tracking changes across weeks, months, or years would provide insight into the persistence or transformation of particular frames and their potential influence on public perception, emotional response, and government policy. A comparative analysis of multiple mass shooting events could assess whether framing strategies differ by context, geographic location, racial identity of involved parties, or political climate.

Methodology. Further refinement of SUFA is necessary. Integrating fine-tuned large language models (LLMs) could enhance clustering accuracy by better capturing abstract or overlapping frames, thus reducing misclassifications. This could improve the contextual awareness of framing components. Additionally, future research can expand SUFA's capabilities by encompassing a wider range of framing components, including metaphors, structural cues (e.g., headline prominence, article placement), narrative sequences, and visuals. This would enable scholars to move beyond lexical modifiers and capture a broader spectrum of frames.

AI integration. Beyond GPT-based clustering, future research can explore the potential of end-to-end, AI-driven systems that might incorporate LLMs with coreference resolution and discourse parsing capabilities. These systems could automate the detection and classification of frames across texts and images, enabling richer framing insights. AI models fine-tuned for media discourse could also enhance the semantic granularity of SUFA. Human-AI collaboration, where researchers guide LLMs through prompt engineering or co-clustering, could lead to more adaptive and iterative approaches to framing analysis.

Interdisciplinary work. Interdisciplinary collaboration will also be critical to advancing SUFA. Partnerships between computational and social science researchers can help refine the model beyond entity-centric frames, enabling exploration of problem-definition, causal, and solution-oriented frames.

Open-source SUFA platform. The development of an open-source SUFA platform equipped with automated NLP pipelines and an interface would make this novel approach more accessible to a broader range of scholars, professionals, journalists, and the general public. Journalists, researchers, educators, and crisis communication professionals could utilize such a tool to analyze real-time framing trends, monitor narrative shifts, and design more effective communication strategies during crises. SUFA can also facilitate evidence-based decision-making across sectors.

E. CHAPTER 5: CONCLUSION

This dissertation presents a novel computational approach to framing analysis by introducing and applying *Semantic Relations-based Unsupervised Framing Analysis (SUFA)*. Through three interconnected research studies, this work advances the field of computational framing by moving beyond traditional bag-of-words and frequency-based methods toward a semantic relations-based approach. While developing and advancing the SUFA, this study systematically examines how mass shooters and victims are framed in U.S. news media coverage. This unveils significant patterns in media bias and attribution of responsibility for gun violence. This study also advances the theoretical understanding of media bias and framing effects by integrating computational methods with attribution theory.

This dissertation is structured into three major studies, each contributing to the introduction and advancement of the computational framing analysis approach, SUFA. The first study in Chapter 2 critically reviews existing computational framing methods. Notably, this project identifies limitations in traditional bag-of-words and frequency-based approaches and emphasizes the need for a semantic relations-based method.

Building on the insights from the Chapter 2 study, study 2 in Chapter 3 introduces SUFA as a novel computational approach. SUFA demonstrates feasibility in a mixed-methods study that integrates qualitative textual analysis and computational techniques to extract entity-centric emphasis frames from news media. Finally, Chapter 4 applies SUFA to a large-scale dataset on U.S. gun violence. This study analyzes how different news media outlets frame shooters and victims. This study leverages dependency parsing, coreference resolution, and LLM-assisted clustering, pushing the boundaries of unsupervised computational framing analysis beyond existing bag-of-words and frequency-based framing analysis approaches.

The results reveal significant differences in how media outlets across ideological biases frame mass shooters and their victims. Right-leaning media tend to emphasize action attribution, threat, and allegation certainty, which reinforces internal attributions that blame the perpetrator as an active agent of violence. In contrast, left-leaning media employ allegation doubt, mental health, and race/ethnicity frames, shifting toward external attributions that associate the shooting with broader systemic factors such as mental illness, racial bias, or societal failures.

Across all three media groups analyzed in Chapter 4, victims are frequently framed through dehumanization, reducing them to statistics. However, left-leaning media use more personalized and emotional frames (e.g., “our victims,” “your victims,” etc.), whereas right-leaning ones portray them through a sacrificial or service-oriented lens. Additionally, the framing components were found to significantly predict news media bias. These results support the argument that news media bias might shape public perception through different framing components.

In terms of contributions, this dissertation’s major contributions include:

i) The dissertation first provides a comprehensive understanding and resources on existing computational framing analysis methods. It also consolidates these resources for interested scholars and professionals to gain deeper knowledge and begin building upon that foundation. While presenting critical reviews, the dissertation also contributes new insights to the ongoing discussion about advancing the computational methods of framing analysis based on semantic relations.

ii) The dissertations developed SUFA to explore entity-centric emphasis frames. Unlike traditional topic modeling or co-occurrence-based methods, SUFA leverages semantic relations to capture contextually meaningful frames, moving beyond word frequency and bag-of-words limitations.

iii) This research integrates NLP and LLM techniques in framing analysis by applying dependency parsing, coreference resolution, and GPT-4o-assisted clustering to improve frame detection accuracy in large-scale datasets.

iv) The project also advances computational methods for media bias detection through their frames, especially in the context of gun violence in the United States. Specifically, the study demonstrates how OLS regression can predict media bias based on framing components and offers a quantifiable approach to measuring news outlets' ideological bias in their coverage.

v) In Chapter 4, GPT and human annotators labeled two subsets of mass shooting data (shooter and victim). GPT's performance was evaluated against human judgment, producing separate labeled datasets that can be used as ground truths in future research, particularly for computational analysis using the SUFA approach.

Overall, this dissertation marks a significant step forward in computational framing analysis by demonstrating how semantic relations-based methods can enhance media research. By integrating NLP, LLMs, and statistical modeling, this dissertation provides a scalable and replicable approach to understanding media bias, crisis communication, and public perception through a novel computational framing analysis approach, namely *Semantic Relations-based Unsupervised Framing Analysis (SUFA).*

Building on the advancements presented in this dissertation, future research can aim to further develop SUFA into a robust, open-source platform for real-time computational framing analysis. This next phase can focus on (1) enhancing SUFA's ability to detect frames across modalities, including metaphor, visual, and structural framing components, by integrating multimodal large language models (LLMs), (2) expanding its application to a wider range of public health, political, and crisis communication domains, and (3) exploring cross-national comparisons of media framing using multilingual NLP models.

APPENDICES

Appendices for Chapter 2 (Project 1)

Appendix 2.1. Framing Components

Table 2. 1. Framing Components Used to Construct Frame(s)

	Devices	Sources	Devices	Sources
Content (Texts)			**Content (Visual)**	
1.	Words	Entman, 1993	Metaphors	Fairhurst & Sarr, 1996; Gamson & Modigliani, 1989; Tankard, 2001
2.	Stock phrases	Entman, 1993	Visual images (e.g., picture, icons)	Tankard, 2001; Gamson & Modigliani, 1989
3.	Stereotyped images	Entman, 1993	Charts, and graphs	Tankard, 2001
4.	Sources of information	Entman, 1993		
5.	Sentences	Entman, 1993	**Action**	
6.	Metaphors	Fairhurst & Sarr, 1996; Gamson & Modigliani, 1989	Placement (e.g., front page, editorial page)	Entman, 1993; Swenson, 1990
7.	Jargon/catchphrases	Fairhurst & Sarr, 1996; Gamson & Modigliani, 1989	Repetition	Entman, 1993
8.	Contrast	Fairhurst & Sarr, 1996	Associating them with culturally familiar symbols	Entman, 1993

9.	Spin	Fairhurst & Sarr, 1996
10.	Stories	Fairhurst & Sarr, 1996
11.	Headlines and kickers (small headlines over the main headlines).	Tankard, 2001
12.	Subheads	Tankard, 2001
13.	Photo captions	Tankard, 2001
14.	Leads (beginnings of news stories)	Tankard, 2001
15.	Selection of sources or affiliations	Tankard, 2001
16.	Selection of quotes	Tankard, 2001
17.	Pull quotes (quotes that are blown up in size for emphasis)	Tankard, 2001
18.	Logos (graphic identification of the particular series an article belongs to)	Tankard, 2001
19.	Statistics	Tankard, 2001
20.	Concluding statements or paragraphs of articles	Tankard, 2001
21.	Exemplars (i.e., historical examples from which lessons are drawn)	Gamson & Modigliani, 1989
22.	Depictions	Gamson & Modigliani, 1989
23.	Emotion	Aaroe, 2011
24.	Hashtag	Borah, 2018

Include	Entman, 1993
Omit or hide	Entman, 1993
Show root causes of a problem	Gamson & Modigliani, 1989
Show consequences or effect	Gamson & Modigliani, 1989
Make appeals to principles (moral claims)	Gamson & Modigliani, 1989
Context	
Contextual information	Baden, 2018
Culture	Entman, 1993
Communicator	
Thought	Fairhurst & Sarr, 1996
Forethought	Fairhurst & Sarr, 1996
Being bias (not a frame, but a tool of constructing a frame)	Fairhurst & Sarr, 1996

Appendix 2.2. Summary of Method and Resources

Table 2. 2. Summary of Computational Framing Analysis Methods and Resources

	Citation	**Type**	**Domain**	**Method/ corpora**	**Resource**
1	Boydstun et al. (2013)	Corpus, method	Tobacco, immigration, same-sex marriage	Regression, Policy frames codebook (PFC)	N/A
2	DiMaggio et al. (2013)	Application	Artists & arts	Topic Modeling	N/A
3	Boydstun et al. (2014)	[1]	[1]	[1]	N/A
4	Card et al (2015)	Method	[1]	Media frames corpus (MFC)	GitHub
5	Nguyen (2015)	Method	Congressional debates, reviews	Hierarchical topic modeling	GitHub
6	Nguyen et al. (2015)	Method	Congress speech	[5]	[5]
7	Burscher et al. (2016)	Application	Nuclear power	Cluster analysis	N/A
8	Ji & Smith (2017)	Application	Immigration	Neural network, semantic relations	GitHub
9	Johnson et al., 2017	Application	Abortion, affordable care act	[8]	GitHub
10	Naderi and Hirst (2017)	Application	Immigration, smoking	[8]	N/A
11	Field et al. (2018)	Application	U.S. coverage in Russian a newspaper	[4]	N/A
12	Sturdza, M. D. (2018).	Method	N/A	Operationalization of semantic relations	N/A
13	Khanehzar et al. (2019)	Application	Immigration, same-sex marriage	[8]	N/A
14	Liu et al. (2019)	Method, annotated corpus	Gun violence	Gun violence frame corpus (GVFC)	GitHub
15	Walter & Ophir (2019)	Method	Senate coverage, epidemics	Topic modeling, network analysis	GitHub

16	Akyürek et al. (2020)	Application & extension	[14]	[14]	GitHub 1, GitHub 2
17	Cabot et al. (2020)	Application	Immigration, smoking	[8]	GitHub
18	Kwak et al. (2020)	Application	Fake news	[4]	GitHub
19	Sanderink (2020)	Application	Renewable energy	Frequency and co-occurrence model	Programs
20	Yang and Kang (2020)	Application	Telecom	[19]	N/A
21	Bednarek and Carr (2021)	Method	Lifestyle	[19]	Wordsmith
22	Bhatia et al. (2021)	Application & extension of open-source tool	Gun violence	[14]	[14]
23	Gilardi et al. (2021)	Application	Gov policy	Structured topic modeling	Appendix
24	Jing and Ahn (2021)	Application	Political polarization	FrameAxis	N/A
25	Kwak et al. (2021)	Method	Reviews	[24]	N/A
26	Li et al. (2021)	Application	#MeToo movement	[2]	N/A
27	Mendelsohn et al. (2021)	Application	Immigration	[8]	GitHub
28	Nicholls and Culpepper (2021)	Comparative	Banking	N/A	N/A
29	Ophir et al. (2021)	Application	COVID-19	[15]	N/A
30	Supran and Oreskes (2021)	Application	Gun violence, oil and gas	[2]	N/A
31	Tourni et al. (2021)	Application &extension	Gun violence	[14]	[14]
32	Walter and Ophir (2021)	Application	[15]	[15]	[15]
33	Ylä-Anttila et al. (2021)	Application	Climate change	[2]	N/A
34	Ziems and Yang (2021)	Method	Police violence	Semantic relations	GitHub
35	Guo et al. (2022)	[22]	Gun violence	[14]	[14]
36	Yu (2022)	Method	Refugee crisis	[34]	GitHub
37	Kang and Yang (2021)	Application	Racism, Xenophobia	[19]	[19]

Appendices for Chapter 3 (Project 2)

Appendix 3.1. Semantic Framing Components in News Reports from NYT and CNN

Table 3. 1. Words, phrases, and their relations in news reports of NYT and CNN. Highlighted words and phrases are common in both Tables 3.1 and 3.2.

Salvador Ramos		**Gun**	**Verb**	**Verb Modifier**	**Victim**	**Event modifier**
Modifiers	**History**					
18-year-old		A long rifle	Burst in and killed	After … [another event]	18	30th K-12 shooting
18-year-old man		Assault rifle	Came in an opened fire	Fatally	19	6-year-old son
Armored		Semiautomatic rifle	in [in custody]	Horrifically	Adults	Aftermath
Gunman		Semiautomatic weapons	Kills	Incomprehensibly	Age between 6 and 7 years old	Attack
He		Semitauonic firearms	Left		At least	Deadliest mass shooting
Shooter		Tactical vest	Left dead		Children	Deadly shooting
Suspect		With a rifle	Left killing		Elementary school children	Devastating
			Opened fire		Elementary school students	Elementary school shooting
			Shoots		Kids	Horrific mass murder
			Shot and killed		One	Mass school shooting
			Shot dead		School children	Mass shooting

Stormed into

Students	Massacre
Teachers	Nation reeling
Two	School massacre
Victims	School shooting
	Second deadliest
	Shakes a nation
	Slaughter
	Slayings
	Stealing their lives
	Terrorism
	Terrorist attack
	Tragedy
	Tragic
	Violent
	Worst school shooting

Appendix 3.2. Semantic Framing Components in News Reports from the Wall Street Journal and Fox News

Table 3. 2. Words, phrases, and their relations in news reports of *Wall Street Journal* and *Fox News.*

Salvador Ramos		**Gun**	**Verb**	**Verb modifier**	**Victim**	**Event modifier**
Modifiers	**History**					
18-year-old	Broken family	AR-platform rifle	Accused of shooting	Fatally	14	Aftermath
18-year-old man	Hostile	Handgun	Allegedly committed by		19	Attack
A resident of Uvalde	Unsettled classmates	Legally purchased	Broke into the school		Adults	Deadliest
Active shooter	Violent	Two rifles	Claimed the lives …		At least	Deadliest shooting
Alleged gunman	was 'flashing red'		Entering [the school]		Children	Deadly
Alleged gunman			Shot		Children	Deadly [shooting]
Alleged shooter			Gunned down		One	Devastated the town
Alone			Is accused of		Students	elementary school shooting
Former student at Uvalde High school			Kills		Teacher(s)	Horrific shooting
Gunman			Left		Two	Horrific tragedy
He			Left dead		Victims	Later discovered to be the shooting
Mass shooter			Left killing		Xavier Lopez	Local elementary-school shooting
Now-deceased			Opened fire			Mass casualty incident

now-deceased suspect	Walking into [school]	Mass murder
Shooter		Mass shooting
Suspect		Massacre
Suspected lone gunman		murders
Suspected shooter		School shooting
Suspected Uvalde school shooter		Senseless crime
Texas school shooting suspect		Shocked the country
		Shooting
		Texas elementary school shooting
		Texas mass shooting
		Texas school shooting
		third most deadly
		Tragedy

Appendix 3.3. Keywords and Dependency Relations

Table 3. 3. Keywords and dependency relations used for the shooter, victims, and the event.

	Keywords	**Relations**
Shooter	‘gunman’, ‘gunmen’, ‘man’, ‘Salvador’, ‘Ramos’, ‘shooter’, ‘shooters’, and ‘suspect’.	'acl', "amod", 'appos', "compound", "relc", 'nsubj', 'dobj', and 'nsubjpass'..
Victims	‘adult’, ‘adults’, ‘child’, ‘children’, ‘kids’, ‘schoolchildren’, ‘student’, ‘students’, ‘teacher’, ‘teachers’, ‘victim’, and ‘victims’.	'acl', 'compound', 'nummod', 'relcl', 'amod', 'dobj', 'nsubj', 'nsubjpass', and 'poss'
Event	‘shooting’, ‘shootings’, ‘attack’, ‘massacre’, ‘event’, ‘tragedy’, ‘terrorism’, ‘slaughter’, ‘crime’, ‘slayings’, ‘murder’, and ‘aftermath’.	'amod', 'advmod', 'compound', 'nummod', and 'relcl'

Appendix 3.4. Shooter Framing Components Deployed by News Media Outlets

Table 3. 4. Framing components (with frequencies) deployed by each news media outlet to attribute the shooter, grouped under different associated semantic relations.

	acl clad (2) identified (2) **amod** active (3) old (21) deranged (1) many (1) other (1) suspected (1) alleged (1) grandmother (1) **appos** Ramos (2) **compound** Salvador (7) mass (1)	**acl** approaching (2) barricaded (2) driven (1) **amod** shooting (1) angry (1) armed (2) old (5) **compound** shooting (2) Salvador (2)		**acl** named (1) **amod** grandmother (1) old (8) teenage (1) unhappy (1) young (1) deceased (2) civilized (2) active (4) **appos** himself (1) student (1) resident (1) old (1) man (1) birthday (1) Ramos (2) **compound** school (1) mass (4) Salvador (12)	**acl** accused (1) identified (1) **amod** active (1) alleged (2) bureaudefined (1) deceased (4) lone (2) old (1) suspected (4) upstate (1) red (1) **appos** resident (2) ones (1) gunman (1) Romas (1) Ramos (2) 18 (1) **compound** suspect (1) York (1) resident (1) mass (1) Texas (1) Salvador (14) Ramos (1) school (1)	
Extreme left	**Left**	**Left-center**	**Least-biased**	**Right-center**	**Right**	**Extreme right**
	CNN	**NYT**		**WSJ**	**FOX**	

Appendix 3.5. Victim Framing Components Deployed by News Media Outlets

Table 3. 5. Framing components (with frequencies) used by each news media outlet to attribute victims, grouped under different associated semantic relations.

	acl aged (1) **amod** local (1) young (2) **compound** Parents (1) parents (1) school (1) **nummod** 13 (3) 14 (4) 18 (1) 19 (13) 20 (2) 26 (2) 535 (1) Eighteen (1) Nineteen (5) Two (2) one (3) two (17) **relcl** treated (1)	**acl** killed (1) **amod** dead (1) other (2) several (2) young (3) **compound** Hook (1) Uvalde (1) daughter (1) grade (1) parents (2) roll (1) school (7) **nummod** 14 (2) 18 (1) 19 (14) 20 (3) one (1) two (12) **poss** America (1) Her (1) my (2) our (2) your (2) **relcl** killed (1)		**acl** celebrating (1) killed (1) **amod** former (1) other (1) small (1) **compound** Elementary (1) Robb (1) Trump (1) adult (1) mother (1) **nummod** 16 (1) 17 (1) 19 (15) 20 (1) 21 (3) four (1) two (13) **poss** her (2)	**amod** dead (1) innocent (1) little (1) ofentry (1) old (1) **compound** School (1) asa (1) center (1) school (1) **nummod** 14 (2) 18 (3) 19 (8) 4,000 (1) Two (1) eight (1) one (3) two (7) **poss** our (3) **relcl** missing (1)	
Extreme left	**Left**	**Left-center**	**Least-biased**	**Right-center**	**Right**	**Extreme right**
	CNN	**NYT**		**WSJ**	**FOX**	

Appendix 3.6. Event Framing Components Deployed by News Media Outlets

Table 3. 6. Framing components (with frequencies) used by each news media outlet to attribute the <u>event</u>, grouped under different associated semantic relations.

	advmod ago (1) fatally (1) least (2) **amod** 30th (2) American (2) Deadly (3) deadliest (2) deadly (2) heinous (1) horrific (1) previous (1) second (2) tragic (1) **compound** Hook (2) mass (5) school (10) **nummod** 39 (2) three (1) **relcl** happened (1) left (4)	**advmod** ago (3) far (1) **amod** Latest (1) deadliest (6) deadly (1) horrifying (1) immediate (1) mass (7) next (1) previous (1) recent (1) reported (1) such (1) **compound** Buffalo (1) Newtown (1) School (2) mass (9) school (14) **nummod** 2012 (1) 215 (1) 693 (1) two (1) **relcl** killed (3) say (1) took (1)		**advmod** away (2) **amod** awful (4) deadliest (1) horrific (1) latest (2) local (1) mass (2) new (2) next (1) **compound** Mass (1) mass (20) school (8) **nummod** 2011 (1) claimed (1) died (1) have (1) is (1) rises (1) targeted (1) tolerated (1)	**amod** deadliest (2) deadly (4) fourth (1) horrific (2) last (1) major (1) mass (1) recent (1) senseless (2) **compound** Parkland (1) Texas (1) Tuesday'smass (1) mass (16) preventmass (1) school (9) **nummod** 20 (1) 2018 (1) **relcl** devastated (1) had (1) happened (1) left (2)	
Extreme left	**Left**	**Left-center**	**Least-biased**	**Right-center**	**Right**	**Extreme right**
	CNN	**NYT**		**WSJ**	**FOX**	

Appendices for Chapter 4 (Project 3)

Appendix 4.1. GPT Prompt to Cluster Shooter Framing Components

Prompt for GPT to Cluster Shooter-Framing Components:

“““ The word "{token}" modifies the shooter or is associated with the shooter in a mass shooting incident. Please categorize this word into appropriate cluster(s) by using the following prompts and the word's context in this sentence: "{sentence}". The word appears in the sentence as capitalized and highlighted with double asterisks (**LIKE THIS**). A word can belong to multiple clusters if applicable.

Prompts:

1. Younger age: Words in this cluster modify the shooter in terms of his/her younger age.

2. Older age: Words in this cluster modify the shooter in terms of his/her older age.

3. Race/Ethnicity: Words in this cluster identify or associate with the shooter’s race, ethnicity, country of origin, or other racial/ethnic groups.

4. Mental health: Words in this cluster identify or associate with the shooter’s mental stability, mental health issues, loneliness, anger, or social relationships.

5. Religion and culture: Words in this cluster identify or associate with the shooter’s religious or cultural identities or affiliations.

6. Humanization: Words in this cluster humanize the shooter or portray him/her in a sympathetic or relatable manner.

7. Threat: Words in this cluster identify or associate with potential future threats from the shooter.

8. Terrorism/Terror: Words in this cluster associate or attribute the shooter to terrorism, terror connections, or related identities and affiliations.

9. Immigration: Words in this cluster identify or associate with the shooter's immigration status, or such identities, or affiliations.

10. Family background: Words in this cluster identify or associate with the shooter's family background or family circumstances.

11. Stereotypes: Words in this cluster identify or associate with other stereotypes about the shooter, not covered by other categories.

12. Bullying: Words in this cluster seem to be bullying or demeaning the shooter.

13. Allegation certainty: Words in this cluster suggest a higher degree of certainty regarding the shooter's action or confirm the shooter's action.

14. Allegation doubt: Words in this cluster suggest doubt or uncertainty regarding the shooter's action or do not fully confirm the shooter's action.

15. Action attribution: Words in this cluster at tribute specific actions or a certain severity of action to the shooter, such as killing, murder, or massacre.

16. Other: Words in this cluster were not categorized in any of the above clusters.

So, the 16 clusters' names are: 'Younger age', 'Older age', 'Race/ Ethnicity', 'Mental health', 'Religion and culture', 'Humanization', 'Threat', 'Terrorism/ Terror', 'Immigration', 'Family background', 'Stereotypes', 'Bullying', 'Allegation certainty', 'Allegation doubt', 'Action attribution', and 'Other'.

Your final output is just the appropriate cluster name(s). I do not want any clarification. If you find a single word categorized into multiple clusters, please provide all cluster names separated by a comma.’”””

Appendix 4.2. GPT Prompt to Cluster Victim Framing Components

Prompt for GPT to Cluster Victim-Framing Components:

"""" The word "{token}" modifies the victim(s) or is associated with the victim(s) in a mass shooting incident. Please categorize this word into appropriate cluster(s) by using the following prompts and the word's context in this sentence: "{sentence}". The word appears in the sentence as capitalized and highlighted with double asterisks (**LIKE THIS**). A word can belong to multiple clusters if applicable.

Prompts:

1) Younger Age – Words that modify the victim(s) by emphasizing their comparative younger age. (Example: school, children, young.)

2) Older Age – Words that modify the victim(s) by emphasizing their comparative older age. (Example: 18-year-old, old)

3) Our [Victims] – Words that identify the victim(s) using first-person pronouns. (Example: our kids, our loved ones.)

4) Your [Victims] – Words that identify the victim(s) using second-person or third-person pronouns. (Example: your kids, their kids.)

5) Heroism – Words that modify the victim(s) by portraying them as brave, selfless, or protective in the face of the shooting. (Example: brave, heroic, defender, protector, sacrificed.)

6) Innocence and Vulnerability – Words that modify the victim(s) by identifying them as innocent, defenseless, or unarmed. (Example: innocent, young, and helpless)

7) Tragic Loss – Words that modify the victim(s) by highlighting the unjust nature of the victims’ deaths. (Example: tragic, senseless, irreplaceable)

8) Humanization – Words that modify the victim(s) by referencing their personal positive attributes and roles. (Example: beloved, kind, talented, and father)

9) Dehumanization – Words that modify the victim(s) by identifying them with impersonal statistics or data points (e.g., casualties, bodies, toll) or their traits that contribute to the shooting (Example: reckless, controversial, troublemaker.)

10) Bullying – Words that modify the victim(s) as individuals who had bullied, teased, or rejected the shooter. (Example: bully, rejected, cruel, mean.)

11) Other – Words that do not fit into any of the above clusters.

So, the 11 clusters' names are: Younger Age, Older Age, Our [Victims], Your [Victims], Heroism, Innocence and Vulnerability, Tragic Loss, Humanization, Dehumanization, Bullying, and Other.

Your final output is just the appropriate cluster name(s), with no additional explanations. If you find a single word categorized into multiple clusters, please provide all cluster names separated by comma (e.g., "Cluster 1, Cluster 2"). """

Appendix 4.3. Shooter Framing Components Clustered by GPT into Frames

Table 4. 8. *GPT4o-Clustered Framing Components in Different Semantic Relations with* *<u>Shooter</u>*

Clusters	**Words (frequency)**
Action Attribution (26 words, 922 times)	active(207), shooter(145), gunman(126), mass(111), murderer(71), school(70), armed(52), shooting(42), barricaded(24), dead(13), accused(13), killed(9), arrested(7), carrying(6), killer(6), motivated(4), called(3), deceased(2), suspect(2), multiple(2), bad(2), hospital(1), a.(1), que(1), lone(1), other(1)
Younger Age (10 words, 794 times)	old(485), school(148), 18(68), young(43), teen(17), teenage(14), shooter(10), man(4), faced(4), most(1)
Threat (26 words, 683 times)	active(233), gunman(93), mass(63), armed(52), school(46), potential(28), shooter(25), barricaded(24), crazed(24), madman(20), murderer(15), deranged(12), future(7), shooting(7), bad(6), carrying(4), irate(4), killer(4), other(3), young(3), angry(3), lone(2), a.(2), multiple(1), called(1), suspected(1)
Allegation Certainty (16 words, 363 times)	shooter(118), murderer(71), gunman(44), active(42), identified(20), shooting(15), accused(13), killed(9), arrested(8),

	dead(7), suspect(7), killer(3), school(3), motivated(1), a.(1), deceased(1)
Allegation Doubt (6 words, 185 times)	alleged(106), suspected(61), suspect(8), accused(7), potential(2), active(1)
Mental Health (13 words, 110 times)	deranged(33), crazed(27), madman(20), shooter(7), angry(6), irate(6), school(2), most(2), lone(2), potential(2), gunman(1), other(1), bad(1)
Older Age (3 words, 43 times)	old(35), granddad(7), most(1)
Race/Ethnicity (6 words, 42 times)	white(21), black(10), texas(6), uvalde(3), other(1), florida(1)
Family Background (7 words, 16 times)	granddad(7), shooter(2), school(2), texas(2), most(1), hardworking(1), suspect(1)
Humanization (5 words, 15 times)	hardworking(6), man(5), faced(2), granddad(1), called(1)
Stereotypes (4 words, 11 times)	school(8), motivated(1), bad(1), madman(1)
Terrorism/Terror (3 words, 10 times)	motivated(4), shooting(3), mass(3)

Bullying (4 words, 7 times)	school(3), madman(2), most(1), bad(1)
Religion and culture (1 words, 2 times)	motivated(2)

Appendix 4.4. Victim Framing Components Clustered by GPT into Frames

Table 4. 9. *GPT4o-Clustered Framing Components in Different Semantic Relations with Victims*

Clusters	**Words (frequency)**
Dehumanization (46 words, 2310 times)	19(1251), two(258), dead(158), 20(84), 21(62), nineteen(47), 14(44), six(41), four(30), killed(30), 18(26), 2(26), 17(21), eight(20), nine(19), volatile(16), three(15), 600(14), dozen(11), slaughtered(11), 12(11), 311,000(10), 26(10), 30(8), 1,000(8), 16(7), many(7), 11(6), 10(6), five(6), 15(6), 250(6), multiple(5), 13(4), one(4), deceased(4), more(3), 4(3), slain(3), people(2), die(2), shot(1), most(1), died(1), other(1), shooting(1)
Younger Age (88 words, 1480 times)	school(337), grade(175), young(128), old(96), 19(73), our(63), elementary(62), little(47), four(41), two(31), college(22), younger(18), hook(18), small(18), my(18), other(16), daughter(14), one(14), shooting(12), 18(12), robb(11), children(11), roll(10), fellow(10), students(9), academy(9), people(8), age(8), first(8), 20(8), daniel(7), youngest(7), parents(7), more(7), child(7), student(6), many(6), attending(6), classroom(5), six(5), three(5), parkland(5), your(5), sweet(4), 14(4), a.(4), their(4), 16(4), holding(4), mother(4), aged(4), 13(3), are(3), honor(3), were(3), shot(3), 11(3), 4(3), richardson(3), us(2), living(2), female(2), eight(2), most(2), nineteen(2),

	family(2), running(2), uvalde(2), 600(2), five(1), loved(1), whose(1), former(1), 17(1), transgender(1), 30(1), taught(1), surviving(1), third(1), own(1), teachers(1), families(1), lost(1), math(1), her(1), is(1), several(1), throwing(1)
Innocence and Vulnerability (94 words, 1287 times)	school(286), young(125), innocent(85), grade(62), old(61), elementary(49), little(47), 19(36), terrified(32), injured(21), wounded(19), our(19), small(18), shot(18), younger(18), murdered(15), four(15), shooting(13), my(12), desperate(12), killed(12), survived(11), slaughtered(10), gunned(10), one(10), trapped(10), daughter(9), two(8), undocumented(8), taken(8), called(8), child(7), slain(7), surviving(7), locked(7), first(7), more(7), children(7), people(7), traumatized(7), beautiful(6), youngest(6), age(6), sweet(6), 20(6), attending(6), 18(6), massacred(6), eyed(5), lived(5), violence(5), students(5), student(5), many(4), running(4), holding(4), hook(4), mother(4), other(4), classroom(3), living(3), massacre(3), a.(3), alive(3), nineteen(2), dead(2), families(2), transgender(2), parents(2), die(2), his(2), several(2), 16(2), streaming(2), were(2), 4(2), six(2), 11(2), are(2), whose(1), college(1), her(1), family(1), their(1), your(1), lost(1), aged(1), loved(1), eight(1), crime(1), three(1), 13(1), run(1), 14(1)

Tragic Loss (62 words, 1099 times)	killed(304), dead(199), shooting(75), innocent(59), murdered(50), slain(50), died(50), 19(21), more(20), shot(19), gunned(19), wounded(16), lost(16), hook(15), slaughtered(15), injured(12), desperate(11), beloved(11), massacre(10), two(10), massacred(10), tragic(7), one(6), survived(6), violence(6), deceased(6), first(5), die(5), four(5), surviving(5), many(4), families(4), people(3), parents(3), alive(3), lived(3), 18(3), several(2), multiple(2), 20(2), parkland(2), our(2), beautiful(2), are(2), taken(2), mireles(1), garcia(1), called(1), 14(1), identified(1), couple(1), other(1), 4(1), her(1), loved(1), were(1), confirmed(1), nineteen(1), crime(1), school(1), traumatized(1), my(1)
Your [Victims] (47 words, 768 times)	their(171), your(123), her(108), parents(87), other(45), his(39), fellow(22), own(20), one(17), four(14), whose(13), families(12), two(12), school(9), mother(8), daughter(7), child(4), were(4), couple(4), family(4), people(4), six(3), older(3), students(3), former(2), many(2), 16(2), garcia(2), female(2), 4(2), daniel(2), are(2), young(2), teacher(1), 19(1), three(1), my(1), children(1), more(1), lost(1), identified(1), 15(1), robb(1), america(1), grade(1), teachers(1), aged(1)

Our [Victims] (39 words, 561 times)	our(269), my(190), we(16), us(8), own(7), two(7), parents(5), daughter(5), children(4), one(4), america(4), fellow(4), other(3), american(3), school(2), more(2), innocent(2), were(2), little(2), small(2), uvalde(2), parkland(1), is(1), loved(1), age(1), most(1), are(1), child(1), 16(1), family(1), four(1), mother(1), six(1), her(1), grade(1), many(1), youngest(1), young(1), five(1)
Humanization (68 words, 504 times)	grade(105), four(33), school(21), parents(20), my(17), two(15), mother(12), roll(12), daughter(12), honor(12), teacher(12), beloved(11), loved(10), her(10), garcia(9), elementary(9), befriended(9), beautiful(8), outgoing(8), longtime(8), sweet(7), former(7), retired(6), substitute(6), veteran(6), renowned(6), hero(6), teachers(6), daniel(6), math(5), couple(5), eyed(5), fellow(5), taught(5), his(5), college(5), mireles(4), history(4), one(4), english(4), families(3), slain(3), family(3), classroom(3), robb(3), other(2), student(2), run(2), three(2), six(2), called(2), people(1), innocent(1), traumatized(1), survived(1), living(1), wounded(1), shooting(1), old(1), female(1), our(1), we(1), running(1), a.(1), small(1), male(1), was(1), uvalde(1)
Older Age (26 words, 365 times)	two(211), old(60), one(12), 2(11), older(10), longtime(8), retired(7), veteran(6), teacher(6), four(4), teachers(3), parents(3), 18(3), three(3), a.(2), former(2), six(2), aged(2), history(2), taught(2), college(1), 19(1), our(1), many(1), your(1), 4(1)

Heroism (34 words, 110 times):	called(18), leading(10), hero(10), her(10), survived(6), died(6), teacher(5), grade(4), garcia(4), young(3), school(3), armed(3), mother(2), their(2), one(2), parkland(2), veteran(2), lived(2), uvalde(1), daughter(1), mireles(1), other(1), fellow(1), 19(1), taken(1), classroom(1), are(1), his(1), surviving(1), teachers(1), befriended(1), parents(1), daniel(1), lost(1)
Bullying (2 words, 3 times)	other(2), one(1)

Appendix 4.5. Predicting Media Bias Scores Based on Framing Components and Frames

Table 4. 10. *Regression Analysis: Predicting Media Bias Scores—left (-1), least (0), and right (1)—Based on Shooter Framing Components.*

Effect	Estimate (B)	SE	t	p	95% CI LL	95% CI UL
Intercept	-0.0658	0.049	-1.334	0.183	-0.163	0.031
Old	0.7039	0.395	1.783	0.075	-0.071	1.479
Active	-1.0856	0.43	-2.527	0.012	-1.929	-0.242
School	0.53	0.372	1.427	0.154	-0.2	1.26
Shooter	2.8034	0.632	4.439	0	1.563	4.043
Texas	0.0252	0.438	0.058	0.954	-0.835	0.885
Uvalde	0.122	0.607	0.201	0.841	-1.069	1.313
Gunman	1.9581	0.448	4.366	0	1.078	2.839
Mass	-0.7078	0.665	-1.065	0.287	-2.013	0.597
Alleged	0.1693	0.575	0.294	0.769	-0.96	1.298
18	0.3168	0.215	1.472	0.142	-0.106	0.739
Suspected	-0.3456	0.724	-0.478	0.633	-1.766	1.075
Armed	0.0242	0.317	0.076	0.939	-0.598	0.646
Young	-0.7399	0.427	-1.734	0.083	-1.578	0.098
Shooting	-0.7889	0.385	-2.048	0.041	-1.545	-0.033
Buffalo	0.6588	0.485	1.359	0.175	-0.293	1.611

Deranged	1.0528	0.439	2.397	0.017	0.19	1.915
Potential	-0.4977	0.872	-0.571	0.568	-2.21	1.214
Crazed	0.0443	0.671	0.066	0.947	-1.273	1.362
Barricaded	0.1281	0.464	0.276	0.783	-0.784	1.04
Identified	-0.6246	0.377	-1.657	0.098	-1.365	0.116
White	-0.9156	0.352	-2.6	0.01	-1.607	-0.224
Accused	-0.9983	0.546	-1.828	0.068	-2.07	0.074
Teen	1.2785	0.607	2.105	0.036	0.086	2.471
Dead	-0.2745	0.401	-0.684	0.494	-1.063	0.514
Resident	-0.321	0.421	-0.762	0.446	-1.148	0.506
Teenage	-0.1492	0.434	-0.344	0.731	-1.002	0.704
Lone	0.2539	0.403	0.63	0.529	-0.537	1.045
Gendron	-0.1988	0.29	-0.686	0.493	-0.768	0.37
Killed	-0.3152	0.291	-1.084	0.279	-0.886	0.256
Killed	-0.3152	0.291	-1.084	0.279	-0.886	0.256

DV: Media bias score—left (-1), least (0), and right (1)

R-squared: 0.173

Adj. R-squared: 0.137

Df Model: 29

F-statistic: 4.755

Prob (F-statistic): < 0.001

N = 687

Note: CI = Confidence Interval, LL = Lower Limit, UL = Upper Limit. Bias scores range from -1 (left-centered) to 1 (right-centered).

Table 4. 11. *Regression Analysis: Predicting Media Bias Scores Based on* *<u>Victim Framing Components</u>*

Effect	Estimate (B)	SE	t	p	95%CI LL	95%CI UL
	-0.0043	0.04	-0.108	0.914	-0.083	0.074
Two	1.3668	1.069	1.279	0.201	-0.73	3.464
19	-0.1008	1.394	-0.072	0.942	-2.836	2.635
School	1.1801	0.732	1.613	0.107	-0.256	2.616
Killed	-0.4028	0.219	-1.838	0.066	-0.833	0.027
Our	-1.1553	0.513	-2.251	0.025	-2.162	-0.148
Shooting	-2.0216	1.083	-1.867	0.062	-4.147	0.103
Dead	0.2745	0.242	1.134	0.257	-0.2	0.749
My	-0.1099	0.865	-0.127	0.899	-1.808	1.588
Grade	0.5411	0.505	1.071	0.285	-0.451	1.533
Their	-0.3886	0.849	-0.458	0.647	-2.054	1.277
Old	-0.1797	0.694	-0.259	0.796	-1.542	1.183
Uvalde	0.1787	0.405	0.441	0.659	-0.616	0.974
Young	-0.4968	0.436	-1.138	0.255	-1.353	0.36
Your	-1.2828	0.596	-2.152	0.032	-2.452	-0.113
One	0.432	0.446	0.968	0.333	-0.443	1.308
Four	-0.4022	0.671	-0.599	0.549	-1.719	0.915
Her	-0.0988	0.442	-0.223	0.823	-0.967	0.769

Other	-0.2837	0.585	-0.485	0.628	-1.431	0.864
Parents	-1.2698	0.502	-2.529	0.012	-2.255	-0.285
20	-0.7415	0.419	-1.768	0.077	-1.564	0.082
Innocent	-0.3469	0.251	-1.381	0.168	-0.84	0.146
Many	-0.5974	0.347	-1.722	0.085	-1.278	0.083
More	0.463	0.397	1.166	0.244	-0.316	1.242
21	1.0512	0.306	3.438	0.001	0.451	1.651
Elementary	0.2612	0.834	0.313	0.754	-1.374	1.897
Texas	1.6065	0.804	1.998	0.046	0.028	3.185
Three	0.0291	0.309	0.094	0.925	-0.577	0.635
Nineteen	0.1059	0.312	0.339	0.735	-0.507	0.719
Six	-0.0735	0.384	-0.191	0.848	-0.827	0.68
Died	-0.2109	0.437	-0.483	0.629	-1.067	0.646
Murdered	-0.0274	0.348	-0.079	0.937	-0.711	0.656
Several	0.6487	0.433	1.498	0.134	-0.201	1.498
Slain	0.3527	0.387	0.911	0.362	-0.407	1.112
Little	0.2567	0.571	0.45	0.653	-0.864	1.377
14	-0.6262	0.331	-1.891	0.059	-1.276	0.024
His	-1.1861	0.532	-2.228	0.026	-2.231	-0.142
Fellow	0.1819	1.29	0.141	0.888	-2.35	2.714
18	1.447	0.743	1.947	0.052	-0.011	2.905
First	0.3043	0.841	0.362	0.718	-1.346	1.955

Terrified	1.5429	0.76	2.029	0.043	0.051	3.035
Were	0.0735	0.499	0.147	0.883	-0.906	1.053
Called	-1.052	0.69	-1.524	0.128	-2.407	0.303
2	-0.1093	0.339	-0.323	0.747	-0.774	0.555
Shot	1.1573	0.617	1.877	0.061	-0.053	2.367
American	-0.3305	0.755	-0.438	0.662	-1.812	1.151
Former	-0.8099	0.53	-1.529	0.127	-1.849	0.229
Own	-0.1902	0.328	-0.579	0.563	-0.835	0.454
Armed	-0.6121	0.395	-1.549	0.122	-1.388	0.163
Survived	-0.6154	0.482	-1.277	0.202	-1.561	0.33
People	0.1774	0.358	0.496	0.62	-0.525	0.88
17	0.3936	0.361	1.09	0.276	-0.315	1.102
Are	-0.2139	0.512	-0.418	0.676	-1.219	0.791
College	-0.4599	0.538	-0.855	0.393	-1.515	0.596
Eight	0.0433	0.346	0.125	0.9	-0.635	0.722
Injured	-0.7738	0.475	-1.628	0.104	-1.706	0.159
Nine	-0.1425	0.651	-0.219	0.827	-1.421	1.136
Gunned	0.676	0.759	0.891	0.373	-0.813	2.166
Hook	-0.828	0.622	-1.331	0.184	-2.049	0.393
Small	-0.559	0.469	-1.192	0.234	-1.479	0.362
Lost	-0.0951	0.418	-0.228	0.82	-0.915	0.725
We	-1.0803	0.975	-1.108	0.268	-2.994	0.833

600	0.8703	0.476	1.829	0.068	-0.064	1.804
Most	-0.6279	0.47	-1.336	0.182	-1.55	0.294
Multiple	-0.4796	0.251	-1.914	0.056	-0.971	0.012
Dozen	-0.7928	0.464	-1.708	0.088	-1.704	0.118
Families	-0.5871	0.27	-2.178	0.03	-1.116	-0.058
Roll	-0.209	0.285	-0.732	0.464	-0.769	0.351

DV: Media bias score—left (-1), least (0), and right (1)

R-squared: 0.108

Adj. R-squared: 0.053

Df Model: 67

F-statistic: 1.947

Prob (F-statistic): < 0.001

N = 1143

Note: CI = Confidence Interval, LL = Lower Limit, UL = Upper Limit. Bias scores range from -1 (left-centered) to 1 (right-centered).

Table 4. 12. *Regression Analysis: Predicting Media Bias Scores Based on Victim Clusters*

Effect	Estimate (B)	SE	t	p	95% CI LL	95% CI UL
Intercept	-0.0539	0.037	-1.453	0.147	-0.127	0.019
Younger age	0.6555	0.537	1.221	0.222	-0.398	1.709
Older age	-2.1169	0.935	-2.264	0.024	-3.952	-0.282
Our [victims]	-1.8296	0.541	-3.38	0.001	-2.892	-0.767
Your [victims]	-1.4993	0.596	-2.516	0.012	-2.668	-0.33
Innocence and vulnerability	0.1587	0.638	0.249	0.804	-1.093	1.41
Tragic loss	-0.8062	0.459	-1.758	0.079	-1.706	0.094
Humanization	-0.3888	0.396	-0.983	0.326	-1.165	0.387
Dehumanization	1.8085	0.878	2.06	0.04	0.086	3.531

DV: Media bias score—left (-1), least (0), and right (1)

R-squared: 0.023

Adj. R-squared: 0.017

Df Model: 8

F-statistic: 3.428

Prob (F-statistic): < 0.001

N = 1154

Note: CI = Confidence Interval, LL = Lower Limit, UL = Upper Limit. Bias scores range from -1 (left-centered) to 1 (right-centered).

Appendix 4.6. Evaluation of GPT's Clustering Performance Against Human Judgments

Table 4. 13. *Evaluation of GPT-4o's Clustering Performance for Shooter Framing Components (Without Context)*

	Cohen's Kappa (between human coders)	**GPT's performance against human judgment**		
Clusters		**Precision**	**Recall**	**F1-score**
Older age	0.718	0.67	0.8	0.73
Young age	0.93	0.57	1	0.73
Other	0.723	0.82	0.59	0.68
Allegation doubt	1	1	0.4	0.57
Family-background	0.827	0.5	0.6	0.55
Mental health	0.86	0.44	0.73	0.55
Threat	0.674	0.42	0.56	0.48
Allegation certainty	0.787	0.28	0.62	0.38
Action attribution	0.833	0.28	0.54	0.37
Bullying	1	0.33	0.33	0.33
Race/ethnicity	0.906	0.21	0.8	0.33
Religion & culture	1	0.17	1	0.29
Stereotypes	0.642	0.4	0.22	0.29
Humanization	1	0.08	1	0.15
Immigration	No value	0	0	0
Terrorism/Terror	No value	0	0	0

Table 4. 14. *Evaluation of GPT-4o's Clustering Performance for Shooter Framing Components (With Context)*

Clusters	**Cohen's Kappa** (between human coders)	**GPT's performance against human judgment**		
		Precision	**Recall**	**F1-score**
Other	1	0.9	0.62	0.74
Allegation doubt	1	0.86	0.56	0.68
Mental health	1	0.53	0.83	0.65
Younger age	0.935	0.48	0.98	0.64
Allegation certainty	1	0.67	0.44	0.53
Bullying	1	1	0.33	0.5
Threat	1	0.34	0.97	0.5
Action attribution	1	0.14	0.9	0.24
Older age	0.658	0.5	0.1	0.16
Family background	No value	0	0	0
Humanization	No value	0	0	0
Immigration	No value	0	0	0
Race/ethnicity	No value	0	0	0
Religion and culture	No value	0	0	0
Stereotypes	No value	0	0	0
Terrorism/terror	1	0	0	0

Table 4. 15. *Evaluation of GPT-4o's Clustering Performance for Victim Framing Components*

Clusters	**Cohen's Kappa** (between two human coders)	**GPT's performance against human judgment**		
		Precision	**Recall**	**F1-score**
Our [victims]	0.919	0.8	0.98	0.88
Your [victims]	0.857	0.73	0.94	0.82
Tragic loss	0.958	0.74	0.87	0.8
Younger age	0.961	0.36	1	0.53
Other	0.87	0.88	0.37	0.52
Heroism	-0.011	0.36	0.83	0.5
Innocence and vulnerability	1	0.13	0.74	0.22
Older age	0.796	0.17	0.32	0.22
Humanization	1	0.1	0.56	0.16
Dehumanization	0	0.01	1	0.02
Bullying	No value	0	0	0

Bibliography

Aarøe, L. (2011). Investigating frame strength: The case of episodic and thematic frames. *Political Communication*, *28*(2), 207–226. https://doi.org/10.1080/10584609.2011.568041

Ad Fontes Media. (n.d.). Retrieved from https://adfontesmedia.com/

Akyürek, A. F., Guo, L., Elanwar, R., Ishwar, P., Betke, M., & Wijaya, D. T. (2020, January). Multi-label and multilingual news framing analysis. In *Proceedings of the 58th annual meeting of the association for computational linguistics.* https://par.nsf.gov/servlets/purl/10217943

Ali, M. & Hassan, N. (2022). A survey of computational framing analysis approaches. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).* https://aclanthology.org/2022.emnlp-main.633/

Ali, M., Hassan, N. (August, 2023). *Semantic-based unsupervised framing analysis (SUFA): A Novel approach for computational framing analysis.* Top Method Paper presented at the *Association for Education in Journalism and Mass Communication (AEJMC) Conference* in D.C., USA.

American Public Health Association. (n.d.). *Gun Violence.* https://www.apha.org/topics-and-issues/gun-violence

Baden, C. (2018). Reconstructing frames from intertextual news discourse: A sematic network approach to news framing analysis. In P. D'Angelo (Eds), *Doing News Framing Analysis II* (2nd ed., pp. 3–26). https://doi.org/10.4324/9781315642239

Barry, C. L., McGinty, E. E., Vernick, J. S., & Webster, D. W. (2013). After Newtown—public opinion on gun policy and mental illness. *New England Journal of Medicine, 368*(12), 1077–1081. https://psycnet.apa.org/doi/10.1056/NEJMp1300512

Bednarek, M., & Carr, G. (2021). Computer-assisted digital text analysis for journalism and communications research: Introducing corpus linguistic techniques that do not require programming. *Media International Australia, 181*(1), 131-151. https://doi.org/10.1177/1329878X20947124

Bhatia, V., Akavoor, V. P., Paik, S., Guo, L., Jalal, M., Smith, A., ... & Wijaya, D. T. (2021, November). OpenFraming: Open-sourced Tool for Computational Framing Analysis of Multilingual Data. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 242-250). https://aclanthology.org/2021.emnlp-demo.28

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Boesman, J., & Van Gorp, B. (2018). Driving the frame: How news values, news pegs, and story angles guide journalistic frame building. In *Doing news framing analysis II* (pp. 112–134). Routledge.

Borah, P. (2008). Examining media content: a case study of newspaper coverage of dowry in India, 1999–2006. *Asian Journal of Communication*, *18*(4), 379–395. https://doi.org/10.1080/01292980802413110

Boydstun, A. E., Card, D., Gross, J., Resnick, P., & Smith, N. A. (2014). *Tracking the development of media frames within and across policy issues.* https://kilthub.cmu.edu/ndownloader/files/11903384

Boydstun, A. E., Gross, J. H., Resnik, P., & Smith, N. A. (2013, September). Identifying media frames and frame dynamics within and across policy issues. In *New Directions in Analyzing Text as Data Workshop,* London. http://faculty.washington.edu/jwilker/559/frames-2013.pdf

Brennen, B. S. (2017). *Qualitative research methods for media studies*. Routledge.

Browne, S.H. (2009). *Close textual analysis: Approaches and applications* (J. Kuypers, Eds.). *Rhetorical criticism: Perspectives in action* (pp. 63–76). Lexington Books.

Burscher, B., Vliegenthart, R., & Vreese, C. H. D. (2016). Frames beyond words: Applying cluster and sentiment analysis to news coverage of the nuclear power issue. *Social Science Computer Review, 34*(5), 530-545. https://doi.org/10.1177/0894439315596385

Card, D., Boydstun, A., Gross, J. H., Resnik, P., & Smith, N. A. (2015, July). The media frames corpus: Annotations of frames across issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing* (Volume 2: Short Papers) (pp. 438–444). https://aclanthology.org/P15-2072.pdf

Cassidy, W. P., La France, B. H., & Babin, S. (2018). Routine adjustments: How journalists framed the Charleston shootings. *International Journal of Communication,* 12, 21.

Center for Gun Violence Solutions at John Hopkins. (n.d.). https://publichealth.jhu.edu/center-for-gun-violence-solutions/research-reports/the-public-health-approach-to-prevent-gun-violence

Chyi, H. I., & McCombs, M. (2004). Media salience and the process of framing: Coverage of the Columbine school shootings. *Journalism & Mass Communication Quarterly, 81*(1), 22–35. https://journals.sagepub.com/doi/pdf/10.1177/107769900408100103

Clyde, D. & Miranda, S. (2022, June 25). Biden signs gun safety bill into law. *The National Public Radio.* https://www.npr.org/2022/06/25/1107626030/biden-signs-gun-safety-law

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407. https://psycnet.apa.org/doi/10.1037/0033-295X.82.6.407

Creswell, J. W., & Báez, J. C. (2020). *30 essential skills for the qualitative researcher.* Sage Publications.

D'angelo, P. (Ed.) (2018). *Doing news framing analysis II. Empirical and theoretical perspectives*. Routledge.

D'Angelo, P. (2017). Framing: Merua frames. In P. Rossler, C. Hoffner, & L. van Zoonen (Eds.), 17, International encyclopedia of media effects (pp. 634–644). Hoboken, NJ: WileyBlackwell. doi: 10.1002/9781118783764.wbieme004

D'Angelo, P. (2018). Prologue. In P. D'Angelo (Ed.), *Doing news framing analysis II: Empirical and theoretical perspectives* (2nd ed., pp. xxiii-xl). Routledge. https://doi.org/10.4324/9781315642239

Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: A neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–220. https://doi.org/10.1093/cercor/bhr111

Dickerson, D. L. (2001). Framing "political correctness": The New York Times' tale of two professors. In *Framing public life: Perspectives on media and our understanding of the social world* (pp. 163–174). Routledge. https://doi.org/10.4324/9781410605689

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics, 41*(6), 570-606. https://doi.org/10.1016/j.poetic.2013.08.004

El-Bawab, N. (2022, June 07). America's gun violence problem by the numbers. *The ABC News.* https://abcnews.go.com/US/americas-gun-violence-problem-numbers/story?id=85136358

Emelu, M. N. (2023). The US cable televisions' framing of mass shooting: a grounded discovery of competing narratives. *Frontiers in communication*, *8,* 1174946. https://doi.org/10.3389/fcomm.2023.1174946

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Entman, R. M. (2003). Cascading activation: Contesting the White House's frame after 9/11. *Political Communication, 20*(4), 415-432. https://doi.org/10.1080/10584600390244176

Fairhurst, G. T. (2005). Reframing the art of framing: Problems and prospects for leadership. *Leadership, 1*(2), 165–185. https://doi.org/10.1177/1742715005051857

Fairhurst, G., & Sarr, R. (1996). The art of framing. San Francisco: Jossey-Bass.

Festinger, L. (1957). *A Theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Field, A., Kliger, D., Wintner, S., Pan, J., Jurafsky, D., & Tsvetkov, Y. (2018). Framing and agenda-setting in russian news: a computational analysis of intricate political strategies. arXiv preprint arXiv:1808.09386. https://doi.org/10.48550/arXiv.1808.09386

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American journal of sociology*, *95*(1), 1-37. https://www.jstor.org/stable/2780405

Gamson, W. A., Croteau, D., Hoynes, W., & Sasson, T. (1992). Media images and the social construction of reality. *Annual review of sociology*, *18*(1), 373–393. https://doi.org/10.1146/annurev.so.18.080192.002105

Ghanem, S. (1997). Filling in the tapestry: The second level of agenda setting. In M.E. McCombs, D.L. Shaw, & D.H. Weaver (Eds.), *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory* (pp. 3–14). Routledge. https://doi.org/10.4324/9780203810880

Gilardi, F., Shipan, C. R., & Wüest, B. (2021). Policy diffusion: The issue-definition stage. *American Journal of Political Science, 65*(1), 21-35. https://doi.org/10.1111/ajps.12521

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience.* Harvard University Press.

Golan, G. (2021, February 19). What is News framing? An informal conversation among framing scholars. *The YouTube.* https://www.youtube.com/watch?v=mArApGS-p1I&t=57s

Greussing, E., & Boomgaarden, H. G. (2017). Shifting the refugee narrative? An automated frame analysis of Europe's 2015 refugee crisis. *Journal of Ethnic and Migration Studies, 43*(11), 1749–1774. https://doi.org/10.1080/1369183X.2017.1282813

Guo, L., Su, C., Paik, S., Bhatia, V., Akavoor, V. P., Gao, G., ... & Wijaya, D. (2022). Proposing an Open-Sourced Tool for Computational Framing Analysis of Multilingual Data. *Digital Journalism, 1*-22. https://doi.org/10.1080/21670811.2022.2031241

Heider, F. (1958). *The psychology of interpersonal relations (2nd edition)*. New York: Wiley.

Hertog, J. K., & McLeod, D. M. (2001). A multiperspectival approach to framing analysis: A field guide. In *Framing public life: Perspectives on media and our understanding of the social world* (pp. 141–163). Routledge. https://doi.org/10.4324/9781410605689

Hillston, J. (2024, May). What does Performance Mean for Large Language Models? In *Proceedings of the 15th ACM/SPEC International Conference on Performance Engineering* (pp. 118–118). https://doi.org/10.1145/3629526.3649130

Iyengar, S. (1994). Is anyone responsible? Chicago: University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/I/bo3684515.html

Jacobo, J. & El-Bawab, N. (2022, December 12). Timeline: How the shooting at a Texas elementary school unfolded. *The ABC News.* https://abcnews.go.com/US/timeline-shooting-texas-elementary-school-unfolded/story?id=84966910

Ji, Y., & Smith, N. (2017). Neural discourse structure for text categorization. *arXiv preprint* arXiv:1702.01829.

Jing, E., & Ahn, Y. Y. (2021). Characterizing partisan political narrative frameworks about COVID-19 on Twitter. *EPJ Data Science, 10*(1), 53. https://doi.org/10.1140/epjds/s13688-021-00308-4

Johnson, K., Jin, D., & Goldwasser, D. (2017, July). Leveraging behavioral and social information for weakly supervised collective classification of political discourse on Twitter. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (pp. 741-752). http://dx.doi.org/10.18653/v1/P17-1069

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: the attribution process in social psychology, in L. Berkowitz (ed.), *Advances in experimental social psychology* (Volume 2, pp. 219–266), Academic Press.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341–350. https://psycnet.apa.org/doi/10.1037/0003-066X.39.4.341

Kang, Y., & Yang, K. C. (2021). Communicating racism and xenophobia in the era of Donald Trump: A computational framing analysis of the US-Mexico cross-border wall discourses: Special Issue on Donald Trump Era and Communicating Race in America. *Howard Journal of Communications*, 1-20. https://doi.org/10.1080/10646175.2021.1996491

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107–128. https://psycnet.apa.org/doi/10.1037/h0034225

Kellner, D. (2022). The Uvalde, Texas school shooting massacre. *Educational Philosophy and Theory, 57*(2), 91–95. https://doi.org/10.1080/00131857.2022.2094763

Khanehzar, S., Turpin, A., & Mikolajczak, G. (2019). Modeling political framing across policy issues and contexts. In Proceedings of *the 17th Annual Workshop of the Australasian Language Technology Association* (pp. 61-66). https://aclanthology.org/U19-1009.pdf

Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, *160*(1), 3-24. https://informatica.si/index.php/informatica/article/viewFile/148/140

Kwak, H., An, J., & Ahn, Y. Y. (2020, July). A systematic media frame analysis of 1.5 million New York Times articles from 2000 to 2017. In *12th ACM*

Conference on Web Science (pp. 305-314). https://dl.acm.org/doi/abs/10.1145/3394231.3397921

Kwak, H., An, J., Jing, E., & Ahn, Y. Y. (2021). FrameAxis: characterizing microframe bias and intensity with word embedding. *PeerJ Computer Science, 7,* e644. https://peerj.com/articles/cs-644/

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., ... & Van Alstyne, M. (2009). Computational social science. *Science, 323*(5915), 721–723. https://doi.org/10.1126/science.1167742

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological bulletin*, *85*(5), 1030. https://psycnet.apa.org/doi/10.1037/0033-2909.85.5.1030

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational behavior and human decision processes*, *76*(2), 149–188. https://doi.org/10.1006/obhd.1998.2804

Li, P., Cho, H., Qin, Y., & Chen, A. (2021). # MeToo as a Connective Movement: Examining the Frames Adopted in the Anti-Sexual Harassment Movement in China. *Social Science Computer Review, 39*(5), 1030–1049. https://doi.org/10.1177/0894439320956790

Liu, S., Guo, L., Mays, K., Betke, M., & Wijaya, D. T. (2019, January). Detecting frames in news headlines and its application to analyzing news framing trends surrounding U.S. gun violence. In *Proceedings of the 23rd Conference on*

Computational Natural Language Learning (CoNLL). https://par.nsf.gov/biblio/10219163

Livingston, A. (2022, May 29). Uvalde school shooting: ‘Do something!’ Biden is urged as he leaves Uvalde church. *The Texas Tribune.* https://www.texastribune.org/2022/05/29/biden-uvalde-visit/

Mass shootings. (n.d.). *The Gun Violence Archive.* https://www.gunviolencearchive.org/mass-shooting?sort=desc&order=Victims%20Killed

McCombs, M., Llamas, J. P., Lopez-Escobar, E., & Rey, F. (1997). Candidate images in Spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly*, *74*(4), 703-717. https://doi.org/10.1177/107769909707400404

McGinty, E. E., Webster, D. W., Jarlenski, M., & Barry, C. L. (2014). News media framing of serious mental illness and gun violence in the United States, 1997-2012. *American Journal of Public Health*, *104*(3), 406–413. https://doi.org/10.2105/AJPH.2013.301557

McKee, A. (2001). A beginner's guide to textual analysis. *Metro Magazine*, pp. 138–149. https://eprints.qut.edu.au/41993/

McKeever, B. W., Choi, M., Walker, D., & McKeever, R. (2022). Gun violence as a public health issue: Media advocacy, framing and implications for communication. *Newspaper Research Journal*, *43*(2), 138–154. https://doi.org/10.1177/07395329221090497

McLeod, S. A. (2010). *Attribution theory.* Retrieved from https://www.simplypsychology.org/attribution-theory.html

Media Bias Fact Check (n.d.). Retrieved from https://mediabiasfactcheck.com/

Mendelsohn, J., Budak, C., & Jurgens, D. (2021). Modeling framing in immigration discourse on social media. *arXiv preprint* arXiv:2104.06443.

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38*(11), 39-41. https://doi.org/10.1145/219717.219748

Morin, A. (2016). Framing terror: The strategies newspapers use to frame an act as terror or crime. *Journalism & Mass Communication Quarterly, 93*(4), 986–1005. https://doi.org/10.1177/1077699016660720

Naderi, N., & Hirst, G. (2017). Classifying frames at the sentence level in news articles. *Policy, 9*, 4-233. https://doi.org/10.26615/978-954-452-049-6_070

Nelson, T. E., Clawson, R. A., & Oxley, Z. M. (1997). Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review, 91*(3), 567–583. https://doi.org/10.2307/2952075

Nguyen, V. A. (2015). *Guided probabilistic topic models for agenda-setting and framing* (Doctoral dissertation, University of Maryland, College Park). https://www.proquest.com/docview/1707358055?pq-origsite=gscholar&fromopenview=true

Nguyen, V. A., Boyd-Graber, J., Resnik, P., & Miler, K. (2015, July). Tea party in the house: A hierarchical ideal point topic model and its application to republican legislators in the 112th Congress. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International*

Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 1438–1448). https://aclanthology.org/P15-1139.pdf

Nicholls, T., & Culpepper, P. D. (2021). Computational identification of media frames: Strengths, weaknesses, and opportunities. Political Communication, 38(1-2), 159-181. https://doi.org/10.1080/10584609.2020.1812777

Odhner, K. (2024, July 10). MediaBiasFactCheck.com as a Tool for Lateral Reading. *The News Literacy Institute at Pennsylvania State University*. https://newsliteracy.psu.edu/news/mediabiasfactcheck-com-as-a-tool-for-lateral-reading

Ophir, Y., Walter, D., Arnon, D., Lokmanoglu, A., Tizzoni, M., Carota, J., ... & Nicastro, E. (2021). The framing of COVID-19 in Italian media and its relationship with community mobility: a mixed-method approach. Journal of health communication, 26(3), 161-173. https://doi.org/10.1080/10810730.2021.1899344

Pan, Z., & Kosicki, G. M. (1993). Framing analysis: An approach to news discourse. *Political Communication, 10*(1), 55–75. https://doi.org/10.1080/10584609.1993.9962963

Park, S. Y., Holody, K. J., & Zhang, X. (2012). Race in media coverage of school shootings: A parallel application of framing theory and attribute agenda setting. *Journalism & Mass Communication Quarterly*, *89*(3), 475–494. https://doi.org/10.1177/1077699012448873

Peck, J. & Goodman, J.D. (2022, May 24). Uvalde elementary school shooting: Shooting at elementary school devastates community in South Texas. *The*

New York Times. https://www.nytimes.com/live/2022/05/24/us/shooting-robb-elementary-uvalde#shooting-texas-elementary-school

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology*, 57, 27–53. https://doi.org/10.1146/annurev.psych.56.091103.070234

Reese, S. D., & Shoemaker, P. J. (2018). A media sociology for the networked public sphere: The hierarchy of influences model. In *Advances In Foundational Mass Communication Theories* (pp. 96-117). Routledge.

Reese, S. D., Gandy Jr, O. H., & Grant, A. E. (Eds.). (2001). *Framing public life: Perspectives on media and our understanding of the social world*. Routledge.

Reese, S.D. (2001). Prologue—Framing Public Life: A Bridging Model for Media Research. In Framing public life (pp. 7–31). Routledge. https://doi.org/10.4324/9781410605689

Remarks by President Biden on the school shooting in Uvalde, Texas. (2022, May 24). *The Biden White House Achieve.* https://bidenwhitehouse.archives.gov/briefing-room/speeches-remarks/2022/05/24/remarks-by-president-biden-on-the-school-shooting-in-uvalde-texas/

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American journal of political science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Saldaña, J. (2016). *The coding manual for qualitative researchers.* SAGE.

Sallam, M. (2021). COVID-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates. Vaccines, 9(2), 160. https://doi.org/10.3390/vaccines9020160

Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health*, *18*(2), 179–183. https://doi.org/10.1002/nur.4770180211

Sanderink, L. (2020). Shattered frames in global energy governance: Exploring fragmented interpretations among renewable energy institutions. Energy research & social science, 61, 101355. https://doi.org/10.1016/j.erss.2019.101355

Sandoval, E. (2023, May 24). A year after the Uvalde massacre: Did anything change? *The New York Times.* https://www.nytimes.com/2023/05/24/us/uvalde-shooting-fallout.html

Sanfilippo, A., Franklin, L., Tratz, S., Danielson, G., Mileson, N., Riensche, R., & McGrath, L. (2008). Automating frame analysis. In *Social computing, behavioral modeling, and prediction* (pp. 239-248). Springer US.

Slovic, P. (2007). "If I look at the mass I will never act": Psychic numbing and genocide. *Judgment and Decision making*, *2*(2), 79–95. https://psycnet.apa.org/doi/10.1017/S1930297500000061

Sturdza, M. D. (2018). Automated framing analysis: A rule based system for news media text. *Journal of Media Research-Revista de Studii Media, 11*(32), 94-110. https://www.mrjournal.ro/docs/R2/32jmr7T.pdf

Supran, G., & Oreskes, N. (2021). Rhetoric and frame analysis of ExxonMobil's climate change communications. One Earth, 4(5), 696-719. https://doi.org/10.1016/j.oneear.2021.04.014

Swenson, J. (1990). News coverage of the abortion issue. Framing changes in the 1980s. Paper presented to the Committee on the Status of Women. *Association for Education in Journalism and Mass Communication*.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S.

Tankard, J.W., Hendrickson, L, Silberman, J., Bliss, K., & Ghanem, S. (1991). Media frames: Approaches to conceptualization and measurement. Paper presented to the annual meeting *of the Association for Education in Journalism and Mass Communication*, Boston, MA.

Tewksbury, D., & Riles, J. M. (2018). Framing in an interactive news environment. In *Doing News Framing Analysis II* (pp. 137–162). Routledge. https://doi.org/10.4324/9781315642239

The Economic Cost of Gun Violence. (2022, July). *The Everytown Research.* https://everytownresearch.org/report/the-economic-cost-of-gun-violence/#introduction

Tourni, I. C., Guo, L., Hu, H., Halim, E., Ishwar, P., Daryanto, T., ... & Wijaya, D. T. (2024). Detecting frames in news headlines and lead images in US gun violence coverage. *arXiv preprint arXiv:2406.17213*.

Tourni, I., Guo, L., Daryanto, T. H., Zhafransyah, F., Halim, E. E., Jalal, M., ... & Wijaya, D. T. (2021, November). Detecting Frames in News Headlines and

Lead Images in US Gun Violence Coverage. In *Findings of the Association for Computational Linguistics: EMNLP 2021* (pp. 4037–4050). https://aclanthology.org/2021.findings-emnlp.339

van Atteveldt, W., & Peng, T. Q. (2018). When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science. *Communication Methods and Measures*, *12*(2-3), 81-92. https://doi.org/10.1080/19312458.2018.1458084

van Atteveldt, W., Sheafer, T., & Shenhav, S. (2013, May). Automatically extracting frames from media content using syntacting analysis. *In Proceedings of the 5th Annual ACM Web Science Conference,* pp. 423–430. https://dl.acm.org/doi/pdf/10.1145/2464464.2464504

van Dijk, T. A. (2023). Analyzing frame analysis: A critical review of framing studies in social movement research. *Discourse Studies, 25*(2), 153–178. https://doi.org/10.1177/14614456231155080

van Dijк, T. (1985). Structures of News in the Press. In T. Dijk (Ed.), *Discourse and Communication: New Approaches to the Analysis of Mass Media Discourse and Communication* (pp. 69–93). Berlin, New York: https://doi.org/10.1515/9783110852141.69

Walter D. & Ophir Y. (2019). News frame analysis: An inductive mixed method computational approach*.* *Communication Methods and Measures, 13*(4), 248–266. http://dx.doi.org/10.1080/19312458.2019.1639145

Walter, D., & Ophir, Y. (2021). Strategy framing in news coverage and electoral success: An analysis of topic model networks approach. *Political*

Communication, 38(6), 707–730. https://doi.org/10.1080/10584609.2020.1858379

Wilson, R. F. (2023). Unintentional firearm injury deaths among children and adolescents aged 0–17 years—National Violent Death Reporting System, United States, 2003–2021. *MMWR. Morbidity and Mortality Weekly Report*, *72*.

Yang, K. C., & Kang, Y. (2020). Framing national security concerns in mobile telecommunication infrastructure debates: A text mining study of Huawei. In *Huawei Goes Global* (pp. 319-339). Palgrave Macmillan, Cham. https://link.springer.com/chapter/10.1007/978-3-030-47579-6_14

Ylä-Anttila, T., Eranti, V., & Kukkonen, A. (2021). Topic modeling for frame analysis: A study of media debates on climate change in India and USA. *Global Media and Communication, 18*(1), 91–112. https://doi.org/10.1177/17427665211023984

Yu, Q. (2022, July). “Again, Dozens of Refugees Drowned”: A Computational Study of Political Framing Evoked by Presuppositions. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop* (pp. 31-43). http://dx.doi.org/10.18653/v1/2022.naacl-srw.5

Zheng, J., & Scardino, F. (2021). News Media Coverage of Mass Shootings: How the Political Stance of News Media Influences the Coverage of Mass Shooters. *Journal of Student Research, 10*(1).

https://doi.org/10.47611/jsrhs.v10i1.1332Ziems, C., & Yang, D. (2021). *To protect and to serve? Analyzing entity-centric framing of police violence.* arXiv preprint arXiv:2109.05325.

Ziems, C., & Yang, D. (2021). *To protect and to serve? Analyzing entity-centric framing of police violence.* arXiv preprint arXiv:2109.05325. https://doi.org/10.48550/arXiv.2109.05325