

THE LINGUISTIC FEATURES OF BIAS IN THE NEWS

By Amanda Black

A Dissertation

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in Applied Linguistics

Northern Arizona University

January 2024

Approved:

William Crawford, Ph.D., Co-chair

Douglas Biber, Ph.D., Co-chair

Jesse Egbert, Ph.D.

Eniko Csomay, Ph.D.

ABSTRACT

THE LINGUISTIC FEATURES OF BIAS IN THE NEWS

AMANDA BLACK

This dissertation aims to determine what linguistic characteristics can reveal about bias in news and whether variation in linguistic characteristics, namely grammatical and lexical features, happen systematically across clines of political leaning and extent (i.e., severity of bias). Through an examination of key features and keywords in news texts divided into groups of political leanings of left and right and divided into groups by extent (extreme bias and no bias), this aim is accomplished. This study demonstrates that media bias, namely gatekeeping, coverage, and presentation bias can be examined via corpus linguistic methods and suggests that lexical and grammatical information are contributing to perceptions of direction and extent of bias. Importantly, it also demonstrates ways in which a researcher can avoid making subjective decisions about bias by relying on exploratory methods of register variation.

The research in this study is conducted through six major steps: (1) compilation of a corpus of newspaper publications for which topic and time are closely controlled, (2) the collection of reader perceptions on the extent and direction of bias of each collected newspaper text, (3) a nuanced examination of key grammatical features and a comparison of their functions across biased and unbiased texts, (4) an analysis and comparison of the keywords that occur in both biased and unbiased texts (5) a comparison of key features across texts perceived as right (a.k.a. conservative) to those perceived as left (a.k.a. liberal) (6) and an analysis and comparison of keywords that occur in texts perceived as conservative to those perceived as liberal.

Results suggest systematic linguistic differences in the following ways: key features of left leaning texts show that this group maintains a formal tone, involvement at the group level, and covertly reveals stance via non-finite clauses while the keywords reveal a concern for social issues (particularly those related to equality), a sentiment of aggression and political unrest, and a more frequent on Trump. Conversely, right leaning texts are colloquial in tone, involved at the individual level, and covertly reveal stance via reported speech. The keywords for the right leaning group demonstrate a concern for political issues (particularly illegal immigration and a loss of constitutional freedoms), the actions of former presidents, and a negative evaluation of the liberal left. Texts perceived as extreme in their bias include features marked for epistemic and attitudinal stance, clausal and phrasal elaboration, description, emphasis, and evaluation and the keywords reveal a concern for both political and social issues, political actors, and opposing negative evaluations of both democrats and republicans. Key features and keywords for texts rated as no bias include features related to reporting what happened (who, what, when, and where). The differences are, for the most part, attributable to a functional or evaluative difference related to direction or extent of bias.

TABLE OF CONTENTS

ABSTRACT	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
DEDICATION	viii
Chapter 1. Introduction.....	1
1.1 Background & Motivation.....	1
1.2 Academic Research on Biased News.....	4
1.2.1 Misinformation, disinformation, fake news, and media bias.....	4
1.2.2 Social sciences and media bias.....	6
1.2.3 NLP and media bias.....	8
1.2.4 Linguistics and media bias.....	12
1.3 Goals of the Dissertation.....	15
1.4 Outline of the Dissertation.....	16
Chapter 2. Literature Review of Media Bias.....	18
2.1 Introduction.....	18
2.2 Misinformation, Disinformation, Fake News, and Media Bias.....	18
2.3 Media Bias.....	22
2.3.1 Problems with defining media bias.....	22
2.3.8 Frameworks for identifying media bias.....	25
2.4 Empirical Research on Media Bias.....	28
2.4.1 Studies of gatekeeping bias.....	28
2.4.2 Studies of coverage bias.....	31
2.4.3 Studies of presentation bias.....	34
2.5 Conclusion.....	36
Chapter 3. Literature Review of Language and Bias.....	36
3.1 Introduction.....	36
3.2 NLP and Media Bias.....	37
3.3 Linguistics and Media Bias.....	42
3.3.1 Linguistic studies of social bias in news.....	42
3.3.2 Linguistic studies of political bias.....	44
3.4 The Present Study.....	46
Chapter 4. <i>SPIN</i>	48
4.1 Introduction.....	48
4.2 Compilation of the Corpus.....	48
4.2.1 Selection of publications.....	49
4.2.2 Selection of texts.....	50
4.2.3 <i>SPIN</i>	56
Chapter 5. Methods of Measuring Bias.....	57
5.1 Introduction.....	57
5.2 Defining Bias.....	58
5.3 The Collection of Perceptual Ratings.....	62

5.3.1 Mechanical Turk.....	62
5.3.2 The instrument.....	64
Chapter 6. The Raters.....	67
6.1 Introduction.....	67
6.2 The Demographics of Average Online News Readers.....	67
6.3 The Demographics of the Raters.....	67
Chapter 7. Defining Bias with Ratings.....	70
7.1 Introduction.....	70
7.2 Using Ratings to Determine Bias Groupings.....	71
7.2.1 Bias groups.....	73
7.3 Group Profiles.....	76
7.3.1 Overlap in number of texts between groups.....	77
7.3.2 Bias group and article type.....	80
7.3.3 Bias groups and publication.....	84
7.3.4 Bias groups and topic.....	88
7.4 Conclusion.....	94
Chapter 8. Key Feature Analysis.....	95
8.1 Introduction.....	95
8.2 Grammatical Feature Tagging and Selection.....	95
8.3 Methods of Key Feature Analysis.....	96
8.4 Key Features: Left V Right.....	98
8.4.1 Right-Rated articles.....	99
8.4.2 Left-Rated articles.....	104
8.5 Key Feature: No Bias Vs Extreme Bias.....	108
8.5.1 No extent articles.....	109
8.5.2 Extreme extent articles.....	111
8.6 Conclusion.....	117
Chapter 9. Keyword Analysis.....	118
9.1 Introduction.....	118
9.2 Methods of Keyword Analysis.....	120
9.2.1 Corpus frequency keyness.....	120
9.2.2 Text dispersion keyness.....	122
9.2.3 Combining both keyword methods.....	124
9.3 Results of the Keyword Analyses: Left Vs Right.....	125
9.3.1 Left -Rated texts	125
9.3.2 Right -Rated texts.....	134
9.4 Results of the Keyword Analyses: Extreme V No.....	140
9.4.1 Extreme extent texts.....	140
9.4.2 No extent texts.....	147
9.5 Conclusions.....	153
Chapter 10. Synthesis and Conclusions.....	155
10.1 Introduction.....	155
10.2 Orientation to Media Bias Studies.....	155
10.2.1 Gatekeeping bias.....	155
10.2.2 Coverage bias.....	158
10.2.3 Presentation Bias.....	159

10.3 Synthesis of Results.....	161
10.4 Limitations and Future Research.....	165
10.5 Conclusions.....	166

LIST OF TABLES

Table 4.1 Publications included in the corpus.....	50
Table 4.2 Topic + Time per year.....	51
Table 4.3 Overview of SPIN.....	57
Table 5.1 Questions that loaded positively on relevant factors from Spinde et al. (2021).....	61
Table 7.1 Agreement for Dissertation Raters and Outside Ratings for Direction of Bias	73
Table 7.2 Number of Texts (Overlapping and Unique) in each Bias Group.....	78
Table 7.3 Number of Texts per Article Type for the Entire Subsample.....	82
Table 7.4 Number of Texts per Publication for the Entire Subsample.....	84
Table 7.5 A list of Topics with Significantly higher Number of texts per Bias Group.....	92
Table 8.1 Sample size interpretations.....	97
Table 9.1 Contingency Table for Log-likelihood: Corpus Frequency.....	121
Table 9.2 Contingency Table for Log-likelihood: Text Dispersion.....	123
Table 9.3 Keyword List for Texts Rated Left	125
Table 9.4 Semantic Domains for Left Bias Keywords	126
Table 9.5 Keyword List for Texts Rated Right.....	134
Table 9.6 Semantic Domains for Right Bias Keywords.....	135
Table 9.7 Keyword List for Texts Rated Extreme Bias.....	140
Table 9.8 Overlap of Keywords between Directional Bias Groups and Extreme Bias.....	141
Table 9.9 Semantic Domains for Extreme Bias Keywords.....	141
Table 9.10 Keyword List for Texts Rated as No Bias.....	147
Table 9.11 Semantic Domains for No Bias Keywords.....	148
Table 10.1 Combined Results for Left and Right Bias Groups.....	162
Table 10.2 Combined Results for Extreme and No Bias Groups.....	163
Table B1 Articles Rated as Extremely Biased.....	196
Table C1 Articles rated as no bias.....	207
Table D1 Articles rated as Right leaning.....	212
Table E1 Articles rated as left leaning.....	224

LIST OF FIGURES

Figure 1.1 Popular memes depicting distrust in News	2
Figure 1.2 Schematic representation of study design for this dissertation	17
Figure 2.1 Venn diagram of misinformation, disinformation, and malinformation (adapted from Santos-D'amorim & Miranda, 2021).....	20
Figure 2.2 Illustrated representation of media bias (adapted from Hamborg, Donnay, and Gipp, 2019)	26
Figure 5.1 Example of Filtering Questions	65
Figure 5.2 Question on Agreement with Article's Perspective	65
Figure 5.3 Items Inquiring about Perspectives on Extent and Direction of Bias	66
Figure 6.1 Division of raters by self-identified political leaning	68
Figure 6.2 Division of Raters by Age	69
Figure 6.3 Division of Raters by Sex	69
Figure 6.4 Division of Raters by Education	70
Figure 7.1 The Division of Texts from SPIN by Extent of Bias	74
Figure 7.2 The Division of Texts from SPIN by Extent of Bias	75
Figure 7.3 Overlap of Texts for No Extent Group	79
Figure 7.4 Overlap of Texts for Extreme Extent Group	79
Figure 7.5 Overlap of Texts for Left Leaning Group	79
Figure 7.6 Overlap of Texts for Right Leaning Group	79
Figure 7.7 Illustration of Percentage of Article Type per Bias Group.....	83
Figure 7.8 Breakdown of Publication by Bias Groupings	85
Figure 7.9 Illustration of Breakdown of Groups by Publication	87
Figure 7.10 Division of Bias Groups by 2016 Topic	90
Figure 7.11 Division of Bias Groups by 2017 Topic	91
Figure 7.12 Division of Bias Groups by 2018 Topic	92
Figure 8.1 Key grammatical features for texts rated left and right	99
Figure 8.2 Key grammatical features for texts rated extreme bias and no bias	108
Figure 10.1 Gatekeeping bias per publication	157
Figure 10.2 Illustrated representation of media bias (adapted from Hamborg, Donnay, and Gipp, 2019)	167

DEDICATION



CHAPTER 1. INTRODUCTION

1.1 Background & Motivation

The political climate in the U.S. has become increasingly divisive since the 2016 election, further polarizing our already bipartisan political system and heightening accusations of ‘fake news’. A Gallup Poll conducted in 2017 demonstrates that 62% of people surveyed believe that US news sources maintain heavy political biases. Reputable newspapers like *The New York Times* and TV news stations such as *CNN* are being hailed as fake news by former President Trump while channels such as *Fox* are being deemed a functional campaign arm for Republicans and further, *Fox* is being sued for airing claims about election fraud they knew to be false. Although some argue that it is not the job of journalists to remain objective, many news sources pride themselves on impartiality as can be observed by the following quotes from the editors of *The New York Times* and *The Wall Street Journal*, respectively:

“The Op-Ed page editors [of *NYT*] do an evenhanded job of representing a range of views in the essays from outsiders they publish...”

“The *Journal's* [*WSJ*] editors stress the independence and impartiality of their reporters.”

The question of whether or not it is the responsibility of news to remain objective, or indeed whether it is possible is of course something to consider, but the fact remains there are many who contend objectivity is a “necessity” for news media (McQuail, 1992, p. 187). In addition, the news media themselves promise to uphold certain journalistic canons which include truth, fairness, objectivity, and impartiality (Society of Professional Journalism Code of Ethics, 2014).

Observe the following quote from the *Associated Press*:

“The good newspaper is fair, accurate, honest, responsible, independent and decent. Truth is its guiding principle. It avoids practices that would conflict with the ability to report and present news in a fair, accurate and unbiased manner.” (*Associated Press*, Code of Ethics)

Despite the promises of impartiality, the terms “liberal media” “leftist media”, “fake news”, “pseudoscience”, or “conspiracy news” are likely not terms that have gone unheard by most Americans as charges of biased media have been heightening since the 1980s and have recently become highly polarized (Syverson, 2017). As figure 1 below shows, popular depictions of news institutions are not positive with shows such as *South Park* directly calling out *CNN*, and interestingly, the same “we make it up” sentiment being attributed to both *Fox* and *The New York Times* in memes posted on social media sites.



Figure 1.1 Popular memes depicting distrust in News

Given that the current ethics of journalism focus heavily on impartiality and given the claims of pervasive media bias, it is important to determine whether the media is in fact displaying bias, and when they are, to understand *how* the news may be spinning information; this is important for two reasons: 1) the media has been shown to influence public policy; and 2) misplaced mistrust in the media can have detrimental effects.

The media plays a vital role in informing citizens of a deliberate democracy so understanding not only how media can display bias, but how that bias affects the public is crucial (Lippman, 1946). Traditionally, the media has been considered “the sword arm of democracy” with the paramount responsibilities of acting as a government watchdog, protecting public interests, and informing the public’s understanding of events that we cannot observe directly (Deane, 2015). Furthermore, media has been shown to influence public decisions on policy. For

example, Kahn and Kenny (2002) examined 60 senatorial campaigns and consistently demonstrated that candidates shown to be favored by newspapers were also favored by voters. This was most significant for candidates that received heavy attention in the press and by voters who attested to reading their local newspaper daily. Similar results were found by Druckman & Parkin (2005) who demonstrated that even a relatively slight positive spin towards one candidate could influence voters for senate elections. Thus, understanding whether we have access to unbiased information is integral to forming a well-balanced understanding of current events as the press plays a significant role in determining the nation's political agenda and spreading values (Chilton, 2004: 24).

Unsubstantiated mistrust in the media can also lead to detrimental effects such as one-sided information seeking as well as nonconsumption and inattention; all of which are practices that lead to filter bubbles, echo chambers, polarization of public opinion, and a generally skewed understanding of events (Kiousis, 2001; Sunstein, 2001). This issue is even more paramount today as social media creates unique circumstances for confirmation bias, and researchers are recognizing this as an era of “post-fact” – but only in regard to political and civil discourse (Rich, 2018). Misinformation, fake news, amplification of internal biases on social media, and traditionally factual sources blurring the lines between fact and opinion without clear demarcation (i.e., News-page columns and sponsored articles written to sound like news) are creating a society in which news literacy is perhaps more essential than ever but is severely lacking. This creates a scenario best put into words by Rich (2018):

The groups on each side [of the political spectrum] can become insular in their thinking and communication, creating a closed environment in which false information proliferates. Data suggest that political, social, and demographic polarization are not only

severe and worsening in the contemporary United States but also overlapping and reinforcing one another (p. 15).

If news presents information objectively, or if we as consumers make it our responsibility to be aware of the spin our news is being presented with, we can break these filter bubbles and in turn strengthen the democratic values that a free and fair press were founded upon (Dean, 2011). This dissertation can help build awareness of newspapers' spin by identifying how linguistic patterns of language use contribute to perspectives of biased reporting.

1.2 Academic Research on Biased News

1.2.1 Misinformation, Disinformation, Fake News, and Media Bias

As the previous section has shown, popular claims of media bias are becoming more prevalent in today's society. This is the case because social media has given rise to the dissemination of information at unparalleled speeds which has, unfortunately, also led to the proliferation of false information, information that has been altered to affect an individual's beliefs, or information presented to appeal to emotions over truth (Nakov & Da San Martino, 2021). According to research, news that is fabricated, altered, or misrepresentative (i.e., fake or biased news) has been shown to spread six times faster than 'real' news, and partisanship, the degree of novelty, and the emotional reactions of readers are strong predictors of belief in deceptive news (Tangri, 2021; Uscinski et al., 2016; Vosoughi et al., 2018). This situation has created unprecedented challenges for traditional news media to remain relevant, making the potential for biased reporting, especially to appeal to the partisanship or emotions of a reader, likely. The current Dominion lawsuit against *Fox News* for knowingly spreading false

information about 2020 election fraud is the perfect example of this situation occurring in the real world; all of this taken together renders examinations of misinformation pressing in academic disciplines.

The academic study of media bias and fake news fall under the larger category of *misinformation* which presents itself in a multitude of ways in news media. The most important distinction is that between *misinformation*, *disinformation*, and *malinformation* of which Fake News and Media Bias are practical incidences. *Misinformation* includes any false or misleading information regardless of intent, *disinformation* refers only to information that is intentionally false, and *malinformation* is the sensitive information (true) that is strategically used to cause harm/disadvantage (Santos-D'Amorim & Miranda, 2021; Stahl, 2006; Tandoc et al., 2018). To clarify, if a news outlet unknowingly or accidentally misleads an audience through biased reporting, this would fall under *misinformation*, but if a news outlet deliberately misleads or deceives an audience through biased reporting, this would fall under both *misinformation* and *disinformation*, if a news agency published a story they knew to be false (Fake News), this would fall only under *disinformation*, and if a news outlet knowingly used sensitive information to harass or spread hate speech, this would be *malinformation*. Due to the overlap that could potentially occur between these three distinctions, Wenzel (2019) distinguishes further between *disinformation* and *truthful disinformation* the latter of which includes information that is true, but important information is purposely excluded to manipulate the reader, a.k.a. bias by selective reporting or bias by omission (a.k.a. coverage and gatekeeping bias).

The distinction between these three is an important one in academic research as all three are closely related and refer to news that is suspected to be inaccurate, one-sided, misleading, or fabricated with the distinction between the three coming down to intent (Grieve & Woodfield,

2023); *disinformation* is the presentation of untrue information with the intent to deceive, *misinformation* is the presentation of untrue or misrepresented information without the explicit intent to deceive, and *malinformation* is the presentation of true information with the intent to harm. The most important facets of misinformation for the present study are Fake News and media bias which are, in theory, distinct with Fake News being news which is intentionally untrue, and media bias being news that offers a skewed perspective (Davis, 1990; Santos-D'Amorim & Miranda, 2021). However, this demarcation is unlikely to remain this clear cut in actual instances of news as a story that is mostly fabricated is likely to contain some truthful information, and a story that is 'real news' could easily contain false information (Santos-D'Amorim & Miranda, 2021). Further, it is difficult to always know an author's intent, therefore, the majority of academic papers on fake and biased news are likely examining news that actually falls somewhere in the intersection between *misinformation* and *disinformation*. For this reason, and because the present study will examine professionally published news articles and have no way of collecting author intent, existing empirical research on disinformation (i.e., Fake News) and misinformation (i.e., media bias) will be used to inform this study.

1.2.2 Social Sciences and Media Bias

A good majority of research examining misinformation and media bias come from fields of social science particularly those in fields of communication, politics, and psychology. These fields tend to be concerned with different facets of media bias. For example, psychology is mostly concerned with human perception of bias and identifying personality characteristics that predict belief in falsified news (i.e., how cognitive dissonance and confirmation bias relate to one-sided information seeking). Political and communication studies are concerned with identifying and characterizing media bias. An important characterization from these fields that

will inform this dissertation is that bias can manifest in news texts in three distinct ways: 1) gatekeeping bias or which news stories make it to print versus those that do not, 2) coverage bias or how much time and space is devoted to one issue or actor over another, and 3) presentation bias or how an issue or actor is presented in the news story (D'Alessio & Allen, 2000).

These categorizations of media bias have led to several subsequent empirical examinations that help paint a clear picture of media bias and factors that may cause biased news. For example, examinations of gatekeeping bias have demonstrated that economic factors, political affiliation, preexisting beliefs about social groups, and target audience are important factors in determinations of what makes it into the news and what does not (D'Alessio & Allen, 2000; Groseclose & Milyo, 2005; Gruenwald et al., 2009; Lundman, 2003; McCarthy et al., 1996; and Oliver & Maney, 2000). Studies on coverage bias have demonstrated that news entities can and do favor issues or actors over others in the amount of coverage they devote, and that this is usually done to affect political or social policy (Al-Gamde & Tenbrink, 2020; Entman, 2010; Niven, 2002; Papacharissi and Oliveira, 2008; Shultziner & Stukalin, 2020). Studies on presentation bias demonstrate that news often relies on negative sentiment to express bias (Kweon, 2017; Niven, 2002).

From studies conducted under the three categorizations of media bias, we can see that media bias exists and that it has real world consequences. However, these studies share the common limitations of looking at very specific, and usually politically charged, instances of biased news and then overgeneralizing their results back to all news media. In addition, these charges of bias often relate to a political affiliation, but very little control (other than examining time and space devoted to political candidates of opposing affiliations) is put in place to determine whether these are actual reflections of political bias.

In addition, two of the aforementioned types of media bias, coverage bias and presentation bias, directly involve examinations of language, yet very few of these studies draw on previous linguistic research to conduct their examinations of language. This is limiting because linguistic methodologies, particularly corpus methods and methods of register variation, could be very informative for examinations into these types of bias coverage. For example, a corpus built to represent news of a wide variety of topics and across multiple publications could provide more generalizable results for gatekeeping bias while examinations of key linguistic features and keywords would statistically highlight whether issues or actors are focused on more than any other. Key linguistic features and keywords would also demonstrate how entities and actors are presented in news articles. In addition, proper control for topic, political leaning, and register would allow for identifications in linguistic variation to be more reliably attributable to political bias. Thus, the present study aims to combine linguistic methodologies with previous characterizations of media bias to determine what these methods can add to examinations of media bias but also to add to our understanding of how news expresses opinion, prejudice, and factual information (i.e., media bias).

1.2.3 NLP and Media Bias

Computer science studies add to the study of media bias by applying methods of Natural Language Processing (or NLP) to the automatic detection of bias in texts. Often, these methods rely on some type of manual identification of bias as a first step in training their models; in bias studies these are usually perceptual ratings of bias at the word, sentence, and article level or a reliance on existing lexicons or feature data bases that have already been bias-labelled by raters. NLP studies of bias provide many useful results that demonstrate that bias can be predicted based on lexical features and some even demonstrate that bias can be predicted based on

grammatical features that raters perceive as biased. However, a major issue in these studies for linguistic researchers is that several of the language-based methods employed in these studies operate as black-boxes, and the specific lexical and grammatical features used as predictors of bias are much less transparent as a result. In other words, NLP researchers often refer to these features as consolidated groups without discussing which specific features or words fall into their categorizations, and even if they identify them at first, the features and their contribution to the output are completely lost once they are introduced to the computer model. In these studies, you are likely to see phrases like ‘complexity features’, ‘stance features’, or ‘biased words’ without any indication, except the occasional example, of what counts as a biased word or a complexity feature. This makes understanding how and why language is manipulated for reasons of bias difficult to understand from studies of this type.

Other types of NLP studies do not operate as black boxes, and it is possible to know which features are included in their tools, however, the goals of these studies are drastically different than my goal. In NLP studies, the goal is usually to perfect, test, or compare an automatic detection model of some kind and not to linguistically describe what is happening in texts that are biased. In other words, their goal is to use features and lexical items to quickly identify biased texts whereas my goal is to identify linguistic patterns that exist in biased and unbiased texts. Thus, examining the collective findings of NLP studies is unhelpful in trying to determine the state of the art of the linguistic patterns of a biased news article as they do not paint a comprehensive picture of a biased text— what their findings could tell us is the state of the art for the best automated detection tool and what tweaks to make to make it faster or more accurate. Findings from their studies are more helpful for my purposes taken piecemeal. Essentially, I’d like to empirically determine the linguistic patterns of biased news reporting, so I

will describe pieces of NLP findings that can shed light on the potential grammatical or lexical items that might be connected to bias below.

As was just stated, certain findings from these studies can still inform subsequent examinations of the language of bias, especially when examples of the features are provided. Lexical features that have been shown to predict bias in texts include lexicon that is connected to morality (Lin et al., 2018), words related to subjectivity (Mukherjee & Weikum, 2015; Pang & Lee, 2004; Recasens & Jurafsky, 2013), lexicon that is connected to negative sentiment (Recasens et al., 2013), words that are opinionated and convey emotion (Spinde et al., 2021), n-grams – particularly those dealing with divisive topics like "trump" and "liberals" (Rashkin et al., 2017), vocabulary richness (Sales et al., 2021), and polarized bigrams (e.g., climate crisis or affordable care) (Bandy, 2020). In terms of lexical features and examinations specific to biased news, lexicon-based NLP approaches have found that emotional, implicative, praising, and perspective specific language (especially negative perspectives) are better predictors of bias in news texts (Horne et al., 2018; Recasens & Jurafsky, 2013).

Grammatical features that have reliably predicted bias in NLP studies include assertive verbs (Recasens et al., 2013; Spinde et al., 2021), hedges (Recasens et al., 2013), factive verbs (Recasens et al., 2013), implicative verbs (Recasens et al., 2013), factuality features (e.g., word frequencies, ratios of quoted words, number of external links, stance markers, and average paragraph length) (Popat et al., 2018; Potthast et al., 2018), structure features (e.g., part of speech counts, pronouns, and prepositions) (Baly et al., 2018), and complexity features (e.g., lexical diversity, word and text lengths, and number of cognitive process words) (Horne et al., 2018). Horne and Adali (2017) demonstrate specifically that part-of-speech tagging, punctuation, stop-words, and measures of syntactic complexity reliably point out that fake or biased news tends to

be more simplistic and more repetitive than ‘real news’. Further, Rashkin et al. (2017) conclude that superlatives (e.g., most and worst), subjectives (e.g., terrible and brilliant), abstract generalities (e.g., truth and freedom) and the second person pronoun were all high indicators of news disinformation.

The important takeaway from NLP studies is that a combination of word choices and grammatical choices perceived as bias can be used to reliably predict bias in texts. However, there are three major limitations in NLP studies that need to be addressed. First, the specific features and words being attributed to bias need to be transparent or else they cannot be understood or examined further. Second, the specific words and features being attributed to bias are often isolated from their contexts and almost never analyzed back in their original contexts once they are identified which is problematic since language is highly context dependent (Spinde et al., 2020). And last, researchers in NLP do not generally control for other forms of linguistic variation – the most important of which is register variation – meaning that they cannot be confident that their identified features are really attributable to bias and not some other factor. Indeed, variation in linguistic structure has been shown most reliably to reflect differences based on context and communicative purpose, thus any considerations of linguistic structures need to control for register and examine the language in context (Biber & Conrad, 2019). However, methods of NLP that are reliable include their collection of perceptions to measure bias. Additionally, there is only one study to my knowledge that attempts to empirically identify the constructs of ‘bias’ which is an NLP meta-analysis conducted by Spinde et al. (2021). They identify 6 latent constructs of bias and demonstrate the reliability of each in rater perception. Based on their findings, and findings from perception studies in NLP, bias will be operationalized in this study by how raters perceive texts as bias along the two most reliable

constructs identified by Spinde et al. (2021) (i.e., extent and direction). This is discussed further in chapters 3 and 5.

1.2.4 Linguistics and Media Bias

There are very few linguistic studies done on misinformation, but the few there are yield very meaningful results for understanding bias as well as how it shapes our perceptions and understanding of news discourse. The majority of the linguistic studies that do exist examine presentations of social bias or political bias in political discourse instead of media bias. However, they will provide valuable information for understanding what is already known about the language of bias that can then be applied to the findings in this study.

Linguistic studies have examined how news can display social biases through their language choices by examining how labels and word choice is directly linked to public perception (Baker, et al; 2018). These studies tend to examine collocates for their social groups of interest (Brouwer et al., 2017; Kim, 2014) and demonstrate that a negative sentiment towards social groups in news is correlative to negative community sentiment. This suggests that language in the news can perpetuate bias by promoting ideas of separation, differentiation, and negativisation. Further studies that have relied on corpora to examine collocates and shed light on how word choice may lead to biased representations of social groups include studies conducted by: Al-Heijn (2015) who examined the semantic macrostructures associated with Muslim women in the British news source BBC; Godbole, Srinivasaiah, and Skiena (2007) who used a corpus of newspapers and blogs to assess subjectivity based on sentiment analysis concerning entities (e.g. *people*); and Baker et al. (2012) who analyzed the representation around the word ‘Muslim’ in the British press. All of these studies demonstrated that the social groups being examined were commonly being presented negatively. Therefore, like with NLP studies,

there is strong indication that negative sentiment, particularly of a group one finds opposing, is an indicator of potentially biased news. These findings, though focused on social biases, are imperative for the study of political bias in the news as the same ideas of alienation and negativisation can perpetuate separation amongst political ideologies.

Other studies of linguistics focus on the language of political bias (outside of news) with the intent of explaining how language choices are reflective of users (i.e., politicians and their agendas) and how they might impact consumers (i.e., how an audience might be manipulated via language choices). Examination into how language choice is manipulated for political effect has shown that lying is marked by an avoidance of factive verbs, that self-protection and projected group responsibility are marked by the use of pronouns “I” and “we”, and othering is marked by pronouns ‘they’ and ‘those’ (Wilson, 2003). Thus, we can see that language choice can be reflective of political motivations.

In a recent linguistic study, Grieve and Woodfield (2023) apply methods of linguistic variation to the study of Fake News. Specifically, they compile a corpus of Fake News and real news, written by the same author, and identify grammatical features that were key in each group. They then functionally analyze the grammatical differences and what they demonstrate about real and fake news. In Fake News, they identify a higher reliance on emphatics, downtoners, adjectives and features associated with interaction over information density (i.e., shorter words, present tense, first-person, second-person, and third-person pronouns, verbs, perfect aspect, etc.). Conversely, real news relied on many features of information density (e.g., nouns, noun-modifiers, nominalizations, longer word length, etc), and interestingly, features of stance that the authors conclude demonstrates conviction (e.g., suasive verbs, possibility modals, by-passives, and Public Verbs). Importantly, the authors identified that methods of linguistic variation,

particularly identifying key grammatical features and compiling comparative corpora that control for register, author, dialect, and political leaning can reliably identify functional variation in real and fake news. Like this study, the present study will also employ methods of register variation to examine media bias. Specifically, key features and keywords will be identified, and comparative (sub) corpora of biased groups will be compiled, and register, political leaning, topic, and time will be controlled for.

Collectively, the academic studies from social sciences, NLP, and linguistics demonstrate that language – both grammar and lexis – are at least tangentially connected to presentations of misinformation (bias and Fake News). From social science studies, we see that bias manifests in news by gatekeeping, coverage, and presentation bias and that coverage and presentation bias specifically, are presented through language. From NLP studies, we see that grammatical and lexical features have been used to automatically identify media bias and fake news. From linguistic studies we see that grammatical and lexical features are functionally related to political agendas, reflect social biases in news, and that grammatical features can functionally describe Fake News. What is missing is a generalizable profile of the language of media bias. In other words, is there a systematic occurrence of language patterns across biased news texts (that is not attributable to another factor like variation in register)? This study will attempt to answer this question by combining methods of rater perceptions and construct definitions from NLP, identifying key language features and corpus collection and control methods from register variation, and functional analysis of language patterns and how they reflect the three frameworks of media bias from a combination of register variation and social science methods. This will add to our overall understanding of expression of information and misinformation in news.

Another important matter to establish at the beginning of this dissertation is what is meant by ‘media bias’. The descriptions above are how media bias is examined academically in different fields of study, but a clear, precise and succinct definition of media bias is not something that exists. Bias by its nature is messy, multifaceted, and complicated – trying to neatly define and study it is very challenging. Because this paper is capturing perceptions of media bias on scales of political leaning and extent, the definition of ‘media bias’ will remain theoretical and will include anything that a person might reasonably consider as biased. To specify, the full range of considerations that raters might be paying attention to as bias could include: issues of credibility, fairness, impartiality, knowledge, accuracy, truthfulness, etc. In addition, they may be perceiving expressions of media bias ideologically, politically, socially, or culturally.

1.3 Goals of the Dissertation

The goal of this dissertation is to try and determine whether systematic patterns of perceived bias can be identified in newspaper texts. Bias in this study will be examined as two constructs based on findings from previous research: bias extent and bias direction (discussed in methods). This study relies on methods of register variation to examine the use of language in newspaper texts perceived as extreme in their bias and comparing them to those perceived as having no extent of bias. Additionally, comparisons of the language in texts perceived as conservative (a.k.a. right) to those perceived as liberal (a.k.a. left) are also examined to determine whether there are patterns that constitute ‘biased language’. In this way, the hope is to take news that is perceived as biased to different extents and directions and determine if any linguistic patterns exist.

To reiterate, the overarching goal is to attempt to identify, analyze, describe, and understand the linguistic variation of biased news and whether variation happens systematically across a cline of political leaning and extent. This goal is accomplished through six major steps: (1) compilation and description of a corpus of newspaper publications for which topic and time are closely controlled, (2) the collection of reader perceptions on the extent and direction of bias of each collected newspaper text, (3) a nuanced examination of key grammatical features and a comparison of their functions across extreme extent and no extent texts, (4) an analysis and comparison of the keywords that occur in both extreme extent and no extent texts (5) a comparison of key features across texts perceived as conservative to those perceived as liberal (6) and an analysis and comparison of keywords that occur in texts perceived as conservative to those perceived as liberal. The results of this study will help contribute to our knowledge of the linguistic nature of political bias in media, and hopefully, facilitate determinations about deceptive and impartial news practices.

1.4 Outline of the Dissertation

This study comprises 10 chapters. Chapters 2 and 3 are literature reviews divided between media bias studies and language studies respectively. Chapter 4 discusses the methods behind the compilation of the corpus in this study (*SPIN*), chapter 5 discusses the methods of measuring and collecting bias ratings, chapter 6 discusses the raters, and chapter 7 focuses on explaining how bias was operationalized with ratings (i.e., methods behind the division of texts into bias groups). Chapter 7 also overviews the characteristics of each bias group of interest (i.e., topic divisions, publication, and article types). Chapters 8 - 9 are a mix of methods, results, and analysis for the key feature and keyword analyses, respectively. Chapter 10 summarizes the results of the empirical analyses carried out in Chapters 8 and 9 and connects the results back to

the types of media bias identified in chapter 2. In addition, the limitations of the study and areas for future research are discussed.

Because this study draws from communication studies, NLP, and linguistics, the methods and the research design are atypical. To try and be as transparent as possible, a bulleted summary of each chapter and a flowchart of the study design are presented below.

- Chapter 2 – Literature review of media bias
- Chapter 3 – Literature review of language studies and media bias
- Chapter 4 – Methods and discussion of *SPIN* (the population of texts relied upon in this study)
- Chapter 5 – Methods of measuring bias (including the instrument used to collect perception)
- Chapter 6 – Overview of the raters and how they compare to average news readers (the population in this study)
- Chapter 7 – Overview of methods for creating bias groups & and overview of group characteristics
- Chapter 8 – Methods, results, and analysis of the key feature analysis
- Chapter 9 – Methods, results, and analysis of the keyword analysis
- Chapter 10 – Synthesis of findings, limitations, and conclusions

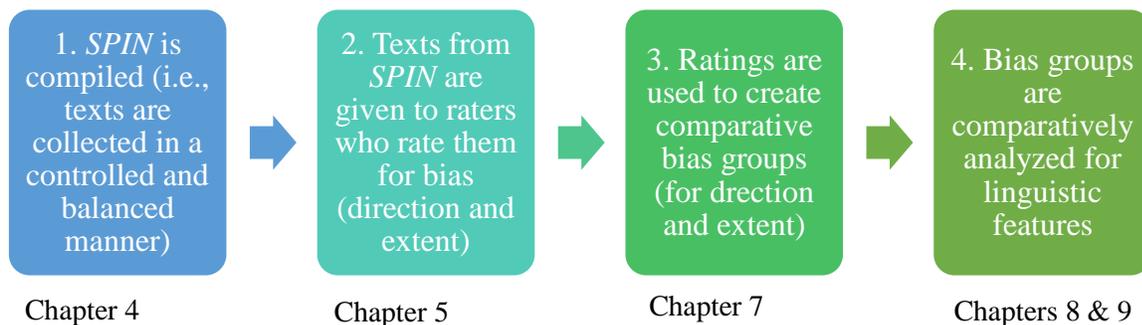


Figure 1.2 Schematic representation of study design for this dissertation

CHAPTER 2. LITERATURE REVIEW OF MEDIA BIAS

2.1 Introduction

This study will focus on identifying the linguistic variation of news bias; however, media bias is part of a much larger field of study known as misinformation. Therefore, the literature review will first provide background on misinformation and how media bias is specifically situated within the larger concept (section 2.2). Then issues in examinations of media bias will be overviewed before the definition of bias media bias and an overview of three common frameworks of examination are provided (section 2.3). Finally, an overview of previous research on media bias is presented.

2.2 Misinformation, Disinformation, Fake News, and Media Bias

As was shown in chapter 1, popular claims of media bias are becoming more prevalent in today's society. This is the case because social media has given rise to the dissemination of information at unparalleled speeds which has, unfortunately, also led to the proliferation of false information, information that has been altered to affect an individual's beliefs, or information presented to appeal to emotions over truth (Nakov & Da San Martino, 2021). According to research, news that is fabricated or altered (i.e., fake or biased news) has been shown to spread six times faster than 'real' news, and partisanship, the degree of novelty, and the emotional reactions of consumers are strong predictors of belief in deceptive news (Tangri, 2021; Uscinski et al., 2016; Vosoughi et al., 2018). This situation has created unprecedented challenges for traditional news media to remain relevant making the potential for biased reporting, especially to appeal to the partisanship or emotions of a reader, likely. The current Dominion lawsuit against *Fox News* for knowingly spreading false information about 2020 election fraud is the perfect

example of this situation occurring in the real world; all of this taken together renders examinations of misinformation pressing in academic disciplines.

The academic study of media bias and fake news fall under the larger academic domain of *misinformation* which presents itself in a multitude of ways in news media. The most important distinction is that between *misinformation*, *disinformation*, and *malinformation* of which Fake News and Media Bias are practical incidences. *Misinformation* includes any false or misleading information regardless of intent, *disinformation* refers only to information that is intentionally false, and *malinformation* is the sensitive information (true) that is strategically used to cause harm/disadvantage (Santos-D'amorim & Miranda, 2021; Stahl, 2006; Tandoc et al., 2018). To clarify, if a news outlet unknowingly or accidentally misleads an audience through biased reporting, this would fall under *misinformation*, but if a news outlet deliberately misleads or deceives an audience through biased reporting, this would fall under both *misinformation* and *disinformation*; if a news agency published a story they knew to be false, this would fall only under *disinformation*, and if a news outlet knowingly used sensitive information to harass, target, or spread hate speech, this would be *malinformation*. Due to the overlap that could potentially occur between these three distinctions, Wenzel (2019) distinguishes further between *disinformation* and *truthful disinformation*, the latter of which includes information that is true, but important information is purposely excluded to manipulate the reader, a.k.a. bias by selective reporting or bias by omission (these are more specifically defined in 2.3).

To help clarify, figure 2.1 below is a Venn diagram designed to illustrate how the three concepts of *misinformation*, *disinformation*, and *malinformation* are distinct and how they overlap. Additionally, the diagram includes examples of practical incidences of news and how they would fall into each category. For example, Fake News falls under *disinformation* while (media) bias

falls under both *misinformation* and the intersection of *misinformation* and *disinformation* (these incidences are circled in red on the diagram because they are of interest in the present analysis). The diagram also contains real and hypothetical examples based on the current Dominion lawsuit against *Fox News* to illustrate how some of these concepts might play out in the real-world news cycle.

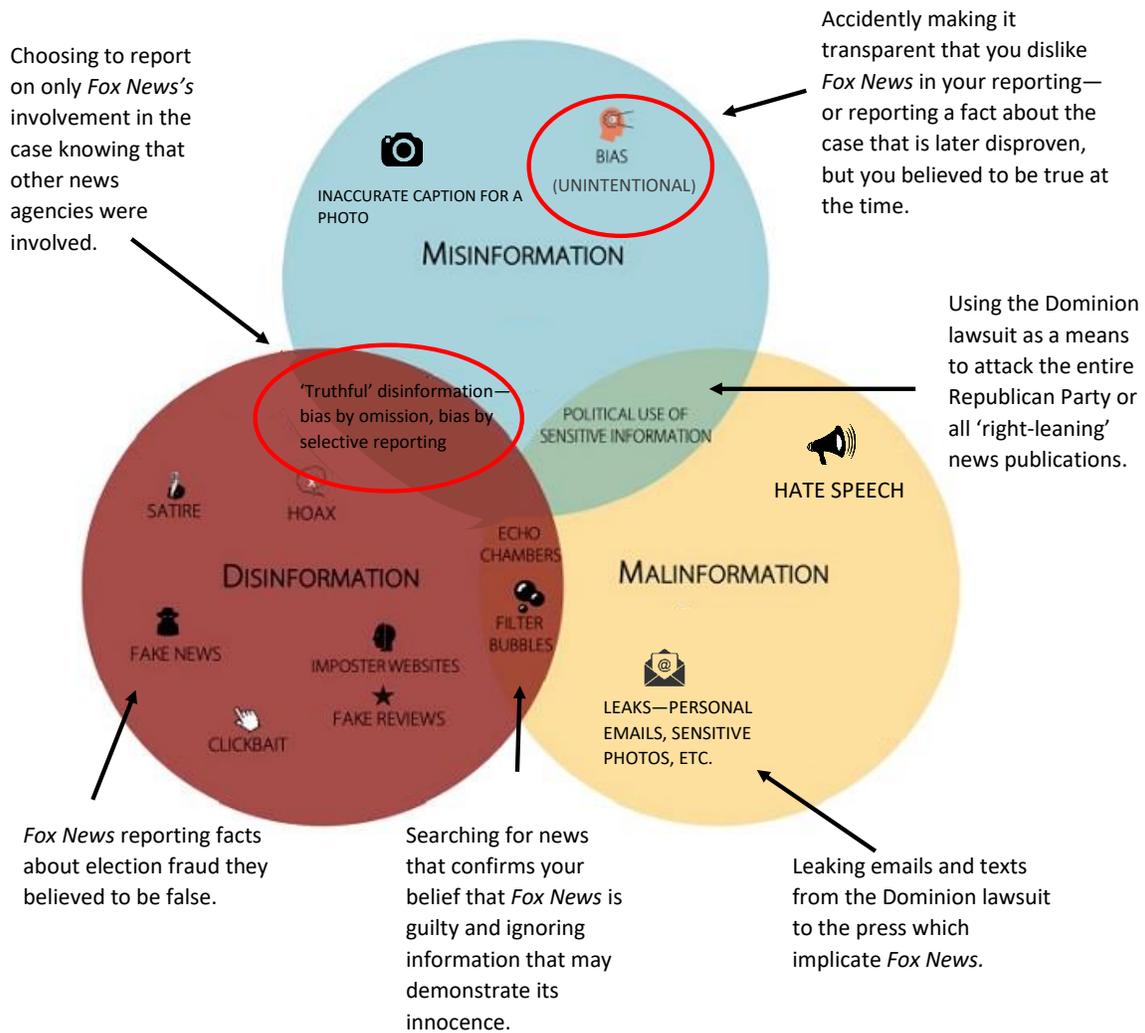


Figure 2.1 Venn diagram of misinformation, disinformation, and malinformation (adapted from Santos-D'amorim & Miranda, 2021)

The distinction between these three is an important one in academic research as all three are closely related and refer to news that is suspected to be inaccurate, one-sided, misleading, or fabricated with the distinction between the three coming down to intent (Grieve & Woodfield, 2023); *disinformation* is the presentation of untrue information with the intent to deceive, *misinformation* is the presentation of untrue or misrepresented information without the explicit intent to deceive, and *malinformation* is the presentation of true information with the intent to harm. The most important facets of misinformation for the present study are Fake News and media bias which are, in theory, distinct with Fake News being news which is intentionally untrue, and media bias being news that offers a skewed perspective (Davis, 1990; Santos-D'Amorim & Miranda, 2021). However, this demarcation is unlikely to remain this clear cut in actual instances of news as a story that is mostly fabricated is likely to contain some truthful information, and a story that is 'real news' could easily contain false information (Santos-D'Amorim & Miranda, 2021). Further, it is difficult to always know an author's intent, therefore, the majority of academic papers on fake and biased news are likely examining news that actually falls somewhere in the intersection between *misinformation* and *disinformation*. For this reason, and because the present study will examine professionally published news articles and have no way of collecting author intent, existing empirical research on disinformation (i.e., Fake News) and misinformation (i.e., media bias) will be used to inform this study. This decision was further made as it is likely that the communicative purposes (e.g., persuasion, deception, emotional appeal, etc.) of biased and Fake News are similar and might, therefore, reveal similar patterns of linguistic variation. It is, of course, expected that if Fake News contains linguistic features functionally similar to biased news, the feature of Fake News will likely happen at a more extreme extent since reputable journalists strive to appear impartial and use language that does

not reveal their opinions directly (Baumer et al., 2015). Nonetheless, it is the belief that studies examining the language of Fake News will be just as informative as studies examining the language of bias, so a presentation of these studies will be included.

2.3 Media Bias

Before presenting studies that examine the language of media bias and Fake News, an overview of studies specific to media bias will be presented. Though these studies do not examine language specifically, they provide the theoretical underpinning for most empirical examinations of media bias as well as a viable framework for identifying how media bias manifests in news. Indeed, I believe that combining a media bias framework with functional theories of language use, particularly those related to register variation (Biber & Conrad, 2019) is crucial for understanding the language of biased news. So, an overview of these studies will be presented below, but first, major issues in the study of media bias will be discussed as these are essential for understanding the definitions of media bias that are widely applied in academic research on media bias (2.4). The chapter will end with a presentation of previous research on media bias done under the framework of gatekeeping, coverage, and presentation bias.

2.3.1 Problems with Defining Media Bias

Media bias is generally considered difficult to analyze because it is messy, complex, and multi-faceted making research on the matter complicated. As was demonstrated in section 2.2, the demarcation between media bias and other types of falsified news is murky at best rendering it difficult to clearly document media bias and to neatly ascribe motivation when/if it is detected. Thus, the biggest problem with media bias is that it is difficult to succinctly define especially in a way that can be meaningfully applied to identify incidences of bias in news. Media bias is

defined by some researchers as “slanted news coverage or internal bias, reflected in news” (Hamborg et al., 2018) but trying to apply this definition as a means of empirically capturing bias in news is likely to yield messy results. Indeed, the literature demonstrates that researchers are employing this definition to incorporate credibility (Choi et al., 2006; Kioussis, 2001), fairness or impartiality (Lee, 2005; Niven, 2002, 2004; Vallone, 1985; Xu, 2022) or truthfulness (Paul & Elder, 2006; Raymond & Taylor, 2021) rendering a consensus in the research very difficult as these are arguably connected, yet separate, senses of the term. In addition, media bias can be demonstrated in a multitude of ways including ideologically (i.e., in favor of or against capitalism, consumerism, patriarchy, etc.), politically (i.e., in favor of or against positions based on a left-right political spectrum), and socially or culturally. It further contains layers of intent and extent. Bias can be intentional and pervasive as well as accidental and infrequent and any mix in-between (i.e., disinformation, misinformation or truthful disinformation).

As with studies of misinformation and Fake News, intent is heavily debated in the media bias literature as a requirement for media to be considered biased; specifically, many contend that it must be intentional and systematic (Williams, 1975). While systematicity is an important quality of media bias, for what harm does an isolated incident really inflict, intentionality is not a required quality of media bias in the mind of the researcher. In fact, one could argue that it is the implicit, systematic yet unintentional bias that is more pervasive in the media as obvious bias is easily recognizable and thus avoided. The bias that is disguised as impartial fact is more believable and thus arguably more harmful. In addition, as was stated in section 2.2, it would be difficult to truly know an author’s intent if bias is disguised as impartial fact.

The next issue in examinations of media bias is the market forces that can drive bias. The economic drivers of media bias, first identified Gentzkow and Shapiro (2004; 2006), have

garnered much attention recently as six corporations control 90% of the media in the US; thus it is important to consider the role that economic drivers may play in media bias (Lutz, 2014). On the one hand are supply-driven motivations which include bias introduced by media firms to influence consumers to take particular action (Gentzkow et al., 2014). Incentives may arise from pressure by governments and advertisers or may be internal to the firm and include preferences of media owners, reporters or editors. On the other hand, are demand-driven motivations which arise from a tendency of consumers to choose media whose biases match their own prior beliefs. In other words, they search for confirmatory news which leads to firms presenting information in a way they think might appease their audience – for example, *Fox's* presentation of election fraud as factual. These determinants of bias are considered in the discussion portion of this paper as they are likely attributable to some of the identified linguistic characteristics of bias.

In addition to media bias being difficult to define, it also includes an underlying assumption that there is an objective baseline. When claims of media bias are being made, we must ask bias relative to what? As Gans (1979) suggests, media bias can be measured as relative to what a “neutral” source would do, as relative to the average citizen’s, voter’s, or target audience member’s preference, or relative to other media outlets. This necessitates that studies of media bias be comparative in nature and establish some baseline for comparison. To illustrate, picture a comparison of an episode of Tucker Carlson to a newspaper article by the *New York Times*. This comparison is likely to skew the perception of bias very heavily in one direction and may even suggest that the *NYT* article is unbiased. However, a comparison of that same *NYT* article to one published by *Associated Press* may yield very different perceptions – and it may not.

The next major issue in examinations of media bias relates to practical limitations which can range from constraints of space, time, and telling a coherent narrative to issues of censorship and government control. Most of these issues can be managed by examining news sources that are practically limited in the same way. For example, a major part of the media bias literature is concerned with degrees of bias in countries like Myanmar or China (MacKinnon, 2009; Simons & Strovsky, 2006; Yin, 2008), but media in these countries are constrained by government censorship and thus cannot fairly be compared to countries like the US in which media bias is viewed negatively as it “contravenes the standards of professional journalism” (Eberl et al., 2017). So, controlling for country in an examination would help with this. Furthermore, comparing news of the same medium (e.g., only print to print, online to online, or TV news to TV news) would control for constraints of space and time since practical limitations are equal.

2.3.2 Frameworks for Identifying Media Bias

Given the issues in defining media bias presented above and all the meanings under which media bias is examined, D’Alessio & Alan (2000) undertook a seminal meta-analysis which overviewed the many frameworks for capturing media bias and reorganized them into three common, succinct, and easily replicable frameworks for identifying media bias. Thus, they take the many references for the physical manifestations of bias, or the forms of bias, including: spin, ideology, framing, priming, agenda-setting, persuasion, gatekeeping, content, distortion, decision-making, label, coverage, presentation, partisan, and statement (see DeMarzo et al., 2003; Hamborg et al., 2018; Hofstetter & Buss, 1978; Hofstetter, 1976; Mullainathan & Shleifer, 2002; Stempel, 1969) and combine all of these into just three succinct frameworks of bias: (1) *gatekeeping bias* for which decision-making, selection, and agenda-setting are common synonymous and refer to stories being selected or ignored on ideological grounds; (2) *coverage*

bias, for which ideology, content, bias by omission, and framing are synonymous and refer to issues or actors being more or less visible in the news; and (3) *presentation bias*, a.k.a. distortion, persuasion, label, statement, priming, and spin, which all refer to coverage being slanted towards particular actors or issues with favorable news production choices. For the purposes of clarity, *gatekeeping*, *coverage*, and *presentation bias* will be the terms used here going forward. It is easy to differentiate between the three types of bias by remembering that gatekeeping is *what* a news publication discusses, coverage is *how much* they discuss a particular topic, and presentation is *how* a topic is discussed (e.g., positively, negatively, etc.). It is also important to note that almost all studies examining the language of media bias fall under the umbrella of presentation bias and some under the umbrella of coverage bias (discussed further below).

In an effort to more clearly illustrate media bias and how these various definitions are connected, a chart from Hamborg, Donnay, and Gipp (2019) has been adapted and is presented below (see figure 2.2).

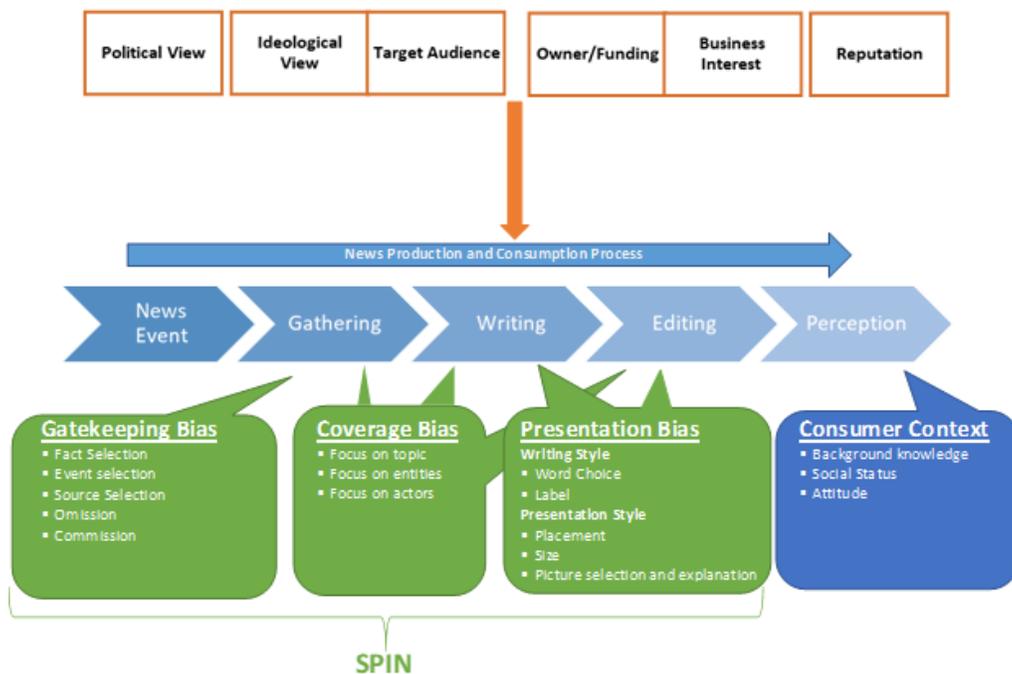


Figure 2.2 *Illustrated representation of media bias (adapted from Hamborg, Donnay, and Gipp, 2019).*

The chart depicts the external economic influences on bias at the top (in orange) which can arguably affect the news production at any point. The light blue portion of the chart depicts the production stages of news, and the green illustrates where each type of bias can enter into the news writing process with the three green boxes collectively equating the spin that is present in the final product. The dark blue box represents what consumers bring to the news story.

As the chart demonstrates, gatekeeping bias can manifest in event selection, source selection, fact selection, complete omission, accepting commission for a piece, or a combination of these. Coverage bias manifests once a story topic is selected in how much focus is on perspectives of the topic, actors, or entities relevant to the topic. Presentation bias manifests as a direct result of choices made during writing and editing and can include word choice, labels, lexico-grammatical choices, and choices made in terms of placement of information, photos, captions, pull-quotes, etc. Consumer context is what the reader brings to the table that may affect their perception of story from previous knowledge to personal biases.

Due to the difficult nature of defining bias, and based on figure 2.2, it was determined that a bottom-up examination of news bias was best. In other words, rather than defining bias or beginning with one of the three frameworks and conducting a top-down examination of news texts, this study will begin with consumer perceptions (the blue box in figure 2.2) of bias in news texts (or the collective spin). Consumers will rate news articles for bias, and linguistic variation across articles marked as biased will be identified and then analyzed to determine how language patterns can reflect the potential biases from production choices of the authors and publications (green boxes in figure 2.2). To clarify, the identified linguistic variation will be analyzed to determine if it reflects presentation bias or coverage bias, and the collection of texts (discussed

below) will incorporate a method of controlling for gatekeeping bias. In addition, the language patterns of articles marked as biased will be analyzed for biases that may emerge from external influences (orange boxes in figure 2.2). Each of these steps will be motivated later in this dissertation, and a clear operational definition of bias in the present study is presented in chapter 5.

2.4 Empirical Research on Media Bias

A good majority of research examining misinformation and media bias come from fields of social science particularly those in fields of communication, politics, and psychology and some from computer science. These fields tend to be concerned with different facets of media bias. For example, psychology is mostly concerned with human perception of bias and identifying personality characteristics that predict belief in falsified news (i.e., how cognitive dissonance and confirmation bias relate to one-sided information seeking). Political and communication studies are concerned with identifying and characterizing media bias, and computer science is concerned with automatic identification of media bias and Fake News. This section will focus on presenting a review of studies from social sciences that conduct examinations with the three frameworks identified by D'Alessio & Allen, 2000: 1) gatekeeping bias or which news stories make it to print versus those that do not, 2) coverage bias or how much time and space is devoted to one issue or actor over another, and 3) presentation bias or how an issue or actor is presented in the news story. Chapter 3 will present studies from computer science (NLP).

2.4.1 Studies of Gatekeeping Bias

Gatekeeping bias, also known as selective reporting, can include event selection, source selection, fact selection, commission, and omission or any combination thereof. To empirically investigate gatekeeping bias, a baseline of newsworthy events and some type of quantifiable measure of news need to be established. The baseline can be anything from the occurrence of a major event (either localized, national, or international depending on the news being examined) or records of facts such as from transcripts, police reports, or court records. Quantifiable measures typically include frequency counts of stories, phrases, citations, names, etc., Two types of bias tend to manifest themselves in news based on findings in studies of gatekeeping bias: social bias and political bias, discussed respectively below.

Several studies examining selective reporting suggest that news sources tend to favor certain ethnic, gender, and socio-economic groups over others in the stories that make it to print versus those that do not (Ditton and Duffy, 1983; Hackett, 1984; Schudson, 1978). Lundman (2003) concluded that ‘newsworthiness’ is heavily affected by cultural typification based on race and gender in his study examining news coverage of homicides in Columbus, OH. The author specifically found that homicides received more media attention if the victim was White-female and the offender Black-male even though this type of homicide was uncommon. This conclusion has subsequently been confirmed by several studies examining gatekeeping bias. Gruenwald et al. (2009) examined the reporting of homicides in Newark, New Jersey by using police reports of homicides as a baseline and measuring reports of the homicides in the major daily local newspaper; both binary measures (present vs not present) and frequency of occurrence (articles per homicide) were collected. The study found that 65% of homicides that occurred were reported in the news, and that even though Black-on-Black homicide was far more common, an overrepresentation of homicide occurrences involving Hispanic offenders and White, Black, and

Hispanic victims existed (Hispanics being the dominant minority population). This study supports previous findings that news decision-makers typify offenders and victims on the intersections of race and gender (see Barak, Leighton, & Flavin, 2007; Daly, 1997). This type of bias displayed by news is arguably detrimental as the news is likely the only way the public is informed about such events.

Political bias, or partisan bias, is also a common type of bias examined in gatekeeping research. Often events such as demonstrations, public speeches, political campaigns, etc. are used as a baseline, and frequency counts (of news articles, reports, phrases, etc.) are used to obtain a quantifiable measurement of whether the event is being fairly covered. In addition to determining whether political events are gatekept, several studies have also tried to determine which factors are related to gatekeeping and have found that the number of participants of an event, relevance of the event issue in the current news cycle, and previous media attention have a high impact on media coverage (see D'Alessio & Allen, 2000; McCarthy et al, 1996; Oliver & Maney, 2000). For example, D'Alessio and Allen (2000) conducted a meta-analysis of 59 studies on partisan media bias in presidential election campaigns and found virtually zero partisan bias in news (save for a small effect in TV news) but did find that demonstration size and popularity in the news cycle were the main factors influencing whether the event was covered.

Examinations related to political bias also examine how often citations, name references, or even phrases are attributable to political organizations, think tanks, or politicians as a means of examining gatekeeping bias. For example, Kenney and Simpson (1993) demonstrate that *The Washington Times* displayed enormous gatekeeping bias in the 1988 presidential campaign towards Republicans. Conversely, Groseclose & Milyo (2005) determined that there was a strong liberal bias in all news outlets examined (except *Fox* and *The Washington Times*) by comparing

the frequency of citations from media outlets to think tanks and policy groups, then comparing this with the frequency that members of Congress cite the same groups. Another study conducted by Gentzkow & Shapiro (2010) examines gatekeeping bias by comparing the use of significant phrases uttered by congressional Republicans and Democrats and analyzing the frequency of those phrases in news sources to estimate the outlet's bias toward one side of the political spectrum.

Gatekeeping bias is important to understand as the news holds the responsibility of disseminating information to the public. At some level in news production, selection and filtering are a requirement as the set of world events is too large to be treated exhaustively, but when it happens systematically along political or social clines, it can have negative consequences.

2.4.2 Studies on Coverage Bias

Coverage bias differs from gatekeeping bias in that it does not focus on whether a story is included in the news but how much it is covered once it is. Specifically, examinations of coverage bias are concerned with the amount topics, entities, and people are focused on in news stories. Empirical investigations of coverage bias are generally concerned with issues that have only two reasonable perspectives as a 50-50 split is then a fair expectation in news coverage; in these investigations, frequency counts and comparisons, size and placement of a story, known as volumetric measurements, or amount of time devoted to an issue are relied upon as quantitative measures (D'Alessio & Allen, 2000).

As with gatekeeping bias, events that are political in nature are often examined in coverage bias since researchers can track how much coverage is devoted to one political party versus the other. Blumberg (1954), for example, found that two major newspapers devoted large

portions of their 1952 campaign coverage to one party (*Indianapolis Star* devoted 77% to Republicans and *The Great Falls Tribune* devoted 72% to Democrats). Druckman & Parkin (2005) examined the amount of space devoted to each candidate of the 2000 Minnesota senate race in the *Star Tribune* and *Pioneer Press* and found that each newspaper devoted equal space to each candidate and most notably, the two papers exhibited extremely similar patterns in their coverage with a majority going to both candidates simultaneously. Entman (2010) examined coverage of Democrats and Republicans after each respective National Convention. He found that *ABC* and *NBC* devoted more paragraphs to negative assessments of Democrats and more positive paragraphs to Republicans. More recently, Kweon (2017) found that *Fox* disproportionately reported favorably on Donald Trump while *CNBC* disproportionately focused on Clinton in the 2016 election.

In addition, many studies of coverage bias will examine how production mechanisms like where articles are placed (Holsanova, Rahm & Holmqvist, 2006), size (McCombs & Shaw, 1972) and headlines (Stempel, 1965) contribute to displays of bias. It is argued that if a news outlet wishes to diminish a story, they can deemphasize it with placement and amount of coverage and by the same means can also amplify a story (Shultziner & Stukalin, 2020). To exemplify, Niven (2001) examined the coverage of Bush and Clinton with regards to unemployment by examining the number of articles published, length of articles, placement of articles, and whether the president was mentioned prominently in the article. The author found no evidence of partisan bias in coverage but did find that the media covers bad outcomes far more than good outcomes. Shultziner & Stukalin, however, found that front-page headlines, the sizing of articles, and accompanying visuals, were “consistently and strategically” employed to align with an outlet’s ideological slant (2020, p. 1).

Studies of coverage bias also examine how specific issues are covered in the news. For example, Smith et al. (2001) examined whether the news coverage of demonstrations reflected the positions of the protestors or the government; based on the amount of coverage devoted to each issue, they found that news tended to favor the perspective of the government. Though not solely a study of coverage bias, Kull et al. (2003) examined how misperceptions about the Iraq War might be related to the presentation of information on US television news and found that more than 40% of Fox News viewers had misperceptions about Iraq War and that self-reported “greater attention” to news increased the likelihood of misperception.

Another study examined coverage bias more generally by comparing how *The New York Times*, *The Washington Post*, and *The Chicago Tribune* reported on the Intifada. This study, conducted by Zelizer, Park and Gudelunas (2002), compared article placement and type, headlines, photographs, graphics, sourcing practices and lead paragraphs to determine how the Intifada was represented in the three newspapers and whether any bias could be identified. Conclusions of this study are important as they identified that bias was shown in an erratic fashion which convoluted the comparison of bias in the three newspapers. Occasionally, all three newspapers showed a pro-Israeli slant and on other occasions this slant disappeared. Overall, they found that *The New York Times* showed a disproportionate amount of pro-Israeli sentiment as compared to the other two newspapers. The authors also concluded that no feature of American print news was free of values and that the study of bias requires development. To exemplify this point, the authors state:

When considering the fact that the uneven appearance of bias here seemed to have no consistent shape across the three newspapers, our analysis underscores how pervasive and embedded [bias] actually is.... Bias study has been limited to certain features of news

presentation, topics of reportage and circumstances for reporting, suggesting that the academy has not yet provided a picture of how bias spans the world of journalistic practice (p 303).

In the present study, coverage bias will be examined via keyword analysis which will highlight whether preselected topics receive equal coverage as a keyword analysis should highlight whether a topic, entity, or source is receiving significantly higher amounts of attention than another. In addition, because the texts will be comparatively examined across bias groups of direction and extent, the keyword analysis will indicate when an entity or topic is focused on for political reasons.

2.4.3 Studies of Presentation Bias

Presentation bias, also commonly called statement bias, is concerned with how something is reported on. It is argued that language is powerful as it can appeal to emotions and ideologies, and therefore, the majority of studies done on presentation bias include some type of examination into how language can demonstrate sentiment (be it positive, negative, opinionated, or emotive). For the most part, studies look at word choice/labeling or ‘language features’ and how they produce a sentiment or affect that displays bias in the news. In the social sciences, analysis requires time-consuming manual coding that is usually topic-oriented or person-oriented and in computer science, automatic detection models are relied upon to extract sentiment or detect bias (demonstrated more clearly in the discussions of NLP and CDA studies below).

The effects of presentation bias can range from concept-level (i.e., how a specific person or concept is presented) to publication-level (i.e., the overall presentation by one publication or firm) to media-level (i.e., how an event is being presented in news overall). Examples of presentation bias on the publication-level include Papacharissi and Oliveira (2008) who

investigated labeling and word choice in media coverage of terrorism in US and UK publications. Overall, they found that US publications focused more on military approaches to terrorism while UK papers focused on diplomatic evaluations of terrorist attacks. Importantly, they noted that *The New York Times* tended to dramatically dehumanize terrorists while *The Washington Post* employed a less dramatic tone, and that *The Guardian* and *Financial Times* focused more on reporting facts. Dooley and Corman (2002) examined temporal patterns in *Reuters* coverage of the September 11 attacks on the world Trade Centers. They identified a reporting pattern they deemed “stimulus-response” in which the triggering event causes an institutional response and is followed by a consideration of the societal impact. Their analysis also demonstrated that for *Reuters* reports of 9/11 the attack and the response were discursively and temporally distinct.

From studies conducted under the three categorizations of media bias, we can see that media bias exists and that it has real world consequences. However, these studies share the common limitations of looking at very specific, and usually politically charged, instances of biased news and then overgeneralizing their results back to all news media. In addition, these charges of bias often relate to a political affiliation, but very little control (other than examining time and space devoted to political candidates of opposing affiliations) is put in place to determine whether these are actual reflections of political bias.

In addition, two of the aforementioned types of media bias, coverage bias and presentation bias, directly involve examinations of language, yet very few of these studies draw on previous linguistic research to conduct their examinations of language. This is limiting because linguistic methodologies, particularly corpus methods and methods of register variation, could be very informative for examinations into these types of bias coverage. For example, a

corpus built to represent news of a wide variety of topics and across multiple publications could provide more generalizable results for gatekeeping bias while examinations of key linguistic features and keywords would statistically highlight whether issues or actors are focused on more than any other and would also demonstrate how entities and actors are presented in news articles. In addition, proper control for topic, political leaning, and register would allow for identifications in linguistic variation to be more reliably attributable to political bias. Thus, the present study aims to combine linguistic methodologies with previous characterizations of media bias to determine what these methods can add to examinations of media bias but also to add to our understanding of how news expresses opinion, prejudice, and factual information (i.e., media bias).

2.5 Conclusion

This chapter has presented an overview of misinformation and how media bias fits within the larger construct and has presented the major issues in studies of media bias. This chapter has also reviewed the three most common frameworks of media bias in communication studies and what research under these frameworks has added to the academic study of media bias before ending with how linguistic methods can make add to the studies of media bias. The next chapter will overview studies on the language of bias, media bias, and fake news.

CHAPTER 3. LITERATURE REVIEW OF LANGUAGE AND BIAS

3.1 Introduction

Because this study is ultimately interested in what we can learn about bias in news by examining linguistic variation, the purpose of this chapter will be to overview empirical investigations of the language of bias, media bias, and fake news. Section 3.2 will overview NLP studies and what they found about language and news bias while 3.3 will overview linguistic studies of bias, media bias, and fake news before discussing details of the present study.

3.2 NLP and Media Bias

Computer science studies add to the study of media bias by applying methods of Natural Language Processing (or NLP) to the automatic detection of bias in texts. Often, these methods rely on some type of manual identification of bias as a first step in training their models; in bias studies these are usually perceptual ratings of bias at the word, sentence, and article level or a reliance on existing lexicons or feature data bases that have already been bias-labelled by raters. NLP studies of bias provide many useful results that demonstrate that bias can be predicted based on a combination of lexical and grammatical features that raters perceive as biased. However, a major issue in these studies for linguistic researchers is that several of the language-based methods employed in these studies operate as black-boxes, and the specific lexical and grammatical features used as predictors of bias are much less transparent as a result. In other words, NLP researchers often refer to these features as consolidated groups without discussing which specific features or words fall into their categorizations, and even if they identify them at first, the features and their contribution to the output are completely lost once they are introduced to the computer model. In these studies, you are likely to see phrases like ‘complexity features’, ‘stance features’, or ‘biased words’ without any indication, except the occasional example, of

what counts as a biased word or a complexity feature. This makes understanding how and why language is manipulated for reasons of bias difficult to understand from studies of this type.

Nevertheless, findings from these studies can still inform subsequent examinations of the language of bias, especially when examples of the features are provided. Lexical features that have been shown to predict bias or falsities in texts include lexicon that is connected to morality (Lin et al., 2018), words related to subjectivity (Pang & Lee, 2004; Mukherjee & Weikum, 2015; Recasens & Jurafsky, 2013), lexicon that is connected to negative sentiment (Recasens et al., 2013), words that are opinionated and convey emotion (Spinde et al., 2021), n-grams – particularly those dealing with divisive topics like "trump" and "liberals" (Rashkin et al., 2017), vocabulary richness (Saleh et al., 2019), and polarized bigrams (e.g., climate crisis or affordable care) (Bandy, 2020). In terms of lexical features and examinations specific to biased news, lexicon-based NLP approaches have found that emotional, implicative, praising, and perspective specific language are better predictors of bias in news texts (Horne et al., 2018; Recasens & Jurafsky, 2013).

Grammatical features that have reliably predicted bias or fake news in NLP studies include assertive verbs (Recasens et al., 2013; Spinde et al., 2021), hedges (Recasens et al., 2013), factive verbs (Recasens et al., 2013), implicative verbs (Recasens et al., 2013), factuality features (e.g., word frequencies, ratios of quoted words, number of external links, stance markers, and average paragraph length) (Popat et al., 2018; Potthast et al., 2018), structure features (e.g., part of speech counts, pronouns, and prepositions) (Baly et al., 2018), and complexity features (e.g., lexical diversity, word and text lengths, and number of cognitive process words) (Horne et al., 2018). Horne and Adali (2017) demonstrate specifically that part-of-speech tagging, punctuation, stop-words, and measures of syntactic complexity reliably point

out that fake or biased news tends to be more simplistic and more repetitive than ‘real news’. Further, Rashkin et al. (2017) conclude that superlatives (e.g., most and worst), subjectives (e.g., terrible and brilliant), abstract generalities (e.g., truth and freedom) and the second person pronoun were all high indicators of news disinformation.

The specific methods used in NLP studies are not always the same. In fact there are several computer models and methods of training that are employed in the automatic detection of bias or fake news in NLP. For example, Recasens et al. (2013) examined how a computational method known as ‘bag-of-words’(i.e., a set of words pre-coded as ‘biased’, ‘neutral’, ‘democrat’, or ‘republican’) in conjunction with indicator features for bias (i.e., hedges, factive verbs, and part of speech features) can be used to accurately detect partisan bias at the sentence level. They specifically identify two types of bias: framing bias (subjective intensifiers and one-sided terms) and epistemological bias (factive verbs, assertive verbs, hedges, and entailment) and determine that subjectivity is strongly linked to features of framing bias. Other NLP studies have demonstrated that features of political bias related to semantico-syntactic structure (i.e.- a combination of certain words and grammatical features) can be used to accurately classify documents about controversial topics like ObamaCare (Conrad et al., 2012) or the Israeli-Palestinian conflict (Greene & Resnik, 2009; Lin et al., 2006).

Most recently, researchers in NLP have been attempting to circumvent the issue of a ‘gold-standard corpora’ for biased language detection by pre-training transformer models in some type of media bias domain (Krieger et al., 2022; Spinde, 2021; Spinde et al., 2020). These media bias domains are typically collections of sentences, words, and/or grammatical features that have been annotated for bias by previous researchers, experts, or by crowdsourcing public perceptions. While several NLP studies have shown that their automatic detection programs

improve with pre-training on better controlled media bias domains, there remains the issue of language being highly context dependent. To combat this issue, several researchers in NLP have attempted to specify their context by examining specific newspapers or specific contexts. For example, Brady and Ma (2003) examined how the ten most liberal and ten most conservative senators were labeled in *The New York Times* and *The Washington Post*. They found that ideological labels were more frequently attributed to conservatives indicating a greater tendency to identify conservative ideology. Spinde et al. (2020) attempt to capture political bias by training word embeddings on a corpus of articles from *HuffPost* (liberal) and *Breitbart* (conservative), and Lazaridou et al. (2017) automatically extracted quotes to demonstrate how quoting patterns can reflect political preference.

In addition, scholars in NLP have used features related to stance to detect and classify political bias. Iyyer et al., (2014) used recursive neural networks, a machine learning model designed to capture semantic and syntactic composition of sentences, to automatically detect partisan bias in a corpus of ideological books and congressional debates based on features of stance. Results show that their approach detected bias more accurately than existing methods (i.e.-framing bias and bag-of-word methods) on both datasets. Yano et al., (2010) examined the linguistic indicators of bias in American political blogs by identifying strongly associated bigrams in liberal and conservative sub-corpora and measuring association using log-likelihood. A bigram was identified as “liberal” or “conservative” if it was among the top 1,000 bigrams from its respective liberal or conservative blog and was not also among the top 1,000 bigrams on the opposite political side. Results showed that readers were able to perceive bias and that certain lexical cues remain consistent in the labeling of political direction.

The important takeaway from NLP studies is that a combination of word choices and grammatical choices perceived as bias can be used to reliably predict bias in texts. However, there are three major limitations in NLP studies that need to be addressed. First, the specific features and words being attributed to bias need to be transparent or else they cannot be understood or examined further. Second, the specific words and features being attributed to bias are often isolated from their contexts and almost never analyzed back in their original contexts once they are identified which is problematic since language is highly context dependent (Spinde et al., 2020). And last, researchers in NLP do not generally control for other forms of linguistic variation – the most important of which is register variation – meaning that they cannot be confident that their identified features are really attributable to bias and not some other factor. Indeed, variation in linguistic structure has been shown most reliably to reflect differences based on context and communicative purpose, thus any considerations of linguistic structures should control for register and examine the language in context (Biber & Conrad, 2019).

One strength that comes from NLP studies of media bias is that they demonstrate that relying on perceptions of bias is a viable way of identifying instances of bias in texts (Baumer et al., 2015; Hamborg et al., 2019; Lim et al., 2018; Lim et al., 2020). Specifically, they've shown high rates of agreement of labeling bias at the word, sentence, and text level with the highest agreement being on bias at the text level (Baumer et al., 2015; Hamborg et al., 2019). They have further shown that the perceptions of bias are particularly reliable when crowdsourcing is used (Mitrović, 2013; Mladenović et al., 2016). Further, a meta-analysis conducted by Spinde et al. (2021) which collected 824 perceptual items to reveal latent constructs of 'bias' showed that political leaning and strength of bias were the two most reliable and agreed upon constructs. To the best of my knowledge, no other study has attempted to separate bias into multiple constructs

even though there is a consensus that bias is very abstract and might mean credibility, truthfulness, etc. So, this dissertation will inform its operationalization of bias based on these findings from NLP studies that show that crowdsourcing perceptions of bias at the text level and inquiring about the constructs of ‘political leaning’ and ‘strength’ are reliable ways of measuring bias. To clarify, crowdsourcing will be used to collect bias ratings of news articles and ratings of bias will be collected for ‘political leaning’, here called direction, and ‘strength of bias’, here called extent. This will be discussed further in chapter 5.

3.3 Linguistics and Media Bias

There are very few linguistic studies done on media bias. The majority of the ones that do exist examine presentations of social bias or newsworthiness instead of political bias. Nonetheless, these will be overviewed here along with linguistic studies that are focused on the language of political discourse outside of news (i.e., the language of politicians). Though these are not directly related to the topic of interest in the present study, they are still considered important, as they will provide valuable information for understanding what is already known about the language of bias that can then be applied to the findings in this study.

3.3.1 Linguistic Studies of Social Bias in News

Because *labels* for social groups are fairly easy to identify and compare, a good majority of studies examine the presentation of social bias through language. Studies done on social bias have shown that the choice of words in the news is directly linked to public perception and that

language in the news can perpetuate bias by promoting ideas of separation, differentiation, and negativisation (Baker, et al; 2018). For example, Kim (2014) examined US news discourses about North Korea and determined that negative sentiment was salient within most publications, and Brouwer et al. (2017) revealed an unfavorable bias towards unauthorized migrants in Netherland news sources; both negative perceptions were also reflected in the respective social communities. Thus, as we saw with NLP studies of bias and Fake News, words expressing negative sentiment about an idea, topic, person, or group of people are associated heavily with linguistic displays of bias.

This is further supported by an examination of social bias in the British press, in which a corpus of texts was compiled to examine representation of refugees, asylum seekers, immigrants, and migrants (Baker et al., 2008). The researchers identified the most common collocations for the terms listed above and determined that these words were often represented in a less than favorable light. Further studies that have relied on corpora to examine collocates and shed light on how word choice may lead to biased representations of social groups include studies conducted by: Al-Heijn (2015) who examined the semantic macrostructures associated with Muslim women in the British news source BBC; Godbole, Srinivasaiah, and Skiena (2007) who used a corpus of newspapers and blogs to assess subjectivity based on sentiment analysis concerning entities (e.g. *people*); and Baker et al. (2012) who analyzed the representation around the word 'Muslim' in the British press. All of these studies demonstrated that the social groups being examined were commonly being presented negatively. Therefore, like with NLP studies, there is strong indication that negative sentiment, particularly of a group one finds opposing, is a strong indicator of potentially biased news. These findings, though focused on social biases, are

imperative for the study of political bias in the news as the same ideas of alienation and negativisation can perpetuate separation amongst political ideologies.

3.3.2 Linguistic Studies of Political Bias (not in news)

Other studies of linguistics focus on the language of political bias (outside of news) with the intent of explaining how language choices are reflective of users (i.e., politicians and their agendas) and how they might impact consumers (i.e., how an audience might be manipulated via language choices). Examination into how language choice is manipulated for political effect has shown that lying is marked by an avoidance of factive verbs, that self-protection and projected group responsibility are marked by the use of pronouns “I” and “we”, and othering is marked by pronouns ‘they’ and ‘those’ (Wilson, 2003). Thus, we can see that language choice can be reflective of political motivations.

Further studies show that indirect speech, e.g., metaphor, innuendos, and circumlocution are used more prominently during discussions of politically risky topics (Obeng, 1997). At the lexical level, critical discourse analysis studies have been done examining how metaphor, loaded words, and euphemisms can display negative sentiment about political actors or ideologies (Carver and Pikalo, 2008; Geis, 1987; Graber, 1981) thereby impressing those opinions onto an audience. At the grammatical level, CDA studies have examined how rhetorical moves (e.g.- transitions, greetings, etc.), the use of pronouns to create “in-groups” and “out-groups”, and persuasive strategies (e.g.-stance markers and qualifiers) are used to favorably or unfavorably relay political messages with an intent to persuade audiences to adopt their respective views (Chilton, 2004; Fairclough, 2000; Lin, 2011; Wilson, 1990). It seems that negative sentiment

(expressed lexically) and othering (expressed pronominally) are repetitively identified as being connected to political bias and Fake News.

A recent linguistic study, Grieve and Woodfield (2023) was conducted examining the language of fake news, which as was shown in chapter 2, is theoretically different from biased news but extremely linked; thus, this linguistic analysis is crucial to the present study since, to the best of my knowledge, it is one of the only linguistic studies examining misinformation. In this study Grieves and Woodfield (2023) apply methods of linguistic variation to the study of Fake News. Specifically, they compile a corpus of Fake News and real news, written by the same author, and identify grammatical features that were key in each group. They then functionally analyze the grammatical differences and what they demonstrate about real and fake news. In Fake News, they identify a higher reliance on emphatics, downtoners, adjectives and features associated with interaction over information density (i.e., shorter words, present tense, first-person, second-person, and third-person pronouns, verbs, perfect aspect, etc.). Conversely, real news relied on many features of information density (e.g., nouns, noun-modifiers, nominalizations, longer word length, etc), and interestingly, features of stance that the authors conclude demonstrates conviction (e.g., suasive verbs, possibility modals, by-passives, and Public Verbs). Importantly, the authors identified that methods of linguistic variation, particularly identifying key grammatical features and compiling comparative corpora that control for register, author, dialect, and political leaning can reliably identify functional variation in real and fake news. Like this study, the present study will also employ methods of register variation to examine media bias. Specifically, key features and keywords will be identified, and comparative (sub) corpora of biased groups will be compiled, and register, political leaning, topic, and time will be controlled for. Unfortunately, author or dialect cannot be controlled as the

interest is in identifying bias in news; thus, representing the population of news will necessarily equate to many different authors. However, this should not be an issue since Greives and Woodfield's major concern in controlling for author relates to the writing background and ability of professional journalists over those random individuals on blogs or social media posts. Most of the authors included in the present analysis are allegedly professional journalists (save for opinion submissions).

3.4 The Present Study

The overview of studies in both chapters 2 and 3 collectively demonstrate that language – both grammar and lexis – are at least tangentially connected to presentations of misinformation (bias and Fake News). From social science studies, we see that bias manifests in news by gatekeeping, coverage, and presentation bias and that coverage and presentation bias specifically, are presented through language. From NLP studies, we see that grammatical and lexical features have been used to automatically identify media bias and fake news. From linguistic studies we see that grammatical and lexical features are functionally related to political agendas, reflect social biases in news, and that grammatical features can functionally describe Fake News. What is missing is a descriptive account of the language of media bias. In other words, is there a systematic occurrence of language patterns across biased news texts (that is not attributable to another factor like variation in register)? This study will attempt to answer this question by combining methods of rater perceptions and construct definitions from NLP (e.g., methods of collecting rater perceptions and construct definitions), register variation (e.g., methods of corpus design and functional analysis of language patterns), and social science (e.g., analysis of the

three frameworks of media bias). This will add to our overall understanding of expression of information and misinformation in news.

To specify, the present study will compile a corpus designed to be representative of online news by collecting news articles on the same topic, published at the same time, and from a balanced array of publications of different political leanings (discussed in chapter 4). Then crowdsourcing will be relied upon to identify the texts from the corpus as biased (both direction and extent) (discussed in chapter 5), and results from the raters will be used to create the comparative (sub) corpora for which the linguistic analysis will be conducted on (discussed in chapter 7). The results of the linguistic analysis will then be analyzed for lexical and grammatical variation that can functionally describe biased news (chapters 8 and 9). Finally, results will be synthesized and connected to the three frameworks of media bias (chapter 10). In this way, it is the hope that a systematic pattern of language use of news bias, at least in terms of direction and extent, can be identified and subsequently analyzed to describe the language of media bias. Further it is the hope that this dissertation demonstrates how methods of register variation can add to examinations of misinformation. The present study will be guided by the following research questions:

- (1) Are there systematic linguistic differences among texts rated as biased in different directions?
 - a) Are there differences in key words?
 - b) Are there differences in key grammatical features?
- (2) Are there systematic linguistic differences among texts rated as biased to different extents?
 - a) Are there differences in key words?
 - b) Are there differences in key grammatical features?

CHAPTER 4. *SPIN*

4.1 Introduction

In order to answer the research questions, several methodological steps were necessary including compiling a corpus, having the texts from the corpus read and rated for bias, creating the bias groups, and conducting linguistic analyses of the texts using keyword and key feature analyses. The focus of this chapter is to describe the compilation of *SPIN*, the corpus used for the present analysis. In Chapter 5 the instrument used to obtain rating perceptions is discussed, in Chapter 6 the raters are discussed, in chapter 7 the ratings and how bias groups were determined are discussed, and in chapters 8 and 9 the linguistic analyses used to examine the texts are discussed.

4.2 Compilation of the Corpus

The compilation of the corpus required several involved procedures to ensure balance and representativeness. These procedures included selecting appropriate publications along a political cline as well as texts that were balanced for topic and time for each publication.

For this dissertation, the target language domain is biased U.S. online newspaper texts. In an ideal world, it would be all online news, however, with the sheer amount of online news and the arguments that exist about what constitutes ‘news’, this language domain would be very difficult if not impossible to capture. Since the interests of this study involve linguistically describing biased newspaper texts, the language domain is more specific. ‘Biased’ in this study includes extent (no bias to extreme bias) and leaning (liberal to conservative); ‘online news’ in this study includes American publications that have been politically rated by the three largest fact checking sites (see below). Thus, this language domain is represented in the corpus for this study

via the balanced selection of newspaper texts across political leaning, publication, topic, and time. The specifics of the compilation of *SPIN* are provided below.

4.2.1 Selection of the Publications

The publications sampled from are listed in table 4.1. These 15 publications were chosen based on several inclusion criteria. The first being that the publication needed to be a US publication targeted at American audiences to avoid issues of government censorship and expectations of objectivism which can differ depending on country. Targeting an American audience offers at least some control over the expectations of professional journalism without restricting the audience too much. Though a standardized expectation of professional journalism and objectivism is likely not something that exists, at the very least, targeting an audience from only one country standardizes the expectations of legal censorship and provides some cultural baseline of expectation.

The second criterion was that the publication has an online interface to ensure that medium was not a confounding factor in variation (i.e., a comparison of print news only to online only was excluded due to the nature of print media having constraints not present online). In addition, it is far more convenient to rely on digitized texts for the purpose of compiling a corpus. The third and final criterion was related to political leaning. Three publications were chosen from five divisions of political leaning: left, left center, neutral, right center, and right; for a publication to be sampled from, its political leaning must have had a similar (if not identical) rating by three different media sources: *Ad Fontes Media*, *Allsides.com*, and *Mediabiasfactcheck.com* (described in section 3.5). Therefore, the fifteen publications listed in the chart below were selected. It is important to note that a perfect match in political leaning for all fifteen publications was not something that occurred. Three publications in particular--

Breitbart, The Washington Times, and USA Today-- were rated slightly differently on *Mediabiasfactcheck.com* than on the other two checked sources.

Table 4.1

Publications included in the corpus

<u>Left</u>	<u>Left center</u>	<u>Center</u>	<u>Right center</u>	<u>Right</u>
AlterNet	CNN (online)	USA Today*	The Wall Street Journal	Breitbart*
The Nation	The Atlantic	NPR	Fox News (online)	NewsMax
HuffPost	The New York Times	Associated Press	The Washington Times*	National Review

**Matched on Ad Fontes Media and Allsides only*

4.2.2. Selection of Texts

Topic and Time. Once the publications were selected, criteria for text selection were also set. First it was determined that topic and time needed to be controlled for so that a fair examination of bias across publication could be conducted. To clarify, a topic (e.g., NBA players taking a knee) was examined within the same time (i.e., published around the same date to ensure that information was as similar as possible. Had I taken publications from different times, the information would likely be different. Thus, 25 texts on the same topic for each publication and year of interest were collected. The years of interest were 2016-2018, and 75 texts per publication were selected for a total of 1,040 texts (see table 4.2).

The selection of twenty-five topics per year was chosen after a search of trending topics for each year via *Ranker.com*, a website that specializes in ranking items based on internet trends. This website contains a year in review in which they examine ‘the best, most important, most influential, and most memorable’ topics of the year as ranked by hundreds of thousands of votes (*Ranker.com*, 2019). Since this website provided the top trends for each of the years of

interest in the present study and relied on systematic methods to rank trending topics, topics were selected based on the top trends listed for each year on this website.

An additional criterion for text selection was that the identified topic had to be something that was published about in several news sources. After checking the topics listed on *Ranker.com* against several news sources, 25 topics per year were deemed viable (see table 4.2). Topics of all types (i.e., not just those that were political or controversial in nature) were included in the study as the goal is to paint a comprehensive picture of newspaper bias and not focus on isolated events as previous studies have done. *Ranker.com* also provided several dates for the trending topics that were used as a starting point for divisions of time. To supplement the dates provided by *Ranker.com*, every trending topic was searched on news websites for possible dates of publication.

Time was divided into two-week periods to allow topics to be as similar as possible while still ensuring that any publications on the topic were captured (see table 4.2). Some topics had several possible dates as they occurred on multiple occasions or across many weeks. For example, the Winter Olympics of 2018, the Amazon Rainforest fires in 2019, and the investigation into Russian collusion in the 2016 election were all events that were published about repeatedly. In cases like this, the two-week time period that yielded the most publications across the 15 sources was selected. For all other topics, the date of occurrence was used.

<u>Topic</u>	<u>Time (2-week period)</u>
2016 Rio Olympics	Aug 8-22
Brexit	June 24- July 6

Black Lives Matter	July 8-22
Orlando Shooting	June 12-26
Gay Marriage	Jan 21-Feb 4
Refugee Crisis	Jan 27-Feb 11
ISIS	June 18-29
Gun Control debates	Oct 20-Nov 3
China as a Superpower	Oct 2-16
Presidential Elections	Nov 9-23
Coup in Turkey	July 15-29
Global Warming hoax?	Oct 28-Nov 11
Anthony Weiner texts	Oct 28-Nov 11
Trump Elected	Nov 9-Nov 23
Trump's Taxes	Oct 1-15
Sean Penn El Chapo Interview	Jan 11-25
Accusations against Trump (sexual assault)	Oct 8-22
Prince's Death	April 21-May 5
Hillary Clinton Emails	Oct 28-Nov 11
Melania Trump's Plagiarism	July 19-Aug 2
NYC Bombings	Sep 17-Oct 1
Zika Virus	Sep 1-15
Harriet Tubman on \$20	April 20-May 4
Saudi Arabia 9/11 Bill	April 15-29
Death of Antonin Scalia	Feb 13-27

	<u>Topic</u>	<u>Time (2-week period)</u>
2017	Solar Eclipse	Aug 21-Sep 5
	#MeToo	Dec 21-Jan 4
	Wonder Woman movie	June 5-19
	Taking a Knee	Sep 28-Oct 12
	Royal Wedding/ Markle	Dec 21-Jan 4

Alec Baldwin SNL	Jan 14-28
Women's March	Jan 21-Jan 6
United drags man off Plane	April 10-24
The Travel Bans	Mar 6-20
Covfefe	May 31-June 14
John McCain Healthcare vote	July 28-Aug 11
Russian investigation	Oct 26-Nov 9
Kathy Griffin	May 31-June 14
Melania moves to WH	June 12-26
Trump's judicial picks	Dec 3-18
Robert Mueller	Dec 11-25
Amazon Buys Whole Foods	June 16-30
Mike Pence Mideast Visit	Dec 18-Jan 1
James Comey Testimony	June 7-21
Pepsi Commercial	April 5-19
Millennials and the economy	Nov 3-17
Betsy DeVos	Feb 10-24
Las Vegas Shooting	Oct 2-16
Kennedy Center Honors	Aug 17-31
Trump's National Security Strategy	Feb 13-27

	<u>Topic</u>	<u>Time (2-week period)</u>
2018	CA Wildfires	Nov 13-27
	Black Panther (movie)	Feb 17-Mar 3
	Death of George H. Bush	Nov 30-Dec 14
	Trump and NK disarmament	July 2-16
	Midterm Elections	Nov 6-21
	Separation of Child Immigrants	June 18-July 2
	Roseann Barr	May 29-June 12
	Hurricane Florence	Sep 14-28

FIFA World Cup	June 14-28
Aretha Franklin Death	Aug 16-30
March for our Lives	Mar 24-April 7
Winter Olympics	Feb 8-22
Kavanaugh Supreme Court	Sep 27-Oct 11
Elon Musk's Tweets	Mar 21-April 4
Net Neutrality	June 11-25
Russian Inquiry	Nov 17-30
Nike & Kaepernick	Sep 4-18
Taylor Swift Politics	Oct 7-21
Serena Williams Tennis sexism	Sep 9-23
'Very Stable Genius': Trump	Jan 6-20
Stormy Daniels	Mar 26-April 10
Fire and Fury book	Jan 8-22
Trump's Tariffs	Mar 2-16
Mira Ricardel Firing	Nov 13-27
Kanye West Politics	Oct 20-Nov 2

Locating texts. Once the topics and times were determined, manual collection of texts was conducted using two databases: Access World News, and U.S. Major Dailies. Access World News is the world's largest research database of news and contains archives of most of the publications in the present study from the 1980s (or earlier) to present. The database allows the user to search for texts within specific publications and specific dates. This database was used where possible, and the search terms used, as well as the dates searched, are those from table 2. Texts from publications that were not accessible on Access World News were retrieved from U.S. Major Dailies as well as Google.

U.S. Major Dailies uses the ProQuest platform to allow access to *The New York Times*, *The Washington Post*, *The Wall Street Journal*, *The Los Angeles Times*, and *The Chicago Tribune* from 1980-present. This database was used to collect texts from *The New York Times* and *The Wall Street Journal*. For all searches conducted, the dates and search terms are those from table 2. For all other text collection, Google and the web pages belonging to the publications were used. Searches on Google consisted of the year, month, topic, and publication to yield results. For example, to find articles on Prince's death from HuffPost "2016 April Prince's death HuffPost" was typed into Google's home page (all searches for this topic looked identical save for the publication). The results yielded by the Google search were then examined to locate an article published within the two-week period. In all cases, the first eligible link was the text collected for the corpus.

When the databases and Google yielded no eligible results, the web pages for the specific publications were searched and eligible texts were collected; however, there were many occasions where there were no texts on the topic of interest within the two-week period; this was recorded for a later examination of gatekeeping bias. In addition, there were several occasions where a publication posted an article taken from a different source (e.g., *The Nation* posts an article taken directly from *Tomsdispatch.com*). These articles were collected when they were the only eligible text on a topic since this is what readers will encounter. In these cases, the original publisher as well as the publication it was found in were recorded.

Once a text was selected, it was copied into a doc file, cleaned, coded for relevant variables and metadata, and saved as three separate files: the original cleaned word doc, an anonymized version for raters, and a .txt version for the linguistic analyses. Cleaning the texts involved removing all hyperlinks, deleting non-linguistic data (photos, graphs, screenshots of

tweets, etc.), and removing any embedded formatting. Further steps were taken to clean the anonymized versions including removing the headings, deleting any identifying information (author's name, publication name, and date), and converting each text into a unified format (12 point, Times New Roman, 1" margins, single spaced). The texts were collected and compiled to form the corpus, *SPIN*, described below.

4.2.3 *SPIN*

The 1,040 collected texts comprise the corpus created for this study which has been titled *SPIN (Survey of Politically Inclined News)*. As table 4.3 below demonstrates, the corpus contains a total of 864,094 words. The goal in compiling this corpus was balance and representativeness rather than size as Biber (1993) recommends focusing on balance and representativeness over size for best corpus sampling practices. Because the interest here is in examining the linguistic features of bias in news, representativeness was incorporated via the collection of texts from American news sources across a political spectrum on the same topic and within the same time-period. It would be ideal to provide a comprehensive overview of what online news *is* or even biased online news for that matter and then evaluate how representative *SPIN* is. However, given the fact that online news is very difficult to classify, there is not much information available to provide a comprehensive overview of what constitutes 'typical online news'. For example, a search about article types (i.e., how many political articles vs opinion are typically published online) yielded no results. Some studies classifying online news have been conducted, but they focus on classifying false vs real information or categorizing blogs vs news. Therefore, evaluating how representative *SPIN* is of the target language domain will not be possible at this time.

Balance was incorporated via the selection of three publications per political leaning and by the collection of the same number of texts per publication (when possible). As text is the unit of analysis, ensuring a collection of a similar number of texts across publication was more important than the number of words; however, due to the close control of topic and time, some publications contain less texts than others as they did not publish on a topic within the required time frame.

Table 4.3
Overview of SPIN

	<u>Publication ID</u>	<u>Publication</u>	<u># of Texts</u>	<u># of words</u>
Left	1	AlterNet	67	53,444
	2	The Nation	48	65,754
	3	HuffPost	75	55,445
Total			190	174,643
Left Center	4	CNN	74	64,338
	5	The Atlantic	68	84,831
	6	The New York Times	75	80,313
Total			217	229,482
Center	7	USA Today	73	52,185
	8	NPR	67	56,607
	9	Associated Press	72	55,514
Total			212	164,306
Right Center	10	The Wall Street Journal	69	54,228
	11	Fox	74	47,076
	12	The Washington Times	72	52,685
Total			215	153,989
Right	13	Breitbart	72	38,675
	14	NewsMax	71	41,560
	15	National Review	63	61,439
Total			206	141,674
Overall Total:			1,040	864,094

CH 5. METHODS OF MEASURING BIAS

5.1 Introduction

Chapter 4 discussed the compilation of *SPIN* and detailed the collection of the texts that were given to raters in this study. The purpose of this chapter will be to motivate and describe how ‘bias’ was operationalized in this study (section 5.2) before discussing the instrument that was used as a means of collecting bias (section 5.3). Chapter 6 will then present the results from the instrument beginning with the population of raters in this study and how it compares to the average online newsreader, before presenting the article bias ratings.

5.2 Defining Bias

As was stated in the literature review, many studies of media bias define bias as “slanted news coverage or internal bias, reflected in news” (Hamborg et al., 2018); however, this definition is vague and difficult to operationalize. For this dissertation it is crucial to understand what raters might imagine bias to be given that the goal is to attempt to identify, analyze, describe, and understand the linguistic variation of news perceived as biased. Further according to Grieve & Woodfield (2023), in examinations of misinformation – in this case bias – the construct needs to be defined in a way that is specific and meaningful and then applicable to identify instances of bias (Grieve & Woodfield, 2023). For this study, bias is not actually operationalized. Instead, it includes anything that could be incorporated when people claim that news publications are “biased”. For my study, this is meaningful as I am concerned with perceptions of media bias. Thus, bias could include anything that occurs when journalists or publications allow their own opinions to affect the news that they report and the way that they report it. This can include one-sided reporting, missing information, positive or negative language use, playing to a home crowd, making topics political, etc. To help highlight what I mean, below are some open-ended responses to the survey where raters could write additional thoughts about why they marked an article as biased or not:

Reasons respondents gave for marking an article as biased:

- “Trump and the White House were gratuitously and negatively included in an article that didn't involve politics or government at all.”
- “When the author uses words like "smarmy", the connotation is negative and reinforces his bias”
- “It almost seems as if the author has a personal vendetta, not against Sean Penn, but the magazine Rolling Stone.”
- “There are assumptions all over”
- “The italic text at the top of the article sets an anti-trump tone which undermines the first three paragraphs that attempt to be unbiased. The article avoids directly stating anything negative about Trump, but the message is that "these people don't like Trump so you shouldn't either.”
- “There were no citations that one could check the veracity of the accounts. There was no explanation of who was fleeing and why. Words like "Islamophobic" are Ad Hominum and not explanatory enough.”
- “It Speaks strongly against his [Trump’s] intelligence, and disrespects him at any given opportunity even insinuating his followers are falsely loyal. He's also painted a conspiracy theorist.
- “People are too focused on the racial aspect of the situation then the message its trying to portray.”

Reasons respondents gave for marking an article as unbiased (or mostly unbiased):

- “I felt like this was very balanced. Just reported the facts and did so in a balanced manner.”
- “Non-bias reporting. Facts are stated without opinion”
- “The only thing I think could really be taken as bias is that they only interviewed the protester and no one from the other side, but that's pretty minor.”
- “It seems fairly neutral because the article describes Trump's confirmations as "wins" but the language in the article (e.g. using terms such as "anti-abortion" rather than "pro-life") makes it sound like the author is probably moderate, if not slightly liberal.

- “I see no bias, I think the article is just summarizing things as they occurred.”
- “I believe this is a well written article that only tells the facts of what happened. This is something I rarely see anymore and I do appreciate it.”
- “It [the article] doesn’t take sides on whether US should have intercepted information”

As is demonstrated by the comments above, perceptions of bias consider a range of things, and so, the ‘definition’ for bias in this paper will remain fairly inclusive, and I will not attempt to succinctly define it.

However, to define the construct of bias for the instrument used in this study (described below), I drew from previous studies in NLP that have relied on crowdsourcing perceptions of bias as well as one study that conducted a factor analysis to identify the latent constructs of bias. Crowdsourcing perception was chosen as a means of identifying bias for two reasons: 1) crowdsourcing has been shown to resolve many complex problems, of which identifying bias is one (Mitrović, 2013; Mladenović et al., 2016), and 2) many studies on media bias as well as news aggregators (Allsides, PolitiFact, etc.) rely on public perceptions of bias; cross-checking between ratings in research to those from aggregators has shown high rates of agreement. Therefore, bias in this study is determined by raters.

In order to determine what to ask raters to reliably collect measures of article bias, I relied on a meta-analysis conducted by Spinde et al. (2021). In this study, the researchers examined 824 questions designed to collect perceptions of media bias with the goal of identifying a scale that can be used as a reliable standard to evaluate article bias. To create this scale, the authors first pulled together a large list of questions from media bias perception studies, conducted a factor analysis on the questions to pull out the underlying constructs of bias, and then subsequently tested the underlying variables’ suitability and reliability as measures of collecting perceptions of bias. Six indicators of bias were identified by the factor analysis in this

study: ‘factuality’, ‘influence’, ‘topic affirmation’, ‘negative emotion’, ‘bias’, and ‘political ideology’; however, the authors note that ‘factuality’ and ‘bias’ were highly interrelated, and ‘influence’ and ‘topic-affirmation’ were heavily context-dependent, so they concluded that the factors of ‘bias’ and ‘political ideology’ were the most useful factors for collecting perceptions of bias - they were also the most heavily agreed upon in their subsequent reliability tests.

Due to the fact the ‘bias’ is pretty abstract and to add clarity to ‘political leaning’, table 5.1 below provides the questions that positively loaded for the factors of ‘political leaning’ and ‘bias’ in Spinde et al.’s (2021) meta-analysis.

Table 5.1
Questions that loaded positively on relevant factors from Spinde et al. (2021)

<u>The ‘Political leaning’ Factor</u>	<u>The ‘Bias’ Factor</u>
Liberal/conservative	How biased is the article?
Where would you place the personal political view of the author responsible for this article?	Do you think a reader would perceive the arguments presented in this article to be stronger on one side of the issue than the other?
	To what extent do you perceive that the author's reporting on the issue reflects a bias?
	Biased/ unbiased
	Strictly neutral/biased
	Not/ favoring a side

As the table shows, both of the positively loading questions for political leaning are requesting perceptions about an article’s liberal/conservative divide while the questions for ‘bias’ are requesting perceptions about the strength of bias (i.e., how biased it is) and, to a lesser extent, the one-sidedness of an article. Based on the results of this analysis, the present study collected perceptions of bias based on these two identified factors of ‘political leaning’ and ‘bias’, but

‘bias’ was reduced only to extent to avoid conflation. To further avoid conflation, the questions about direction and extent of bias are asked only once which allowed for a one-to-one correlation from question to a bias rating.

5.3 The Collection of Perceptual Ratings

The next step in this study was to have the collected texts in *SPIN* read and rated by participants for bias. The purpose of this section is to explain the methods behind the collection of the perceptual ratings relied upon in this study. This includes the description of the crowdsourcing tool used to recruit raters, *Mechanical Turk* (5.3.1), and the instrument used to collect ratings (5.3.2).

5.3.1 Mechanical Turk

The survey created for this study, described below, was posted on *Mechanical Turk* to recruit raters. *Mechanical Turk* is a crowdsourcing website hosted by Amazon designed to allow researchers to quickly get massive amounts of data from human workers around the world. On this platform, workers (here called raters) sign-up on *Mechanical Turk* and are presented with a list of HITS (here called surveys). Requesters (or those listing and paying for the HITS), can post a variety of surveys for raters to complete. Requesters can also set qualifications (or specific requirements) for the raters allowed to complete their surveys. For example, a requester can require that a rater be located in the U.S., speak fluent English and beginner Spanish, and be over 50 years old. *Mechanical Turk* also allows a requester to determine how many surveys an individual rater can complete and allows the requester to review responses before accepting (and

thus paying for) the work. *Mechanical Turk* has been shown to be a viable and generalizable sampling platform when a general population sample is needed (Gerlich et al., 2018). For example, an examination by Clifford, Jewell, and Waggoner (2015) into the validity of data from MTurk for research in political ideology demonstrates that that respondents recruited from MTurk are often more representative of the U.S. population than in-person convenience samples. Additionally, Berinsky et al., (2012) show that the demographic characteristics of domestic MTurk users are more representative and diverse than alternative convenient samples.

A requester has the option of using *Mechanical Turk's* existing code to create a survey or to link to an outside survey. For the purposes of this dissertation, the option of the link to the outside survey was selected and 3 raters per article were requested for a minimum of 3,120 responses. Further qualifications were set requesting that the raters be located in the U.S. and speak English (self-reported). Location was restricted to the U.S. as the interest is in perceptions of news created for American audiences and language was set to high proficiency in English to ensure no language barrier in understanding the news articles.

The reason for the minimal number of exclusions is twofold- first, it was difficult to accurately determine who the 'typical' readers of online news are, and second, the goal was to recruit as many raters as possible and too many restrictions narrowed the result pool. To avoid accepting ratings given by nefarious means, the time it took for each survey to be completed was checked and two questions designed to ensure a human response were included (discussed below). In cases where the survey was completed at inhuman speeds (i.e., milliseconds to a few seconds) or the responses to the open-ended questions were Wikipedia excerpts rather than a coherent response, responses were not accepted. With these restrictions, 3,510 responses were accepted from raters -- the demographics of which are presented in Chapter 6.

5.3.2 *The Instrument*

Once *SPIN* was compiled, raters read and rated each text from the corpus using a perceptual rating instrument created for this study. The perceptual rating instrument was created on Qualtrics and was placed on *Mechanical Turk* for rater recruitment (discussed above). The instrument (see appendix A) consists of fifteen items, the first two of which explain the background of the research, provide instructions, and request consent. The next five items of the instrument request background information from the rater; particularly their worker ID, age range, sex, and highest degree earned. These items were requested so that an analysis and comparison of the raters from this study and the average online newsreader could be conducted. This is presented along with the demographic information about this population in chapter 5. Self-reported political leaning was also collected as part of the background questionnaire as a means of checking representativeness of political affiliation amongst raters. This question was asked for two reasons: first, several studies on the psychology of bias show that political background can be a determinant of bias labeling (Pennycook & Rand, 2021; Van Bavel & Pereira, 2018), and second, collecting perceptions from one self-identifying liberal, conservative, and moderate per article have been shown to mitigate this effect, and so, the goal was to be able to recruit an even mix of raters from differing political leanings.

The next two questions on the survey were designed to provide the article to the rater and to ensure that raters were human and not a computer program designed to take as many surveys as possible. Figure 5.1 shows that question 8 required raters to click a link to download the article, and question 9 asked them to write the number that appeared in the link. This allowed me

to track when articles were downloaded and to check that the answer entered matched the link.

When a response was submitted by a computer program, the answer to question 9 was typically a random Wikipedia entry. All responses to surveys that incorrectly answered question 9 or did not click the link in 8 were thrown out as this indicated that they were not a human rater and/or did not read the text.

<p>Q8 Please click the link to access the article. It will download as a word document. Please read the text in its entirety before answering the following questions.</p> <p>Q9 Please write the full number that appears in the link above. Example 1 (#).</p> <p>_____</p>

Figure 5.1 *Example of Filtering Questions*

The next four items on the instrument asked the raters to provide their perceptions about the article read. First, they were asked whether they agree with the perspective presented in the article. They were presented with a 5-point slider scale (1-Definitely yes, 2-Mostly yes, 3-Might or might not, 4-Mostly not, and 5-Definitely not) (see figure 6). This question was included as a means to determine whether there is any effect from rater’s self-identification of agreement on their ratings of bias since people are less likely to perceive bias in news content that is concordant with their existing beliefs (Pereira et al., 2023).

Q10 Do you agree with the article's perspective?

	Definitely yes	Mostly yes	Might or might not	Mostly not	Definitely not
	1	2	3	4	5
Agreement with Perspective 1 ()					

Figure 5.2 *Question on Agreement with Article’s Perspective*

Next raters were asked to indicate to what extent they believe the article was biased on a 0 to 5 scale (0 being no bias and 5 being extreme bias). Raters were then asked to indicate which direction the article leans also using a 5-point scale (1- strongly left to 5-strongly right) and were asked to add a short description of why they made their selection. This question was given as an open-ended response so raters could provide information about any element that led to their rating.

Q11 Please select the extent to which you believe the article is biased.

	Not at all biased	Slightly biased	Somewhat biased	Pretty biased	Extremely biased	
	0	1	2	3	4	5

Extent to which the article is biased ()	
--	--

Q12 Please select the direction you believe this article leans.

	Strongly Left/Liberal	Left/Lib eral	Neutral (no obvious leaning)	Right/Con servative	Strongly Right/Con servative
	1	2	3	4	5

Direction the article leans ()	
--------------------------------	--

Figure 5.3 *Items Inquiring about Perspectives on Extent and Direction of Bias*

Once all the ratings were collected, information about rater backgrounds were organized and analyzed to create a profile of whether the raters here match the typical online newsreader. This information is discussed and presented in chapter 5. In addition, mean ratings were used to create the bias groupings for the empirical examinations in this study. Both of these are presented in the next chapter.

CHAPTER 6. THE RATERS

6.1 Introduction

The purpose of this chapter is to present a profile of who the raters in this study are and to compare the demographics of this population to that of average online newsreader. In this way, we can determine whether the population of raters in this dissertation are representative of the population of online newsreaders and thus can determine if their ratings of bias might be like that of the average news reader.

6.2 The Demographics of the Average online News Reader

According to Nielson data, 69% of the American population reads a newspaper at least once a month (2016). Demographic information about those readers available online was limited to age, sex, and education which is why that demographic information was requested of the raters in the present study. According to *Pew Research Center* (2011;2016) and Anderson (2018), the average online newsreaders are between the ages of 18 and 49, evenly split between men and women, and tend to be educated at the level of some college or higher. It was difficult to find information about the average political leaning of the online newsreader as it varies widely depending on source. Section 6.3 below will provide the demographic information about the population in this dissertation and compare each demographic point to that of the average online news reader.

6.3 The Demographics of the Raters

As already stated, the demographic information collected about raters included political leaning, age, education, and sex. Political leaning was collected as a means of ensuring a fair

balance of political affiliation amongst raters. The division of age, sex, and education of the raters were collected to determine how they compare to the typical population of an online newspaper and whether they are representative of that population.

In terms of political leaning, although the original goal of recruiting one self-identifying liberal, conservative, and moderate per article was not met, there was a fairly even split of self-identifying liberal, conservative, and moderate raters overall with a total of 3,510 raters. As figure 6.1 illustrates, 1,202 (34%) self-identified as liberal; 1,110 (32%) self-identified as conservative and 1,198 (34%) self-identified as moderate. Figure 6.2 and figure 6.3 illustrate how raters self-identified their age and sex respectively while figure 6.4 illustrates the division of education.

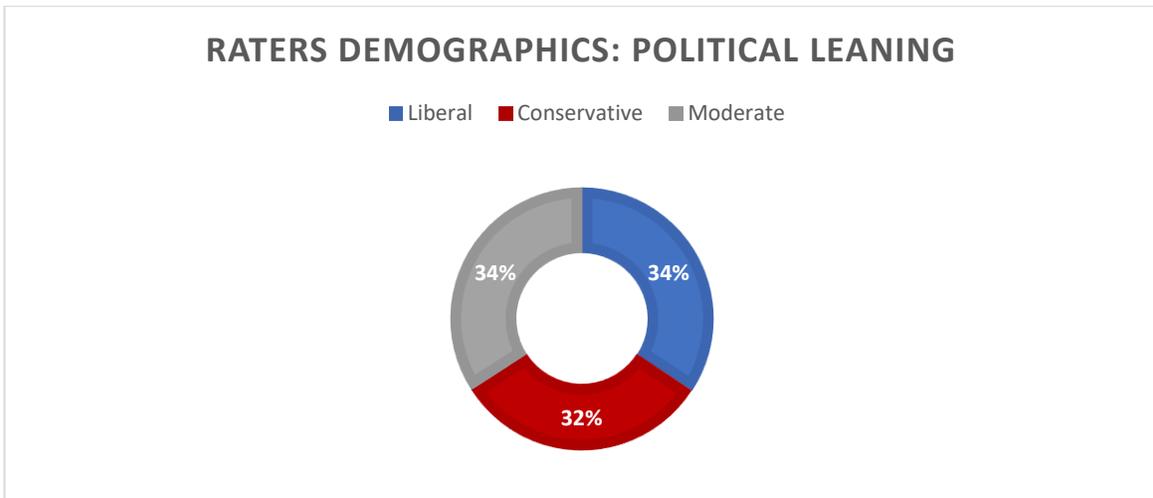


Figure 6.1 Division of raters by self-identified political leaning

According to *Pew Research Center* (2016), the majority of online newsreaders are between the ages of 18 and 49 (~82%). As you can see from figure 8, the majority of the raters in the present study closely match that of the average online newsreader with 82.6% being between the ages of 18 and 49 suggesting that, in terms of age, the population here is representative of online news readers.

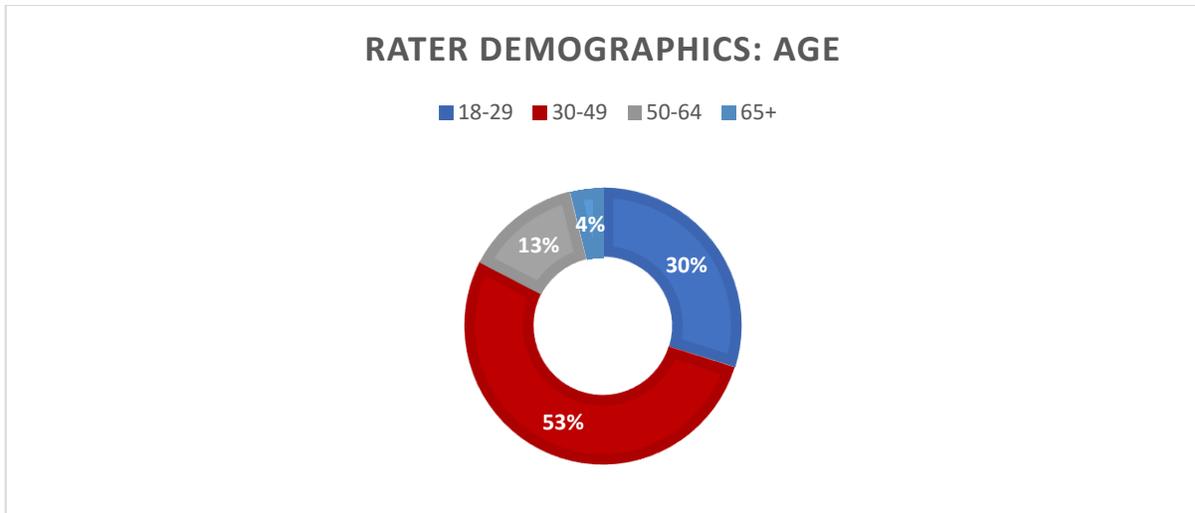


Figure 6.2 *Division of Raters by Age*

As far as demographic information on sex, *Pew Research Center* (2011) reports that overall consumption of online news is fairly even between men and women though it fluctuates dramatically based on the specific source. As figure 6.3 shows, the raters for the present study were nearly evenly split with slightly higher percentage of self-identified female at 51% and male at 48%.

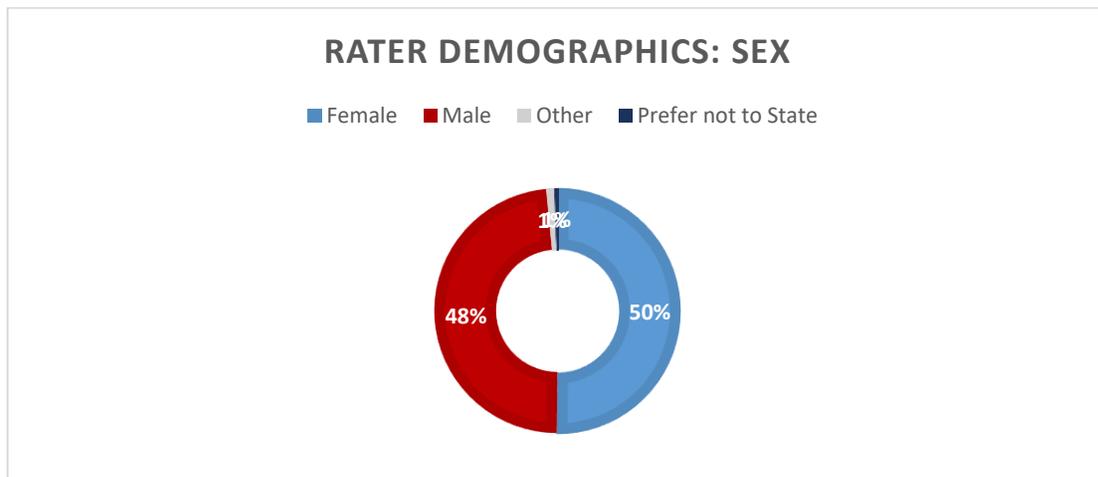


Figure 6.3 *Division of Raters by Sex*

As for education, according to Anderson (2018), the majority of people with an associate degree or higher prefer reading news online while those with lesser education prefer to watch the

news. It is hard to determine how representative the present study’s educational divide is to online newsreaders with this limited information, but as shown in figure 11, the majority of the raters in the present study hold an associate degree or higher with only around 21% reporting a lesser degree.

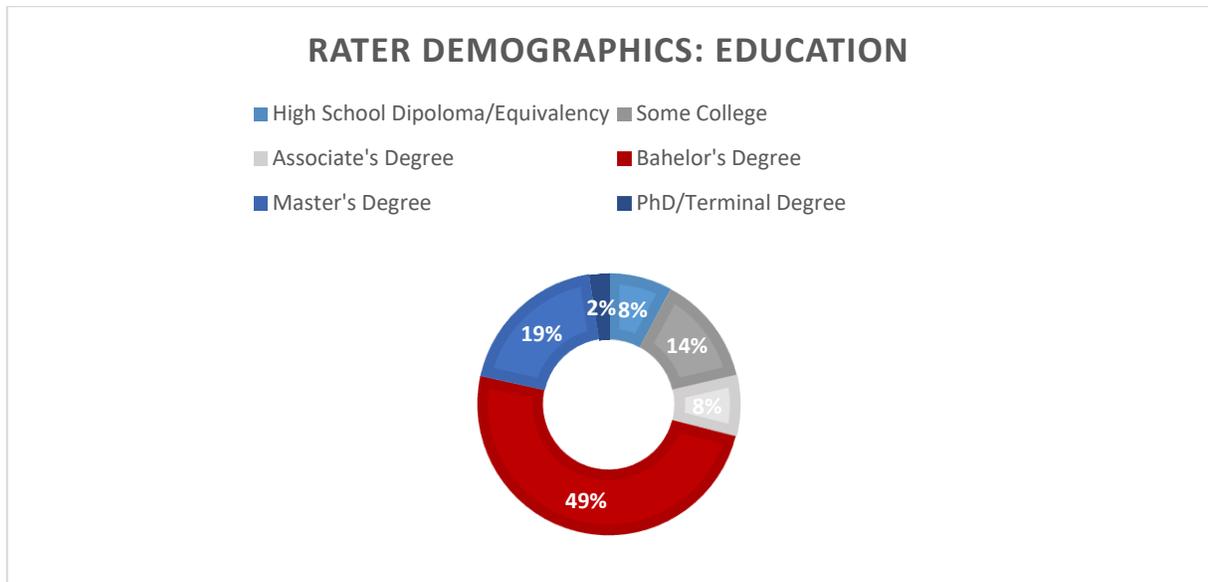


Figure 6.4 *Division of Raters by Education*

Based on this demographic information, we can assume that the raters are at least similar to the population of online news readers with very near identical splits in political leaning, gender and a mostly representative divide in education. This is important to note, as perception of bias from these raters is the baseline for the present analysis; thus, given that the raters closely represent the average reader of online news, we can feel confident that the perceptions of bias are likely representative of the average online newsreader.

CHAPTER 7. DEFINING BIAS WITH RATINGS

7.1 Introduction

Bias in this study was defined by rater perceptions and included collecting perceptions about an article's direction of bias and extent of bias and then using those ratings to group texts into comparable groups of bias. The impetus behind grouping the texts by biased ratings comes from the fact that an examination of linguistic variation requires comparable groups of texts - in this case the comparable groups for direction will be left and right rated texts and extent will be the texts rated as extreme extent and no extent. The methods relied upon to collect perceptions of bias were presented in chapter 5. The purpose of this chapter is to discuss the methods relied upon to group texts based on ratings of bias (section 7.2). Then, in section 7.3, this chapter will detail the characteristics of the texts in each bias group so that the subsequent linguistic analyses in chapters 8 and 9 are clear.

7.2 Using Ratings to Determine Bias Groupings

Once all perceptions were collected from raters, information from survey items were used to rate each text for bias in terms of direction and extent. As was discussed in Chapter 5, raters were asked to select the extent to which they found the text biased and were presented with a sliding scale from zero to five, zero-1 being no bias, 1.1 - 2 being slightly biased, 2.1 - 3 being somewhat biased, 3.1 - 4 being pretty biased, and 4.1- 5 being extremely biased. Each text had 3 individual raters, so ratings for each text were combined and mean ratings were used to determine the extent of bias for each article. Next raters were asked to state which direction the text leans (strongly left, left, neutral, right, and strongly right). Selections were converted into an ordinal scale from 1-5. One indicated left while five indicated right and three indicated neutral. Mean answers to this question were used to determine the direction of bias.

Each article thus received two averaged ratings: one for extent and one for direction, and each of these ratings were used to place an article into a bias group. In other words, every article

was placed into two bias groups – a group of extent and a group of direction. Based on the mean scores, articles were separated into groups for extent and direction as follows: 0-1.3 – no extent, 1.4-3.6 – some extent, and 3.7-5 – extreme extent and 1-2.4 – left leaning, 2.5-3.5 - neutral, and 3.6-5 -right leaning. So, an article that had a mean score of 4 for extent and 2 for direction was put into the left leaning group and the extreme extent group.

Due to the fact that there were 3 raters per text, an important consideration is rater agreement. Two types of agreement were calculated in this study – the first was agreement between raters within the study and the second was agreement between raters in this study to outside ratings. Agreement for raters within this study was calculated by a simple percent agreement for each of the biased ratings (i.e., a percentage of agreement for direction of bias and a percentage of agreement for extent of bias). Agreement for direction of bias was fairly high with 93% inter-rater agreement. Agreement for extent of bias was much lower with only 79% of inter-rater reliability.

Ratings from this study were also compared to ratings from third-party raters that have rated direction of bias of the publications that are examined in this dissertation. Specifically, ratings from the present participants were compared to those from third party media rating companies (*Ad Fontes Media*, *Allsides.com*, and *Mediabiasfactcheck.com*) to determine what level of agreement exists. This was done as a means of calibrating current perceptions to those on a larger scale and to determine whether raters in the present study had an ‘accurate’ sense of a publication’s leanings. Each article’s rating was compared to the publication’s leaning as determined by the third-party media bias companies. It is important to note that third-party media companies use differing scales than the one in the present study (left, neutral, right) so for this comparison to make sense, the 5-point scale from the perception instrument was reduced to a 3

point scale. Results, presented in table 7.1 below, showed that perceptions lined up closely with third-party ratings; the lowest rate of agreement being 52% and the highest being 100%.

Considering that there were over 3,500 raters in the present study, this level of agreement both within this study as well as to the third-party classifications is considerable. It suggests strongly that the articles are marked in some universal way for political leaning -- whether these articles are marked linguistically is exactly what the present study aims to determine.

Table 7.1

Agreement for Dissertation Raters and Outside Ratings for Direction of Bias

Publication	Total # of Articles	Left	Neutral	Right	% correct (w/3 rd party)	Third-party Leanings
AlterNet	67	52 (78%)	15 (22%)	0	78%	Left
The Nation	48	40 (83%)	8 (17%)	0	83%	Left
HuffPost	75	50 (67%)	24 (32%)	1 (1%)	67%	Left
CNN	74	34 (46%)	37 (50%)	3 (4%)	96%	Left Center
Atlantic	68	31 (46%)	37 (54%)	0	100%	Left Center
NYT	75	31 (41%)	42 (56%)	0	100%	Left Center
USA Today	73	15 (21%)	54 (74%)	4 (5%)	74%	Center
NPR	67	14 (21%)	49 (73%)	4 (6%)	73%	Center
AP	72	6 (8%)	63 (88%)	3 (4%)	88%	Center
WSJ	69	0	46 (67%)	23 (33%)	100%	Right Center
Fox	74	2 (3%)	44 (59%)	28 (38%)	97%	Right Center
TWT	72	0	30 (42%)	42 (58%)	100%	Right Center
Bb	72	2 (3%)	22 (30%)	48 (67%)	67%	Right
NewsMax	71	2 (3%)	32 (45%)	37 (52%)	52%	Right
Nat Review	63	1 (2%)	21 (33%)	41 (65%)	65%	Right

7.2.1 Bias Groups

Because an article received a score for direction of bias and extent of bias, overlap between the articles in the extent groups and the direction groups occur; however, it is important to note that groups that are comparatively analyzed, e.g., the left leaning and right leaning group

(a.k.a. direction groups) and the extreme extent and no extent group (a.k.a. extent groups), are completely distinct from one another. In other words, the overlap that occurs does not directly impact the analyses in the present study.

For an illustration of how each text from *SPIN* was rated and subsequently separated into a bias group, see figures 7.1 and 7.2 below.

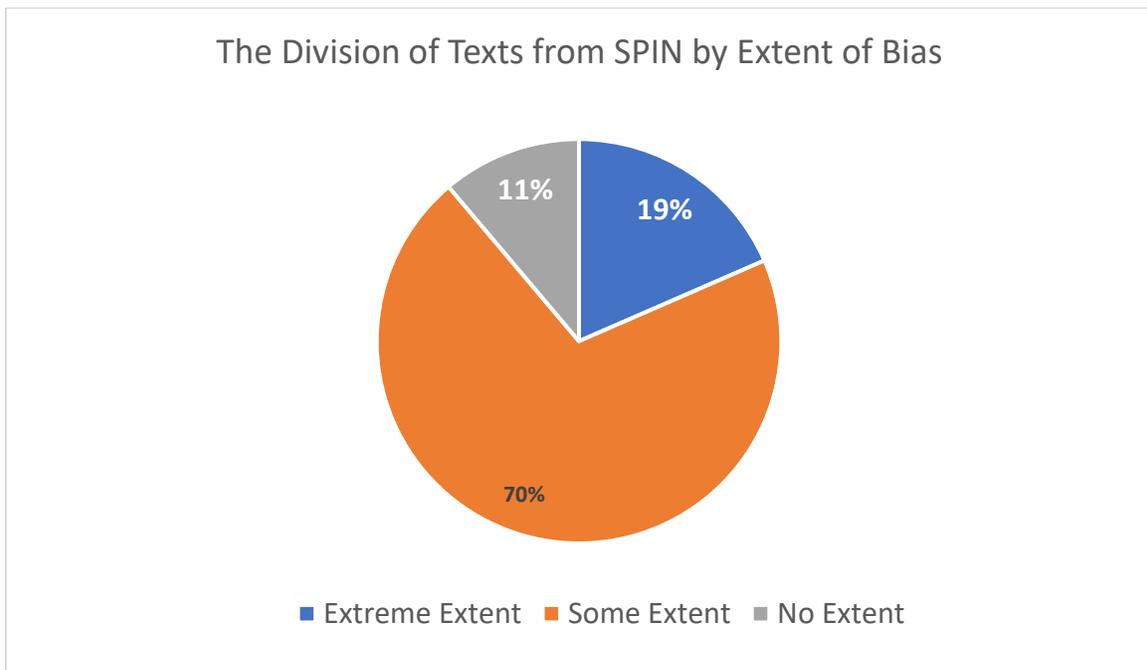


Figure 7.1 *The Division of Texts from SPIN by Extent of Bias*

Figure 7.1 illustrates the division of the 1,040 texts from *SPIN* by extent of bias. In total, 192 texts were rated as having extreme extent, 116 were rated as having no extent, and 732 texts were rated as having some extent. The ‘some extent’ group is rather large with more than 70% of the texts from *SPIN* ending up in this group. This result was not surprising for three reasons: 1) texts were collected from professional and, for the most part, reputable newspapers who strive to appear impartial and use language that does not reveal their opinions directly, 2) the 5 point scale from the perception instrument was reduced to a 3 point scale for this division of bias and the

‘some extent’ scale was by far the largest, and 3) raters typically select a middle ground rating over an extreme rating if they are unsure or have no strong opinion (Kulas et al., 2008).

The Division of Texts from SPIN by Direction of Bias

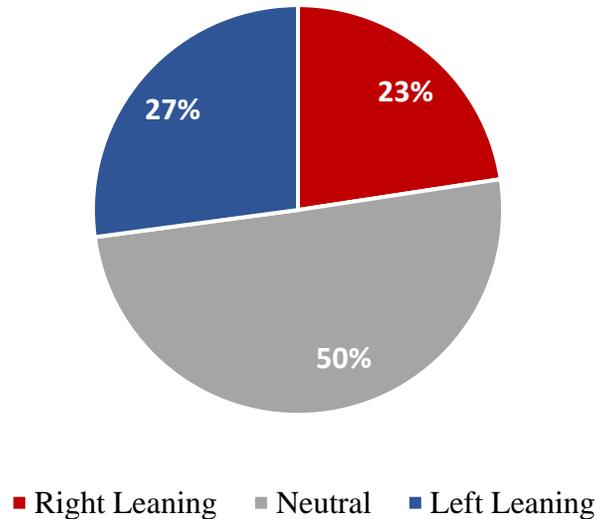


Figure 7.2 *The Division of Texts from SPIN by Extent of Bias*

Figure 7.2 illustrates the division of the 1,040 texts from *SPIN* by direction of bias. In total, 282 texts were rated as left leaning, 517 were rated as having no extent, and 235 texts were rated as right leaning. As with the extent group, the largest group is the middle ground with just under 50% of the texts from *SPIN* being rated as neutral. This, again, was not surprising for similar reasons of journalists striving to be impartial and raters favoring a middle ground rating. In addition, because topic was controlled for, many of the articles here are describing the same real-world event and will thus share many similar details of the event, regardless of their political perspectives.

Although groups were created for extreme extent, some extent, and no extent as well as left leaning, neutral, and right leaning, the linguistic analyses (chapters 8 and 9) only examine

articles that are grouped into the polar categories of extreme extent and no extent as well as left leaning and right leaning. Because the goal of the study is to determine whether there are differences in grammatical and lexical patterns in texts rated as biased in different directions as well as texts rated as biased to different extents, it was decided that including the categories that represented the middle ground would create too much noise in the analyses. Thus, comparisons are only conducted on extreme extent to no extent and left leaning versus right leaning. This is not to suggest that there is nothing interesting to be learned from examinations of all of the bias groups, in fact there are many different ways these groups could be interestingly compared (e.g., neutral to left or right); however, the examinations in this dissertation are a first step in understanding the language of bias present in this data, and I believe that the features of the extreme categories will be most informative. Future examinations will be conducted on all of the collected texts, but the rest of this analysis will focus only on texts that ended up in the extreme extent group, no extent group, left leaning group, and the right leaning group. This means that only 666 texts from *SPIN* are being considered in this analysis, and 374 are being put aside for future examinations.

I acknowledge that this is problematic as the 666 texts only comprise ~66% of the corpus, *SPIN*, that has been described and analyzed here. This is a major problem as the corpus was already questionably representative of the domain of online news, and of course this sub-sample of 666 texts is not representative of the entire domain. However, as we are missing a description of the language of biased news, beginning with the linguistic patterns of texts rated into the most polar categories (extreme bias, no bias, and left, and right) will be most informative at this stage. In other words, it was determined that reducing noise in the data would be more beneficial at this stage in the examination than conducting the analysis on the full collection of texts. Again, future

examinations will be conducted on the entire corpus. To paint a clear picture of the 666 texts that are relied upon in this dissertation, a profile of text characteristics is provided below.

7.3 Group Profiles

Topic and political leaning of publication were carefully controlled for in the original collection of texts in *SPIN*, but naturally, the division of texts into groups based on bias ratings and the exclusion of the middle ground groups has created a different organization of these texts. This section will focus on detailing the characteristics of the texts in each of the bias groups of interest (i.e., extreme extent, no extent, left leaning, and right leaning), so that the analyses comparing these groups in chapters 8 and 9 are clear. Specifically, this section will detail the division of article type, publication, and topic for each group as well as the number of texts that overlap between the extent and direction groups.

7.3.1 Overlap in Number of Texts between Groups

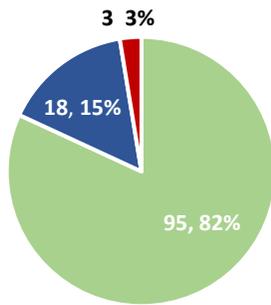
After articles were separated into appropriate groups based on averaged rating scores, there were a total of 192 articles in the extreme extent group, 116 articles in the no extent group, 282 articles in the left leaning group, and 235 texts in the right leaning group (see table 7.1). For a list of the articles included in each group as well as the publication, year, article topic, and article type of each text see appendices B - E.

As was previously stated, because an article received a score for direction of bias and extent of bias, several articles ended up in a direction group and an extent group. Thus, there are articles from the right and left leaning groups that are also in the extreme extent and no extent

groups. It is important to repeat here that direction groups and extent groups are completely distinct from one another as these are the groupings compared in the analyses. In other words, the overlap that occurs does not directly impact the analyses in the present study; however, the researcher felt it important to discuss the overlap so that group descriptions and later evaluations are as clear as possible. Of the texts in these groups, 507 appear in one single group and 318 appear in a direction and extent group meaning that 666 different texts from *SPIN* are included in this analysis (about 65% of the 1,040). Table 7.2 below shows the number of texts in each group as well as the number of unique texts and the number of overlapping texts in each group while figures 7.3 – 7.6 visually illustrate the amount of overlap in each group.

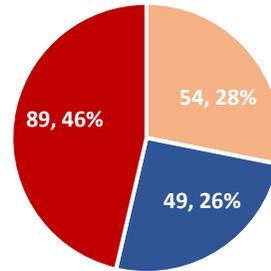
	<u>No Extent</u>	<u>Extreme Extent</u>	<u>Left Leaning</u>	<u>Right Leaning</u>
Number of unique texts	95	54	215	143
Number of overlapping texts	21	138	67	92
Total Number of Texts	116	192	282	235

Overlap of Texts for No Extent



- Texts that appear only in this group
- Texts that are also in the left leaning group
- Texts that are also in the right leaning group

Overlap of Texts for Extreme Extent

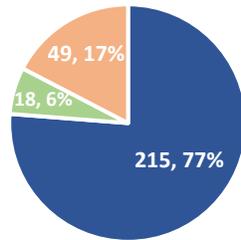


- Texts that appear only in this group
- Texts that are also in the left leaning group
- Texts that are also in the right leaning group

Figure 7.3 *Overlap of Texts for No Extent Group*

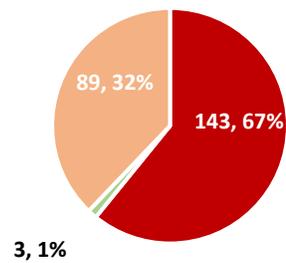
Figure 7.4 *Overlap of Texts for Extreme Extent Group*

Overlap of Texts for Left Leaning



- Texts that appear only in this group
- Texts that are also in the no extent group
- Texts that are also in the extreme extent group

Overlap of Texts for Right Leaning



- Texts that appear only in this group
- Texts that are also in the no extent group
- Texts that are also in the extreme extent group

Figure 7.5 *Overlap of Texts for Left Leaning Group*

Figure 7.6 *Overlap of Texts for Right Leaning Group*

To be clear, figures 7.3 – 7.6 demonstrate where each of the 666 texts were divided by ratings and when there was overlap. As is shown, the no extent group contains 18 texts that are also in the left leaning group and 3 texts that are also in the right leaning group equating to an 18% overlap and leaving 82% as unique texts. The extreme extent group contains 49 of the same texts as the left leaning group and 89 of the same texts as the right leaning group meaning that

72% of this group overlap and only 28% are unique texts. The overlap for the direction groups is such that the left leaning group contains 24% overlap with the extent groups while the right leaning group has 33% overlap. This means that all groups contain mostly unique texts except for the extreme extent group which contains a large amount of overlap. This will be considered in the analysis portion of this paper, but it is important to note that this suggests that articles marked as left or right are not necessarily also considered as extreme in their bias, but that the majority of extreme extent articles are also marked as some kind of directional bias. This adds strength to the idea that the definition of bias needs to be split into multiple constructs in research.

7.3.2 Bias group and Article Type

The next characteristic of these bias groupings that will be discussed is article type (e.g., political, opinion, economy, etc.). Article type is an important characteristic of the domain of newspaper as it informs of the style and of content to expect (original research, review, communication, etc.). This breakdown is intended to provide an illustration of how each bias grouping is divided by article type to highlight how article type might contribute to bias perceptions. In addition, it is going to be important to consider for the analysis portion of this dissertation as some article types, like opinions and politics, might lend themselves well to bias rating. Thus, understanding which article types were more likely to be perceived as biased as well as how article type is balanced within each comparative set of bias groups will provide valuable insight into the data that will be useful for the comparative linguistic analyses in chapters 8 and 9.

For the division of article type, 11 categories were relied upon. Article type was originally collected from the article's publication website and recorded in the tables presented in appendices B - E; however, the 15 publications naturally have differing ways of referring to similar article types. For example, to refer to U.S. news topics, the terms *America*, *US*, *The Nation*, *etc.* were used. In order to streamline the breakdown of groups by article type, the 11 categories used here were the best approximation of a publication's tagging. *Economy* includes any article that was labeled as business, economy, marketplace, or similar. *Entertainment* includes anything labeled as media, TV, Radio, reviews, movies, or similar. *Health, Tech & Science* are all fairly straightforward and include any article labeled as such. *Life* includes articles labeled as Black Voices, Queer voices, culture, religion, women, weddings, life, or similar. Most *opinion* articles were clearly labeled as such, but this category also includes articles labeled as contributed, op-ed, and editorials. This decision was made as an editorial is often allowed to reflect the opinion of newspaper editors and is thus often written in a different manner than 'journalistic' pieces. *Politics* includes any articles labeled as political, the right wing, the left wing, White House, election news, or similar. *Sports* articles were mostly labeled as such but also include articles tagged as events, Olympics, football, or similar. *U.S.* is the most widely varying category and includes articles labeled as US news, America, news, newsfront, hot topics, rights, social justice, the nation, or similar. And last but not least, *world* incorporates articles labeled as global, international, world news, or similar.

To provide an understanding of how article type is dispersed generally in this entire subsample, table 7.3 below shows the number of texts per each category excluding overlap (i.e., across just the 666 texts from *SPIN*), and considering overlap (all texts rated as directional and extent of bias). Both of these divisions are important as excluding overlap details article type for

all of the 666 texts used from *SPIN* and considering overlap provides insight into the group divisions used in the analysis.

Table 7.3
Number of Texts per Article Type for the Entire Subsample

	<u># of texts in a direction and extent group</u>	<u>Total # of texts excluding overlap</u>
Politics	60	276
U.S.	30	89
Entertainment	21	81
Opinion	13	52
World	8	39
Economy	6	37
Sports	3	33
Life	7	29
Environment	5	17
Tech & Science	2	10
Health	4	3
Total	159	666

Table 7.3 is organized in order of the most occurring article type to the least. As is shown, politics is by far the most occurring text type in this subsample with more than triple the occurrence of any other category. It is also likely that political articles are the most common article type in all of *SPIN* suggesting that 1) the top trending topics for each of the examined years might have been written through a political lens for many publications, 2) topics that were political in nature were most likely to be published about (discussed further under gatekeeping bias in chapter 10), and 3) news outlets may publish more political articles than any other category thereby insinuating that political topics are more ‘newsworthy’. Additionally, for this subsample, it suggests that articles that are political in nature are more likely to be perceived as

bias than other article types. The next most common article types were US, entertainment, and opinion. Figure 7.7 below illustrates how these 11 article types are distributed within each bias group, so each group totals to 100%.

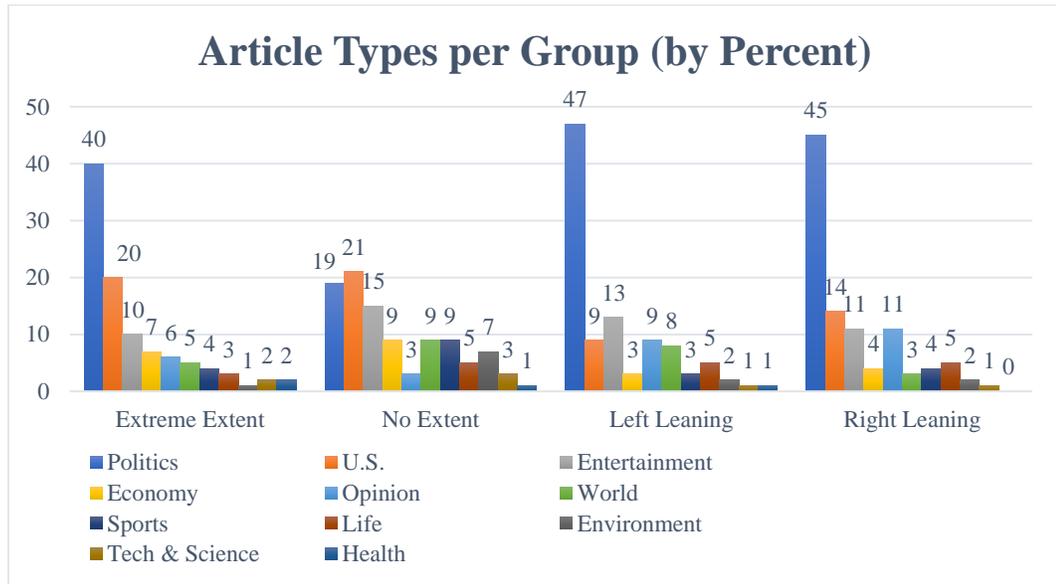


Figure 7.7 Illustration of Percentage of Article Type per Bias Group

As figure 7.7 shows, the group profiles by article type follow the trend of the overall division of articles in the bias subsample but vary widely for the extent groups and are fairly similar for the direction groups. The left leaning group and right leaning group contain nearly even divisions of article type by percentage with politics equating to ~45% of all of the articles in each group, and U.S. and entertainment being second and third most frequent while environment, tech & science, and health are among the least occurring article types in the direction groups. The no extent group and the extreme extent group are less evenly split with the no extent group containing less articles of opinion and politics and the extreme extent group containing less articles of entertainment, environment, sports, and world. The difference in the no extent and extreme extent groups is likely also due to the fact that certain article types lend themselves more easily to perceptions of bias.

7.3.3 Bias Groups and Publication

The next breakdown of the groups to be considered is publication. The original texts from *SPIN* were collected from 15 publications that evenly spanned the left to right political spectrum. As a reminder, *AlterNet*, *The Nation*, and *HuffPost* are the left-most rated publications, *CNN*, *Atlantic*, and *NYT* are rated as left center, *USA Today*, *NPR*, and *AP* are rated as center, *WSJ*, *Fox*, and *TWT* are right center, and *Breitbart (Bb)*, *NewsMax*, and *National Review* are the right-most rated publications (see table 4.1). As above, I will first provide an overview of how all of the texts from this subsample are divided by presenting the total number of texts from each publication (including and excluding overlap) in table 7.4 below. Also provided in table 7.4 is the number of texts from each publication in *SPIN*, the number of texts from this subsample that occur in both a direction and extent group per publication, and the total number of texts per publication excluding overlap (just the 666 texts from *SPIN*).

Table 7.4
Number of Texts per Publication for the Entire Subsample

	<u>Publication</u>	<u># of Texts in <i>SPIN</i></u>	<u># of texts in a direction and extent group</u>	<u>Total # of texts excluding overlap</u>
Left	AlterNet	67	15	58
	The Nation	48	6	47
	HuffPost	75	13	56
Left Center	CNN	74	8	46
	The Atlantic	68	7	38
	The New York Times	75	5	47
Center	USA Today	73	4	37
	NPR	67	8	35
	Associated Press	72	4	15
Right Center	The Wall Street Journal	69	3	40
	Fox	74	8	41
	The Washington Times	72	11	52
	Right Breitbart	72	32	57

	NewsMax	71	17	49
	National Review	63	18	48
Total		1,040	159	666

The publications that have the most articles rated to an extent or direction of bias include *Breitbart* with 89 texts, *AlterNet* with 73 texts, *HuffPost* with 69 texts, *Newsmax* and *National Review* both with 66 texts, and *The Washington Times* with 63 texts while the publication with the least contributing texts is *AP* with only 19 texts; the rest of the publications contribute evenly to the texts in this subsample with between 40-50 texts each. The division of these publication by bias extent and direction is presented below, but from table 7.3, we can already see that publications considered to be the most extreme on the political spectrum contributed the most texts to the bias groupings in this study. Figure 7.8 below illustrates the number of texts contributed from each publication to a bias group.

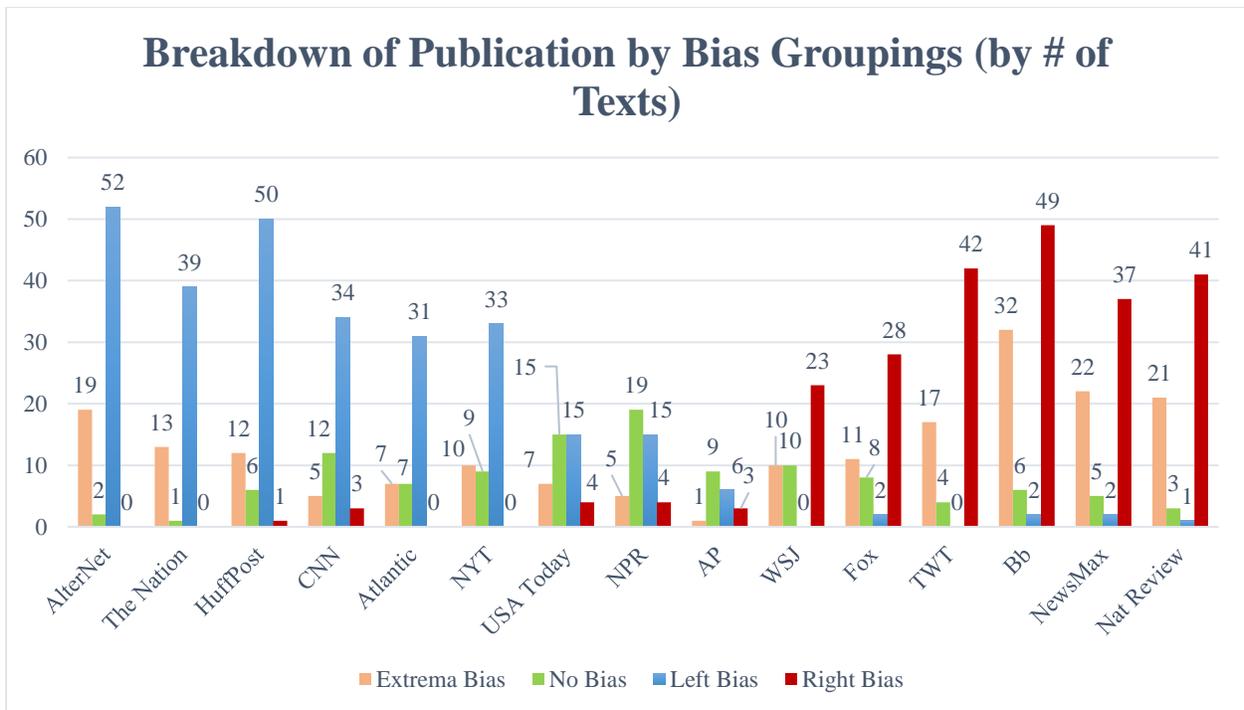


Figure 7.8 Breakdown of Publication by Bias Groupings

Unsurprisingly, the publications that are completely left or completely right have very few or no articles rated in their opposing direction group and have the highest number of texts in the extreme extent group. Also unsurprising is the fact that the publications with the most articles in the no extent group are considered center. This suggests that the perceptions of the raters in the present study align well with the perceptions of the raters relied upon for *AllSides*, *Ad Fontes Media*, and *Media Bias Fact Check*, at least in terms of direction. In addition, because the combined number of raters is quite high, this suggests some universal consensus on the meaning of directional bias that can be further examined and revised in future bias studies.

In addition to how each publication is divided by a bias group, it is important to consider how each group is divided by publication. Figure 7.9 below provides an illustration of bias group broken down by publication. This chart can be understood as how each publication contributes to the overall percentage of texts in each group. So, for example, *Breitbart* makes up 16% of the texts in the extreme extent group and 5% of the no extent group. It is also important to note that figure 7.9 is colored in such a way that the publications rated right, or right-center are shades of red, publications labeled as left or left-center are shades of blue and publications rated as center are shades of green. This was done so that a clear visual of combined left, right, and center could be presented. In addition, each grouping totals to 100% to provide an idea of how each bias grouping is divided by publication.

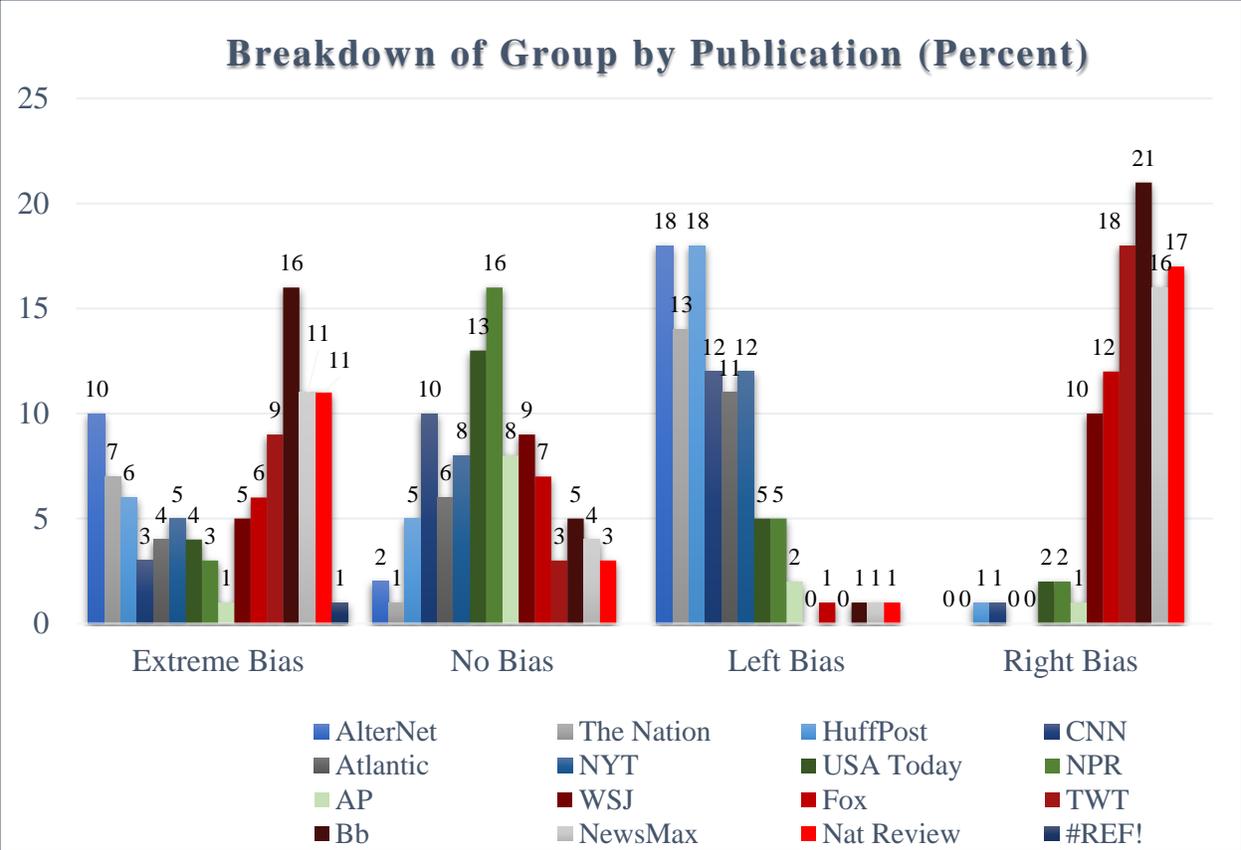


Figure 7.9 Illustration of Breakdown of Groups by Publication

As figure 7.9 shows, the right leaning group is mostly made up of right and right-center publications with the majority of texts coming from *Breitbart*, *The Washington Times*, *National Review*, and *NewsMax*. The right leaning group takes 94% of its texts from right and right-center publications, 5% of its texts from center publications, and only 1% from left and left-center publications. The left leaning group, on the other hand, consists mostly of texts from publications that are left or left-center with *AlterNet*, *The Nation*, *HuffPost*, *CNN*, *Atlantic*, and *NYT* contributing the highest percentage of texts. 84% of the texts in the left leaning group come from left and left-center publications, 12% from center publications, and 4% from right and right-center publications. The no extent group shows a fairly even split between left, right, and center rated publications with the highest text contributions coming from *NPR*, *USA Today*, and *CNN*. The no extent group takes 37% of its texts from center publications, 31% from right and right-

center publications, and 32% left and left-center publications. The extreme extent group is made up of mostly left and right publications with right or right-center contributing over half the texts (57%), left and left-center contributing about 35% of texts, and center rated publications contributing only 8% of the texts.

From this breakdown we can conclude that the most biased publications (based on ratings in this study) are *Breitbart*, *Newsmax*, and *National Review* which are marked as bias to the right and to the extreme, and *AlterNet* and *HuffPost* which are marked as bias to the left and to the extreme. *Associated Press* is the least marked for bias with only 19 texts contributing to the bias groupings and 47% of those ending up in the no extent group. We can also conclude that this aligns heavily with the previous determinations of the publication's leanings which, again, suggests directional bias is pretty well agreed upon.

7.3.4 Bias Groups and Topic

The last division of groups considered here, presented in figures 7.10 – 7.12 below, is by topic. This is to illustrate whether any topic lends itself to direction or extent of bias more than another and, importantly, to note whether any topics occur significantly more in one group than another. A text on each topic was collected up to 15 times (one per publication) in the original collection of texts for *SPIN* meaning that of the collected 75 topics, several occur repeatedly. In this subsample, while there are no stark differences between groups, 75% or more of the articles on the following topics were rated as left or right: 2018 midterm elections, Betsy DeVos's appointment as Secretary of Education, the death of George H. Bush, the *Fire and Fury* book release, Hillary Clinton's emails, Robert Mueller's firing, and Trump's tax release.

Interestingly, these were not always also rated as extreme in their bias which again suggests that a clear left or right bias does not equate to extreme bias in the minds of the raters. In further support of this, several topics contained 0 articles rated for bias extent but several rated for bias direction. Namely, the topics of FIFA World Cup, Kanye West on politics, Mike Pence's visit to the Middle East, and Serena Williams and sexism in tennis contained 0 articles rated as extreme bias or no bias but several rated as left bias or right bias.

The topics with the highest occurrence of texts in the no bias group were Aretha Franklin's death and the Winter Olympics while the topics with the highest occurrence in the extreme bias group were Trump's sexual assault charges, taking a knee, and the #MeToo movement. The highest occurrences in the left bias group are Alec Baldwin on SNL and Fire and Fury book release while the highest occurrences on the right were Betsy DeVos's appointment as Secretary of Education and the death of George Bush; however, these texts do not necessarily occur at differing rates between groups. Below, a comparison of topics and how they contribute to each bias group is presented.

Because there are 75 topics included in this analysis, illustrating how they each contribute to every group is difficult. Therefore, charts representing divisions of bias groups by topic were created for each year of interest – this way, the division of topic can be examined as three manageable chunks. Each difference in topic will also be discussed in the prose, and topic differences per bias group are presented at the end of this section in table 7.4. Figure 7.10 illustrates the division of bias groups by the 2016 topics, figure 7.11 illustrates the division of bias groups by the 2017 topics, and figure 7.12 illustrates the division of bias groups by the 2018 topics (for a reminder of all 75 topics, please see table 4.2).

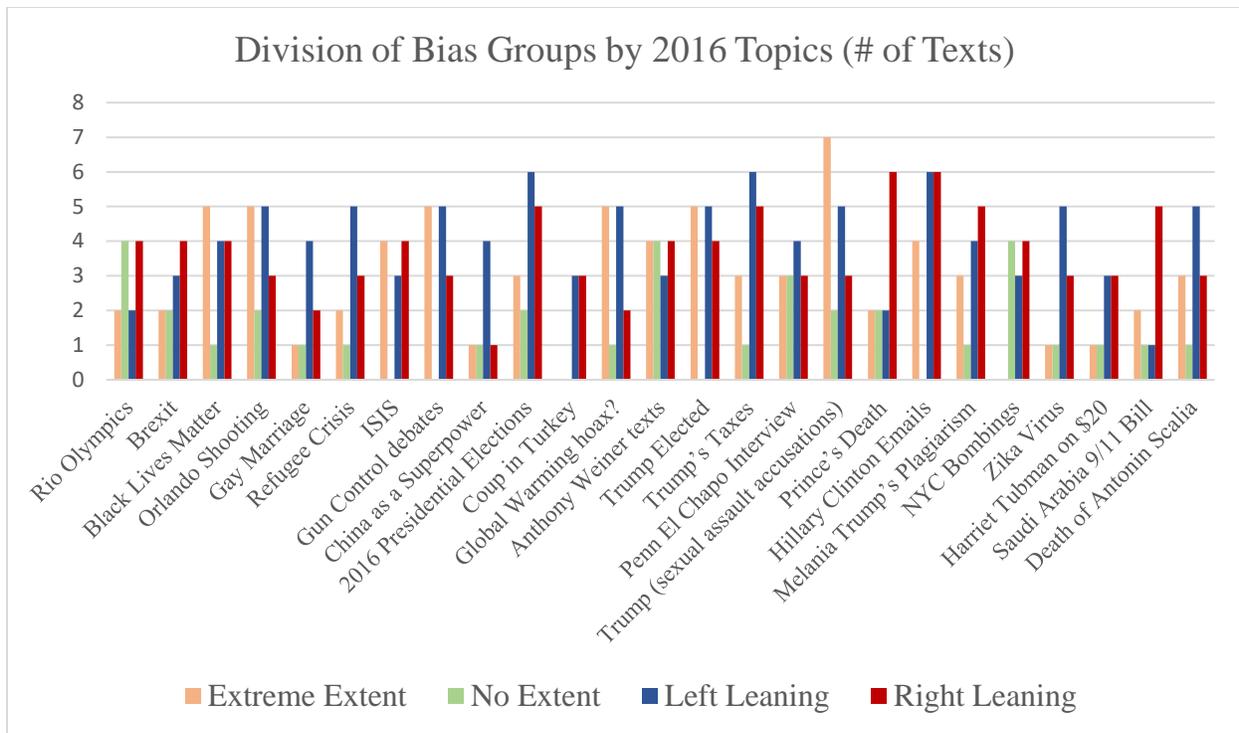


Figure 7.10 *Division of Bias Groups by 2016 Topic*

The highest occurrence in number of texts for any one topic is 7 while the lowest occurrence is 0 – and this is true for all of the divisions by year – thus, differences of 3 or more texts of any one topic will be considered significant, and differences of 2 or fewer texts will be considered as similarly occurring. This is done because it could explain some of the lexical patterns identified in chapter 9.

For the extreme extent and no extent groups there were differences in the following topics from 2016 (note that the number on the left is extreme extent and the left is no extent): Black Lives Matter (5 v 1), the Orlando shooting (5 v 2), ISIS (4 v 0), gun control debates (5 v 0), global warming hoax (5 v 1), Trump elected (5 v 0), Trump sexual assault accusations (7 v 2), Hillary Clinton’s emails (4 v 0), and NYC bombings (0 v 4). For the left leaning and right leaning groups, the following differences in 2016 topics occur: China as a superpower (4 v 1), global warming hoax (5 v 2), Prince’s death (2 v 6), and Saudi Arabia 9/11 bill (1 v 5).

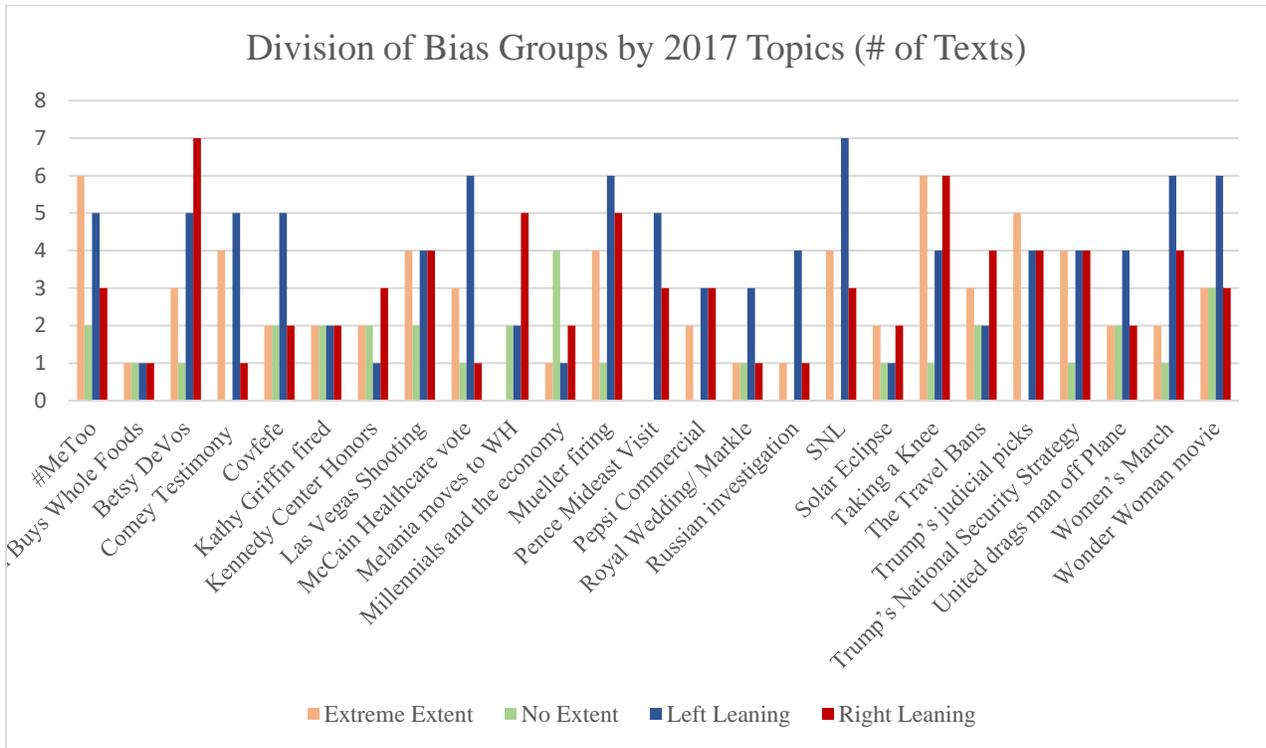


Figure 7.11 *Division of Bias Groups by 2017 Topic*

As the chart above shows, for the 2017 topics, the following differences in extreme extent and no extent were identified: #MeToo (6 v 2), Comey’s testimony (4 v 0), Millennials and the economy (1 v 4), Mueller firing (4 v 1), SNL (4 v 0), Take a knee (6 v 1), Trump’s judicial picks (5 v 0), and Trump’s national security plan (4 v 1). For the left leaning and right leaning groups, the following differences in topics occur: Comey’s testimony (5 v 1), Covfefe (5 v 2), McCain’s healthcare vote (6 v 1), Melania moves to White House (2 v 5), Russia investigation (4 v 1), SNL (7 v 3), and the Wonder Woman movie (6 v 3).

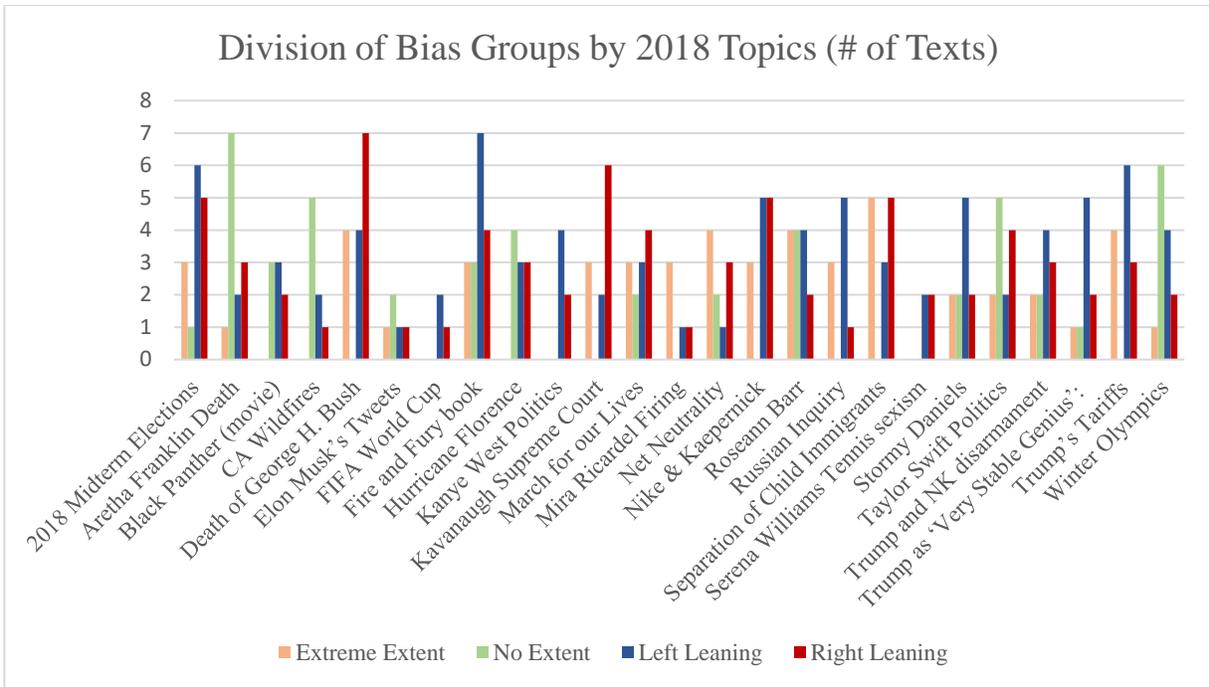


Figure 7.12 *Division of Bias Groups by 2017 Topic*

For the extreme extent and no extent groups presented in figure 7.12, there were differences in the following topics: Aretha Franklin’s death (1 v 7), CA wildfires (0 v 5), Death of George H. Bush (4 v 0), Hurricane Florence (4 v 0), Kavanaugh Supreme Court appointment (3 v 0), Mira Ricardel firing (3 v 0), Nike and Kaepernick (3 v 0), Russia inquiry (3 v 0), separation of child immigrants (5 v 0), Taylor Swift politics (2 v 5), Trump’s tariffs (4 v 0), Winter Olympics (1 v 6). For the left leaning and right leaning groups, there were differences in the following topics from 2018: death of George H. Bush (4 v 7), Fire and Fury book release (7 v 4), Kavanaugh Supreme Court (2 v 6), Russia inquiry (5 v 1), Stormy Daniels (5 v 2), Trump as a ‘very stable genius’ (5 v 2), and Trump’s tariffs (5 v 3). Because the charts are still a bit difficult to follow, table 7.5 below lists the topics per bias grouping that had a significantly higher number of texts.

Table 7.5

A list of Topics with Significantly higher Number of texts per Bias Group

(underline: topics that occur in both extreme extent and a direction)

<u>Bias group</u>	<u>Topics with a significantly higher number of texts</u>
Extreme Extent	<ol style="list-style-type: none">1) Black Lives Matter2) Orlando shooting3) ISIS4) Gun control debates5) <u>Global warming hoax</u>6) Trump elected7) Trump and sexual assault accusations8) Hillary Clinton's emails9) #MeToo10) <u>Comey's testimony</u>11) Mueller's firing, SNL12) Take a knee13) Trump's judicial picks14) Trump's national security plan15) <u>Death of George H. Bush</u>16) <u>Kavanaugh Supreme Court appointment</u>17) Mira Ricardel firing18) Nike and Kaepernick19) <u>Russia inquiry</u>20) Separation of child immigrants21) <u>Trump's tariffs</u>
No Extent	<ol style="list-style-type: none">1) NYC bombings2) Millennials and the economy3) CA wildfires4) Hurricane Florence5) Taylor Swift's politics6) Winter Olympics
Left Leaning	<ol style="list-style-type: none">1) <u>Comey's testimony</u>2) Covfefe3) McCain healthcare vote4) Russia investigation5) Wonder Woman movie6) China as a superpower7) <u>global warming hoax</u>8) Fire and Fury book release9) <u>Russia inquiry</u>10) Stormy Daniels11) Trump as a 'very stable genius',12) <u>Trump's tariffs</u>
Right Leaning	<ol style="list-style-type: none">1) Prince's death2) Saudi Arabia 9/11 bill3) Melania moves to WH

-
- 4) Death of George H. Bush
 - 5) Kavanaugh's appointment to Supreme Court
-

As the table shows, the extreme extent group has 21 topics occurring at a higher rate than the no extent group and the no extent group has 6 topics occurring at a higher rate than the extreme extent group which means that 48 topics occur at a similar rate between the two groups. The left leaning group has 12 topics that occur at a higher rate than the right leaning group and the right leaning group has 5 topics that occur at a higher rate than the left leaning group leaving 58 topics that occur at a similar rate between the two groups. Although the majority of the topics occur at a similar rate between the groups which will serve as the points of comparison, understanding the topics that occur at different rates could be important in the linguistic analysis – particularly the keyword analysis – as these differences could explain some of the identified variation. Thus, these differences will be discussed again in chapter 9.

7.4 Conclusion

The purpose of this chapter has been to describe the methods for using bias ratings to separate texts into bias groups (both direction and extent) as well as to explain the reason behind the exclusion of the middle ground groups and discuss the characteristics of the groups of interest so that the analyses in the next two chapters might make more sense. To summarize, the left and right leaning groups are similar in division of article type and contain only 17 topics that occur at different rates but contain differing divisions of publication – the divisions align heavily with left and right bias. This suggests that a comparison of these two groups can be considered fair since they are quite similar in their makeup. The extreme extent and no extent groups vary widely in article type, contain 26 topics occurring at different rates, and contain differing

divisions of publication. Like with the direction groups, the division of publication align heavily with a perspective of bias. The extent groups are thus much less evenly comparable; however, the differences in these groups likely relate to the fact that these were rated on a scale from no bias to extreme bias so these differences are likely due to the fact that these extremes of bias would present themselves differently.

CHAPTER 8. KEY FEATURE ANALYSIS

8.1 Introduction

Now that the background of the groups of comparison have been presented, the next two chapters will present the methods, results, and analysis of the linguistic examinations. To examine the language features of the texts, methodologies similar to that of Biber & Egbert (2018) were employed; specifically, a keyword and a key feature analysis were done. Although these methods were used by Biber and Egbert (2018) to examine register variation these methods might also be useful for highlighting variation within a register. The combination of these methods was relied upon to comprehensively examine the linguistic variation across texts of different political leanings and extent of bias as each offers different, but complementary, quantitative results. The goal of this chapter is to present the key feature analysis while chapter 9 presents the keyword analysis.

8.2 Grammatical Feature Tagging and Selection

Because there are very few studies on grammatical features related specifically to bias, pre-selecting features would have involved many subjective choices by the author. This, in conjunction with the investigative nature of the research question, led to the key feature analysis being run in a bottom-up fashion on a wide range of lexico-grammatical features.

All features were identified by the Biber tagger which was first developed in 1984 by Doug Biber to identify a broad set of linguistic features in large data sets (Goulart & Wood, 2021). The current version has been developed and revised numerous times, draws from multiple noteworthy dictionaries, and relies on “both probabilistic and rule-based components” (Biber & Egbert, 2018, p. 22). In addition, the Biber tagger analyzes a larger set of linguistic features than other taggers and has been shown to have comparable, if not better, accuracy than other available taggers (Biber & Egbert, 2018). Thus, *SPIN* was tagged with the Biber tagger and tag-checked and fixed using The Biber Tag Checking Tool demonstrating a .95 precision and recall (developed by the CROW Team: <https://writecrow.org/>).

After each text in *SPIN* was tagged, checked, and fixed, a program called TagCount, also developed by Biber, was used to calculate normed rates of occurrence (per 1,000 words) on all features identified by the Biber tagger. Although there is considerable overlap in some of the tagged features (e.g., short passives and all passives), only features with low frequencies were removed as the Cohen’s *d* formula analyses each feature independently. A total of 122 lexico-grammatical features were included in the key feature analysis.

8.3 Methods of Key Feature Analysis

“Keyness” is typically a term used to describe a word that occurs at a statistically higher or lower frequency in one context over another; however, Biber and Egbert (2018) apply keyness to lexico-grammatical features and demonstrate that it is an effective means of identifying linguistic variation across registers. Further studies have demonstrated that it is useful not only as a means of identifying variation across registers but can also highlight nuanced differences within registers (Biber et al, 2021; Goulart, 2022). Because the goal here is to examine linguistic

variation across bias groups within the register of news, this method is highly appropriate; thus, methods similar to those of Biber and Egbert (2018) will be employed. Due to the high frequencies typical of grammatical features, Cohen’s *d* was relied upon to obtain keyness values. The same methods were applied in the present study and keyness was measured using Cohen’s *d* formula:

$$d = \frac{M_1 - M_2}{SD_{pooled}}$$

where:

$$SD_{Pooled} = \sqrt{\frac{SD_1^2 + SD_2^2}{2}}$$

As the formula shows, the mean rate of occurrence for a grammatical feature in the target group (*M2*) is subtracted from the mean rate of occurrence for the same grammatical feature in the reference group (*M1*). The result is then divided by the pooled standard deviation (*SDpooled*), which gives us a standardized effect size. Based on Cohen’s interpretation of sample size (table 8.1), features with scores <.2 were removed (Cohen, 1988; 2013).

Table 8.1
Sample size interpretations

<u><i>d</i></u>	<u>size</u>
>+/- .80	Large
>+/- .50	Medium
>+/- .20	Small

For the purposes of the present study, the key feature analysis was run comparatively meaning that rather than a target-reference comparison, a target-target comparison was done. Thus, features that are key in the positive and negative direction were interpreted with the positive features indicating which features are used more in one group and negative features indicating which features are used more in the opposite group. Specifically, left-rated texts are compared to right-rated texts and extreme extent texts are compared to no extent texts. Features with large positive and negative d values indicate a feature is used much more in the respective group (Biber & Egbert, 2018). In the present study, all features with d values of $\pm .20$ are interpreted and considered valuable. The reason for this is that this is an intra-register comparison, and the differences are not expected to be large since communicative purpose and context are similar. Additionally, features of media bias will likely be subtle, implicit, and intricate since, in general, the news is expected to be factual and impartial (Baumer et al., 2015; Recasens et al., 2013). Instead, the key feature analysis will allow for an understanding of nuanced differences across direction and extent of bias and can detail the distinctive linguistic and functional characteristics of each. The results are presented below.

8.4 Key Features: Left V Right

The results of the key feature analysis revealed that right-rated articles are marked by four grammatical structures while texts rated as left are marked by five grammatical structures. Due to the comparative nature of the study, positive key features represent those that are marked for right-rated texts while negative key features represent features marked for left-rated texts. Figure 8.1 illustrates the nine key features that met the inclusion criteria of $d \geq \pm .20$. The inclusion criteria might seem low, however, given the similarities in situational characteristics of

the texts included in the study, we can reasonably expect no effect, so even a small effect size was interpreted as relevant here.

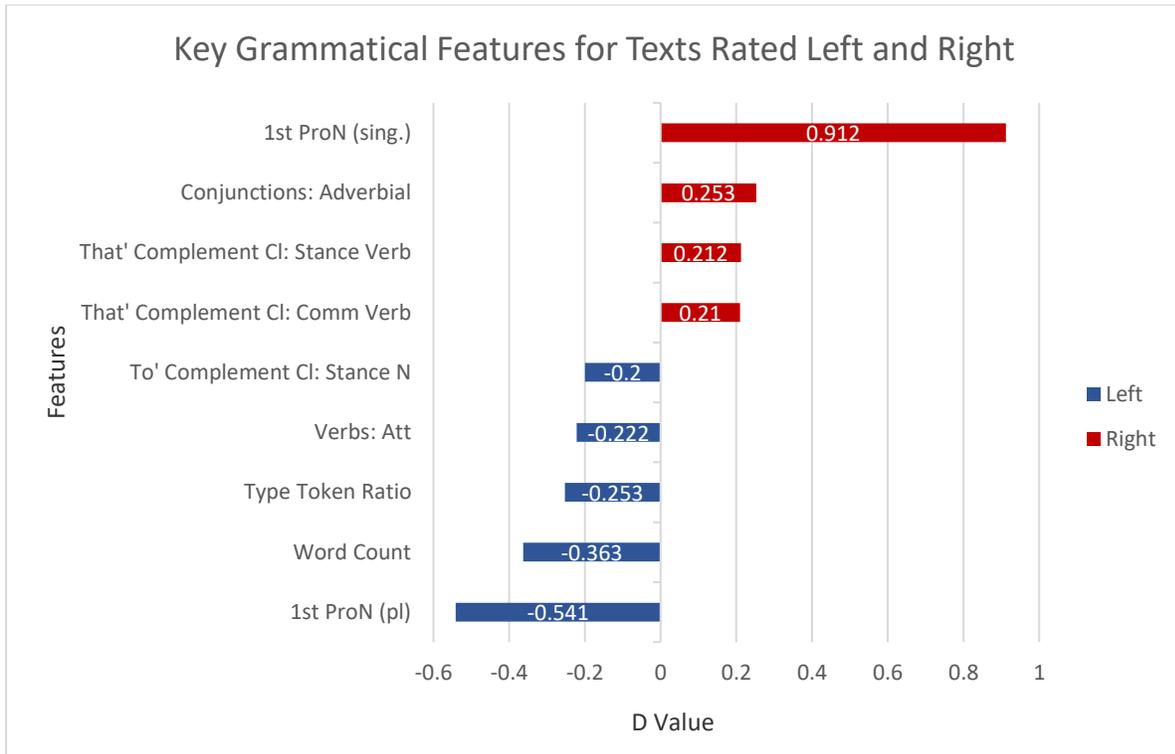


Figure 8.1 Key grammatical features for texts rated left and right.

8.4.1 Right-Rated Articles

Articles rated as right rely heavily on singular first-person pronouns (I, me, my, mine, and myself). Initially, this was quite surprising as the majority of register studies on news demonstrate that personal pronouns typically have low frequencies due to the lack of interaction and involvement in News Reports (see Egbert and Biber, 2018, p. 84). In linguistic studies of fake news, pronouns, particularly first- and second- person, are strong markers of real versus fake news (Rashkin et al., 2017); however, conflicting findings of the first-person pronoun *I* were found by Grieve & Woodfield (2023) who found it to be a marker of fake news while

Taboada (2021) found it to be a feature of real news. Greive & Woodfield (2023) conclude that the use of the first-person pronoun was a reflection of fake news' tendency to be more involved than traditional news discourse while Taboada (2021) concluded that the marked first-person pronoun was a reflection of modern news's move towards informalization. So, while the feature was marked for opposing reasons in these two studies, the authors agree in their conclusions that the use of the first-person pronoun signals discourse that is more involved than informational and that this feature allows the perspective of the writer to be more prominently displayed.

Interestingly, here, right leaning articles rely on singular first-person pronouns while left-rated articles rely on plural first-person pronouns (discussed below). This highlights the importance of using analyses such as these to examine variance within registers as well as across registers as the nuanced use of features can be understood. In this analysis in particular, we see that interaction and involvement via personal pronouns is expressed in news and that it is expressed differently in articles rated as right and left; the first being involved at the individual level and the second being involved at the group level.

Initially, it was thought that the difference in pronoun use in right-rated articles may have been due to a higher number of opinion pieces in the right-rated group versus the left-rated group; however, a closer examination revealed that of the 235 articles rated as right, 23 were opinion articles while of the 282 rated as left, 26 were opinion articles (see appendices D & E). A closer examination of the texts suggests that singular first-person pronouns are key in right-rated articles for two reasons: (1) a heavy reliance on quotes particularly in situations where people are on the defense (see [1] & [2] below); (2) In many of the publications rated as right, there is an almost bloglike personal style of writing (see examples [3] & [4] below) which suggests that this shift towards informalization of news discussed by Taboada (2021) is not only reflected here, but

also that it is associated with directional bias. Interestingly, not all of the articles rated right with this bloglike style are from right leaning publications; however, raters seem to associate this personal style with the right and the more group-oriented articles with the left.

[1] “**I** wrote them down and later included some of the phrasing in the draft that ultimately became the final speech,” she said. “**I** did not check Mrs. Obama’s speeches. This was my mistake, and **I** feel terrible for the chaos **I** have caused Melania and the Trumps, as well as to Mrs. Obama. No harm was meant.” (*Fox*, Melania Trump plagiarism)

[2] “Of course **I** support women,” she told Connell McShane on Tuesday. “And **I** want violence to end against women. **I’ve** been very active in abuse shelters and child advocacy centers. **I’ve** been advocating for women and equal pay since **I** was 19-years-old and making certain that women have the opportunity for maximum pay and have a good record on that.” (*Newsmax*, Taylor Swift politics)

[3] But apparently – so **I** have since learned – the march was about none of these things. According to Madonna it had something to do with blowing up the White House. According to Michael Moore (not a woman) it was about stopping Betty De Vos (a woman: how does this work, exactly?) from becoming Secretary of Education. (*Breitbart*, Women’s March)

[4] At the same time, **I** applaud the president for holding firm on his “zero tolerance” policy and criminally prosecuting any immigrants who come to the U.S. in violation of our laws. (*Fox*, Child immigration)

Examples [1] and [2] demonstrate some of the quoted material that is common in right-rated texts in which the speaker is personally involved in a way that signifies self-defense while in [3] and [4], the authors are personally involving themselves to present their own perspective. Left-rated texts also rely on the first-person pronoun, but rather than the singular *I*, the plural *we* is used much more suggesting that while right-rated news articles are involved at the personal level, left-rated articles are involved at the group level. This seems to occur for three reasons: 1) as a way to place responsibility for something on an entire group as in [5], 2) a group call to action as in [6], and 3) as a presentation of accepted truths rather than supported facts as in [7]. Interestingly, in discourse of political speech ‘I’ and ‘we’ were noted as constructs used by politicians for self-protection and projected group responsibility which are two of the functions

occurring here with self-protection being a function of right leaning texts and projected group responsibility being a function of left leaning texts.

[5] Now **we** need to accept that our long and nefarious involvement in Central America has helped spawn the present crisis and, accordingly, that **we** are responsible for extending a helping hand to the women and children who are its victims. (*The Nation*, refugee crisis)

[6] This was a slaughter of LGBTQIA folks, many of them Latinos and people of color, during Pride month. **We** cannot, and should not, hide from these facts. (*NPR*, Orlando shooting)

[7] It was the half of America, a half that if not bigoted itself seemed mighty fine with being bigotry-adjacent. This is who **we** are. (*The Atlantic*, Trump elected)

Adverbial conjunctions (e.g., however, although) are also relied on more in right-rated articles which suggests a difference in clausal elaboration than left-rated articles (see [8], [9], & [10]). This is likely a result of the difference in tone that appears to exist between the right-rated articles and the left-rated articles with the right maintaining a more colloquial tone and the left maintaining a more formal tone. This is discussed further at the end of this section.

[8] “Swift seems to accept the headlines and marketing of the Left that would have you believe that a vote against anything claiming to be pro-woman or pro-equality is actually so; **However**, voters in Tennessee are likely to be more savvy.” (*National Review*, Taylor Swift politics)

[9] Your ability to watch and use your favorite apps and services could start to change - **though** not right away - following the official demise Monday of Obama-era internet protections. (*The Washington Times*, net neutrality)

[10] A wider pullback from emerging markets --with many currencies sliding against the dollar late Friday as the Turkish lira sank 4.3% after news of the coup emerged--isn't necessarily justified, **however**. (*The Wall Street Journal*, Turkey coup)

The last features that came out as key on the right were ‘that’ complement clauses controlled by verbs of communication (e.g., say, announce, state) (see [11]-[13]) and ‘that’ complement clauses controlled by stance verbs (see [14]-[16]). These are lumped together here, because while

communication verbs facilitate direct speech reporting, they can also indicate a type of stance that allows an author to present an assertion as fact without accepting responsibility for the claim (Biber, 2004). Thus, we can see that stance expression for articles rated as right are ‘that’ complement clauses controlled by different types of stance verbs.

[11] Trump's attorneys have **said that** they will attempt to compel arbitration in the case, prompting Avenatti to demand a jury trial. The White House has **said that** Trump denies having an affair. (*Newsmax*, Stormy Daniels allegations)

[12] When it comes to giving up their iPhone or other communication devices or spending less time with family, millennials were less interested in surrendering those things. The survey **said that** only 16 percent of millennials were willing to surrender their cellphones for their hobbies, and only 17 percent would want to spend less time with family. (*Newsmax*, Millennials and the economy)

[13] "I think that it's very unfair in light of what's happened with other people," said Trump, who **stated that** he wouldn't even consider a pardon if not for the outcome of the Clinton investigation. (*Fox*, Baldwin as Trump on SNL)

[14] That his crowds remain huge and his following loyal, and that he remains competitive, **suggests that** the depth of the detestation of our cultural, political and media elites out there in Middle America is deep. (*Newsmax*, Trump's taxes)

[15] "Look, I spoke with people who were in this world that we're talking about," Trump told **. "They said, 'Absolutely, it works.' Now, General Mattis said that he doesn't intend to use it. I'm with him all the way. Do I **believe that** it works? Yes, I do." (*Fox*, Baldwin as Trump on SNL)

[16] He added, "I do **believe that** we're going to have a much better plan, we're going to have a cheaper plan. I **think** it's going to be a lot less expensive." (*Breitbart*, 2018 Midterm Elections)

Taken together, the features for right-rated texts demonstrate personal involvement through the strong use of the first-person singular pronoun that serves the function of self-protection as well as insertion of author opinion, stance that is expressed through verbal communication, and a reliance on adverbial coordination which are features that indicate a more colloquial style and might normally be expected in face to face communication or personal blogs as they are associated with relatively casual and spontaneous discourse, and lower levels of informational

density (Bier, 1988; Biber & Egbert, 2018; Grieve & Woodfield, 2023). In addition, these are features that allow for the presentation of perspective in a covert manner which aligns well with Taboada's (2021) findings that news is shifting toward a more conversational discourse pattern that mixes opinion with hard news.

8.4.2 Left-Rated Articles

Articles rated as left rely on plural first-person pronouns (we, us, our, ours, and ourselves) which is in contrast to the right-rated articles that relied on singular first-person pronouns. We can see from examples [17]-[16], left-rated articles seem to consider issues from a group perspective rather than the individual one often presented in right-rated articles. As stated above this appears to happen for three distinct reasons. The first is that there is a tendency in left-rated articles to discuss group responsibility for issues happening in the United States as in examples [17] – [20]. Second, articles in this group tend to make group calls to action presented with a 'we're stronger together' sentiment as in [21] – [25]. And third, left-rated articles display a tendency to make assumptions about their readers and then to subsequently rely on these assumptions to then present statements as accepted truths as in [26] – [30].

1) Group responsibility:

[17] Now **we** need to accept that our long and nefarious involvement in Central America has helped spawn the present crisis and, accordingly, that **we** are responsible for extending a helping hand to the women and children who are its victims. (*The Nation*, refugee crisis)

[18] But, there are warning signs that should prevent the purchase of a gun, and **we** are just not doing it. (*NYT*, gun control)

[19] If the response to it continues to be characterized by hype and inflammatory ignorance, **we** are only encouraging ISIS's efforts to stage attacks like the one in Paris and inspire more massacres like that in Orlando. (*The Atlantic*, ISIS)

[20] **We** thought **we** were stronger and more numerous than **we** are, and that made us insular and arrogant....**We** Americans tend to think **we**'re unique, but the election of Trump is a version of what's happening in many parts of the world. (*The Nation*, 2016 presidential election)

2) Group calls to action:

[21] **We** pride ourselves on acceptance. If **we** are truly that country, then it is our obligation to reunite these detained children with their parents, and to stop separating parents and children in the first place. (*The Nation*, separation of child immigrants)

[22] This was a slaughter of LGBTQIA folks, many of them Latinos and people of color, during Pride month. **We** cannot, and should not, hide from these facts. (*NPR*, Orlando shooting)

[23] **We** absolutely cannot allow this debate to become one of 'unity' vs. 'the flag' or a liberal brand of bumper-sticker patriotism ("Protest Is Patriotic") vs. the Trumpian brand ("Stand or Die!"). (*The Nation*, take a knee)

[24] **We** need to re-orient antitrust policy to factor in the harm that economic concentration causes for American workers. **We** also need to be mindful that concentrated industries stifle innovation. (*HuffPost*, Amazon buys Whole Foods)

[25] Courage is the habit that allows us, even as **we** tremble, to do what **we** know is right and necessary - to refuse to torture anyone, to rescue refugees, to recognize that striking first is not self-defense but aggressive war, to realize that others are profiting by the ways **we** let fear immobilize us. If **we** are to successfully resist the Trump juggernaut, **we** will need to be brave, (*Th Nation*, Trump as a 'very stable genius')

3) Establishing assumptions and accepted truths:

[26] **We** know how to help children fleeing violence in Central America. Instead, **we**'re raiding communities and deporting families. (*The Nation*, Refugee crisis)

[27] "**We** might still be in elementary school, but **we** know," she said. "**We** know life isn't equal for everyone. And **we** know what's right and wrong. **We** also know that **we** stand in the shadow of the Capitol." (*CNN*, March for our Lives)

[28] **We** all know the interview script: the runners-up hiding their disappointment as they talk up their training routines; the medalists dutifully crediting coaches and family. (*The Atlantic*, Winter Olympics)

[29] **We** are indeed in the age of destabilization, a phenomenon that is on the rise globally and at present spreading into the Balkans and Eastern Europe (*The Nation*, ISIS)

[30] It was the half of America, a half that if not bigoted itself seemed mighty fine with being bigotry-adjacent. This is who **we** are. (*The Atlantic*, Trump elected)

The next features that came out as key for left-rated articles are word count and type token ratio.

This indicates that articles rated as left are longer and have a wider vocabulary range than those on the right. These features, taken together, indicate that articles rated as left tend to be more verbose which adds to the formal tone we see appearing in left-rated articles as compared to those rated as right.

The last two features that came out as key on the left, like the last two features of the right, also indicate expressions of stance. In left-rated articles, stance is expressed by attitudinal verbs (e.g., believe, hope) (see [31] & [32]), and ‘to’ complement clauses controlled by stance nouns (fact, assertion, duty, etc.) (see [33] & [34]). These expressions of stance indicate that left-rated articles express stance in a denser fashion than right-rated articles by relying more on non-finite clauses which make the meaning less explicit than the finite clauses depended on in right-rated articles. This fits well with the formal versus colloquial tone that appears in left versus right rated articles, respectively.

[31] West tweeted this week that he realizes “I’ve been used to spread messages I don’t **believe in**.” That follows his Oval Office visit last month during which he dropped the F-word while talking with President Donald Trump. (*Associated Press*, Kanye West for Trump)

[32] “The gun industry doesn’t have much to worry about. I **hope** to be proven wrong.” (*The Nation*, Las Vegas shooting)

[33] Still, there was something exceedingly familiar about Mr. Trump's contention that he had a **duty** to others **to keep** his taxes low. (*The New York Times*, Trump’s taxes)

[34] Trump communication adviser Jason Miller called the article "fiction," saying, "To reach back decades **in an attempt** to smear Mr. Trump trivializes sexual assault, and it sets a

new low for where the media is willing to go in **its attempts to** determine this election. (*USA Today*, sexual assault accusation against Trump)

The key feature analysis comparing right-rated articles to left-rated articles highlights that there are indeed grammatical differences based on direction of political leaning. The differences in the two groups relate directly to reporting style with the right being much more reliant on personal involvement, and a colloquial tone whereas the left is involved at the group level, has a wider vocabulary range, and tends to be a bit more formal in tone. Further, the right displays more verbal communication, shorter prose, and higher reliance on coordination while the left is slightly more nominal, verbose, and expresses stance with non-finite clauses. Interestingly, these differences are like those one would expect on an oral versus literate dimension; the right being more oral and the left being more literate (see Biber 2014 for a description of this universal dimension). This again highlights the importance of exploring variation within register via methods typically relied on to explore variation across register. We can see that patterns of variation appear to emerge even on a nuanced scale which suggests that linguistics studies may benefit from examinations at the text level in addition to the register level (Biber, Egbert, & Keller, 2020).

This analysis also demonstrates that both right-rated and left-rated articles rely on features of stance that express author/speaker attitude in a covert manner. While the features they rely on to achieve this goal differ due to style, both use features that allow the author to avoid accepting responsibility for the expressed sentiment; the right relies on communication verbs and reported speech while the left relies on non-finite clauses that make the statements less explicit. This is not surprising since news is expected to uphold tenets of impartiality and too much overt expression of stance would likely reduce a publication's reputation as a valid source of news.

What is interesting is that despite the attempts to be covert and despite the small effect sizes, it is enough for raters to differentiate between left and right bias.

8.5 Key Feature: No Extent Vs Extreme Extent

The results of the key feature analysis revealed that articles rated as no extent are marked by five grammatical features while texts rated as extreme are marked by sixteen grammatical features. As with the analysis above, both positive and negative features are interpreted with positive key features representing those that are marked for no extent texts and negative key features representing those marked for extreme extent articles. Figure 8.2 illustrates the twenty-one key features that met the inclusion criteria of $d \geq \pm .20$.

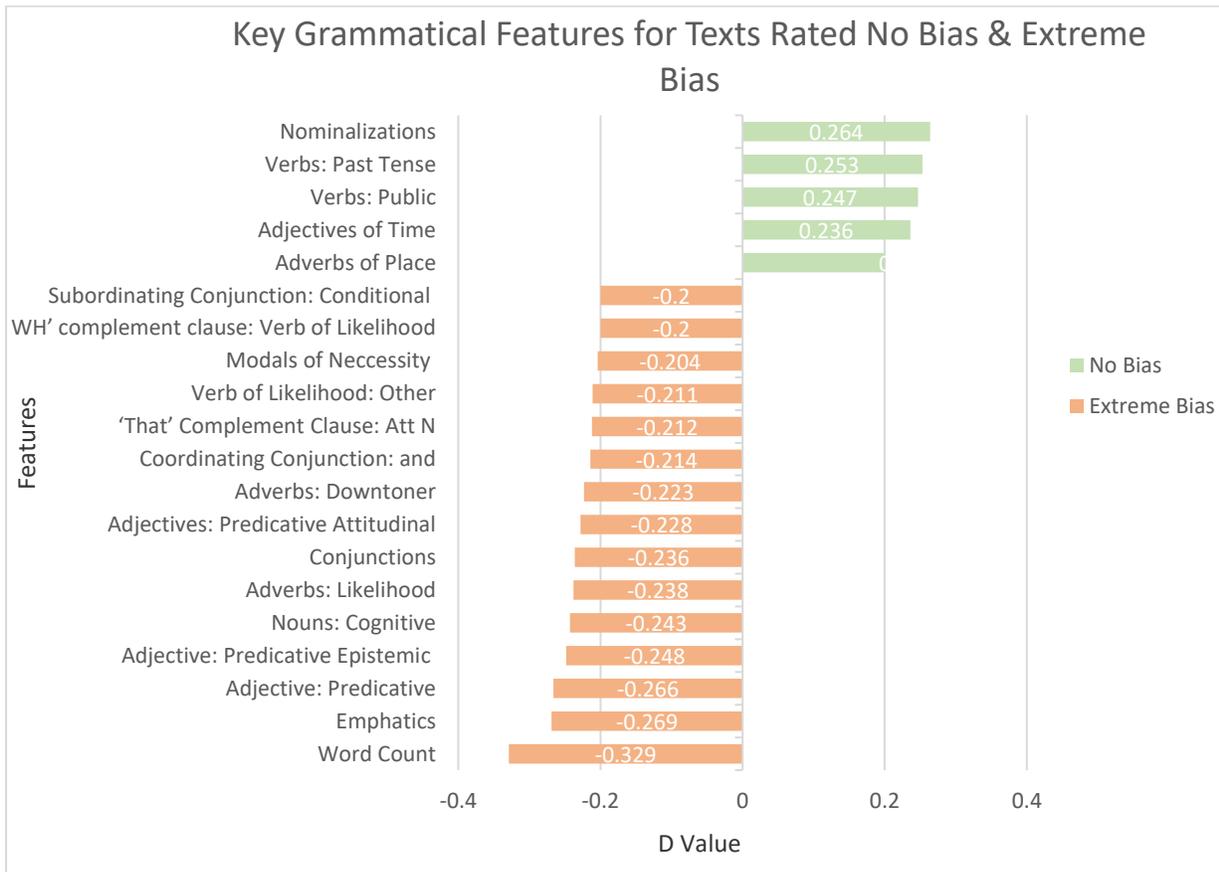


Figure 8.2 Key grammatical features for texts rated extreme bias and no bias.

8.5.1 No Extent Articles

The first feature that emerged as key for articles rated as no bias are nominalizations. Given that nominalizations are typically associated with academic writing as they remove the agent in a sentence and help create an impersonal tone (Baratta, 2010), it is not surprising that this feature is associated with impartiality. In addition, examinations of news have demonstrated that as a register, news tends to exhibit features related to informational density and formality, and interestingly, Grieves & Woodfield (2023) demonstrate that nominalizations (and many other features of related to nouns) were a strong marker of real news. In excerpts [35] & [36], we can see that nominalizations in no extent texts are functioning to create an impartial and impersonal tone.

[35] A new set of rules at the Federal **Communications** Commission went into effect Monday, ending the previous consumer **protections** that prevented Internet service providers from blocking or slowing legal traffic, or charging for faster delivery of some content. (*USA Today*, net neutrality)

[36] Fast Company attributes the economic concerns and **disappointments** young Americans feel to a **combination** of factors, including just starting to seek **employment** concurrently or shortly after the Great **Recession**. (*NPR*, Millennials and the economy)

The next key feature marked in texts rated as no extent are past tense verbs which are used in news reports to describe events or actions that have already happened (see [37]). Past tense verbs are also typically associated with a narrative function that is used in conjunction with public verbs to report statements from public figures or organizations as in [38].

[37] Florence **claimed** five lives, **caused** widespread flooding and **knocked** out power to nearly 900,000 homes Friday. (*USA Today*, Hurricane Florence)

[38] Police **recovered** 23 guns from his Las Vegas hotel room and another 19 guns from Paddock's home in Mesquite, Nevada, Clark County Assistant Sheriff Todd Fasulo **said**. (*CNN*, Las Vegas shooting)

Unsurprisingly, public verbs also emerged as key for no extent texts. Public verbs imply the idea of speaking and are used in news reports to introduce indirect statements as in [39] (Biber, 2014). Like past tense verbs, public verbs are also associated with a narrative function as they are syntactically simple and are used to narrate speech acts that have or will occur as in [40] (Hinkel, 2012).

[39] The high-profile incidents, advocates **say**, are empowering others to **report** their own stories. (*NPR*, #Me Too)

[40] Hong said Pyongyang is expected to **tell** Pompeo what disarmament steps it will take over the next six to 12 months (*Associated Press*, Trump and North Korean disarmament)

Adjectives of time also demonstrated a small effect size in the no extent group. Adjectives of time, or time attributive adjectives, add temporal information to the noun phrase and appear to be used here to add information about the time of an event as in [41] & [42], describe the age of something as in [43] & [44], or provide information about the ordered status of an item or event as in [45] & [46].

[41] NJ Transit service was suspended **early** Monday between Newark Liberty Airport and Elizabeth, and New Jersey-bound Amtrak trains were being held at New York Penn Station, officials told AP. (*Associated Press*, NYC bombing)

[42] Dana Outlaw, the town's mayor, said in the **late** afternoon that there had been 1,200 requests to 911 over the previous 12 hours. (*NPR*, Hurricane Florence)

[43] He settled on an Amazon heroine who hails from an island where women had lived without men since **ancient** times. (*New York Times*, Wonder Woman movie)

[44] The second half of the interview focused largely on online culture, social media and the opportunities and challenges that digitally connected lives place before **young** people. (*The Atlantic*, royal wedding)

[45] In the end, they chose to rescind the **old** order though Spicer maintained the first was 100 percent legal and constitutional. (*Associated Press*, Trump's travel ban)

[46] A week ago, Prince was healthy enough to give what would be his **final** public performance. (*New York Times*, Prince's death)

The last key feature for the no extent group is adverbs of place. Place adverbials provide locative reference about an event and are typically associated with features that indicate situation-dependent reference as “these forms typically mark exophoric reference to places . . . outside the text itself” (Biber & Finegan, 1989). Given that news articles are typically written about events that are widely known by the public, the use of place adverbs to add information about location of an event makes sense (see [47] & [48]).

[47] It comes after reports earlier Sunday that rescue workers are still combing through muddy ash and debris left **behind** by the catastrophic blaze for signs of human remains. (*Fox*, CA fires)

[48] Dozens of more players protested before the Raiders-Redskins game, the final one of the day and not **far** from the White House in Landover, Maryland. (*Associated Press*, take a knee)

Taken together, features of no extent texts demonstrate that functionally, these texts stick to acts of reporting information and do not exhibit any type of involvement, stance, or opinion.

8.5.2 *Extreme Extent Articles*

Many features that came out as key for articles rated as extreme extent are features that mark grammatical stance which is how an author expresses their assessments, attitudes, and opinions towards the information they are communicating (Biber & Finegan 1989; Gray & Biber 2012). While the key features demonstrate only small effect sizes, it is important to consider that the majority of these features are related to functions of opinion expression which suggests that extreme bias in news is associated with the expression of perspective and opinion. Because there are sixteen marked features for the extreme bias group, the discussion around them will be organized by features of similar stance type and will then proceed on to the features not explicitly marked for stance.

For the present study, two main categories of stance will be vital: attitudinal stance which conveys the author's personal beliefs or emotions and epistemic stance which conveys certainty, doubt, or level of commitment to the validity of the expressed information (Arrese, 2015; Gray & Biber, 2012; Biber & Zhang, 2018).

The features of epistemic stance that came out as key for texts rated as extreme bias include likelihood adverbs, epistemic predicative adjectives, 'WH' complement clauses controlled by verbs of likelihood, and verbs of likelihood in contexts other than complement clauses (see [49] – [58]). In all cases, the features of epistemic stance serve to position the author with respect to the assertions made. In [49], [50] & [54], features of likelihood are being used to hedge and indirectly reflect the author's opinion about an assertion. In [51], [52], and [58] features of epistemic stance are used to express certainty or affirm that predictions were accurate, and in [53], [55], [56], and [57] features of epistemic stance are used to make value judgments about the actors in the assertion (i.e., 'reasonable people', 'Trump', 'people from certain demographics').

Adverbs: Likelihood

[49] Nearly half of Americans believe President Trump **likely** committed a crime related to Russian interference in the 2016 presidential election, according to a new poll published Thursday (*USA Today*, Russia sanctions)

[50] They will attack again, maybe at another gay bar, or another office Christmas party, or a coffee house, or a sporting event, or a church. And when they do, there will **probably** be some Americans who excuse their actions out of eagerness to blame other Americans, instead. (*National Review*, Orlando shooting)

Adjectives: Epistemic predicative

[51] We are also **confident** that the judges will rule against Pai's rollback and the way he conducted the proceedings. (*HuffPost*, net neutrality)

[52] Devos identifies her target: a beauty-shop owner (Nathalie Baye) who seems **certain** to have been the driver and yet is too brittle under her strained, professional smile for Devos to hate. (*The Nation*, Wonder Woman movie)

‘WH’ complement clause: Verb of likelihood

[53] Some on the left simply refuse to believe when terrorists make claims about themselves and instead **postulate about what** their intentions are. (*National Review*, Orlando shooting)

[54] Added Gallegos, " It **appears what** we've done is given white, non-ethnic people the pen to write stories for the entire population, but don't empower minorities and give them the opportunity to do the same. (*USA Today*, Pepsi ad)

[55] But "if you've only got people who are from certain demographics **imagining what** normal looks like, this is what you potentially could end up with. (*USA Today*, Pepsi ad)

Verb of likelihood: Other

[56] They could be Baptist missionaries looking for lost souls, it's true, but most reasonable people would **assume** otherwise. (*Fox*, Take a knee)

[57] Trump once lost almost a billion dollars in a single year. He **bet** on the housing market just before history's biggest real estate crash. (*The Atlantic*, Clinton's emails)

[58] The media ridiculed conservatives for that argument, but might have to pay attention to a respected voice on the left. As **predicted**, the #metoo movement has finally reached a reckoning. (*Breitbart*, #Metoo Movement)

The two features of attitudinal stance that came out as key for articles rated as extreme extent include 'that' complement clauses controlled by attitudinal nouns and predicative attitudinal adjectives (see [59] – [63]). Attitudinal features of stance can either reflect the author's belief about a situation as in [61] and [63], or they reflect the author's attitude about an actor in a situation as in [59], [60], and [62].

‘That’ complement clause: Attitudinal noun (express the attitudes of the actors in the story)

[59] Of the 49% of Americans that think Trump likely acted illegally, 30% hold **the belief that** he's guilty on suspicion alone while 19% think there's already solid evidence to back it up, according to the * poll. (*USA Today*, Russia sanction)

[60] Osama bin Laden couldn't have really attacked the World Trade Center **in the hope that** he would avenge Christians' 15th-century conquest of Muslim Spain. (*National Review*, Orlando shooting)

Predicative attitudinal adjectives

[61] Tuesday's cancellation of the comedienne Roseanne Barr's show was simultaneously **surprising and unsurprising**, if that is at all possible. (*Newsmax*, Roseann Ambien tweet)

[62] "Trump's decision does no good to everyone except a few American steel enterprises," Li Xinchuang, the vice chairman of the China Iron and Steel Association, said, while calling Trump's move "**stupid**." (*Newsmax*, Trump's tariffs)

[63] It is **astonishing** that Scalia, a conservative Catholic with rigid views on sexual morality, whose judicial philosophy of \x93original intent\x94 specifically barred interpreting the Constitution in the light of modern circumstances, so successfully imprinted his views on the legal world. (*The Nation*, Scalia's death)

Stance nouns and cognitive nouns also demonstrated a slightly higher use in texts rated as extremely biased. Like the other features we have seen thus far, these features are used to express epistemic stance (see [64]-[66]) and cognitive attitude (see [67]-[69]), respectively.

Nouns: Cognitive

[67] In **fact**, we expressed grave reservations about his fitness for office. Those **concerns** remain, particularly in the national security arena, where presidents have fewer constraints on their power. (*USA Today*, Trump elected)

[68] More than any president in **memory**, Trump will need wise hands at his side. (*USA Today*, Trump elected)

[69] "The Muslim woman in their ad operates as nothing more than a signifier for diversity and a vague **notion** of resistance." (*HuffPost*, Pepsi ad)

The next two features that came out as key are necessity modals and conditional subordination.

While these two features are not explicitly marked for stance, they are two features commonly seen as expressions of personal stance and often fall on dimensions of overt expressions of argumentation in MD analyses (Biber, 2014). As examples [70] and [71] demonstrate, modals of necessity certainly seem to be used to express an author's opinion, and in opposition to hedging, place obligation on the actors of the statement. Conditional subordination, as seen in [72] and [73] describe certain limits or conditions for the situations.

Modal of necessity

[70] Republicans **must** adapt to a diversifying electorate or lose their influence. (*National Review*, 2016 Presidential election).

[71] Voters **ought** to punish that poor judgment when evaluating Clinton, but if their vote flows from a cumulative comparison of both candidate's flaws, rather than reflexive disgust at the one that they read about most recently, Trump should easily lose to Clinton even if her emails did violate the law. (*The Atlantic*, Clinton's emails)

Subordinating conjunction: Conditional

[72] The news comes just a few weeks after in early September the New York Times' executive editor Dean Baquet said he would go to jail to print Trump's tax returns, **even if** he obtained them illegally. (*Breitbart*, Trump's taxes)

[73] It's a sad commentary on the state of conservative politics that the only thing standing between the United States and the death of its founding document was a brilliant 79-year-old jurist. But **unless** Republicans stand up on their hind legs now, that will certainly be the case. (*Newsmax*, Scalia's death)

Word count and conjunctions also came out as key for this group. These features are not directly related to expressions of opinion but do demonstrate that texts rated as extremely biased are longer and rely on clausal and phrasal elaboration more than texts rated as no bias. This is unsurprising considering many of the articles in each group are on the same topic, and typically there is more room to express opinion, even if accidentally, in a text that is longer. In further support of this idea is the fact that features of density (i.e., non-finite clauses, nouns, etc.) did not come out as key for the no extent group which suggests that texts in that group are indeed able to provide the same information in a shorter manner – and that density of information is not to blame. Interestingly, subordinators and modals have also been demonstrated to be markers of fake news, so it is unsurprising that they are associated with extreme bias (Grieves & Woodfield, 2023; Torabi, 2019).

The last three features that emerged as key are emphatics, downtoners, and predicative adjectives. Again, while these features are not expressly related to stance, these features can be

evaluative and used to express stance in a more lexical manner. As seen in [74] and [75], emphatics are used to emphasize a sentiment while downtoners like those in [76] and [77] are used to remove the validity from something in a sentiment. Predicative adjectives, like those in [78] and [79] can be used to express evaluations.

Emphatics

[74] But that's not **nearly** as frustrating as the story behind where these emails were discovered. (*AlterNet*, Anthony Weiner sexting scandal)

[75] "I love Prince because he put out **such** great music and he was a great performer," Obama said at a press conference in London. "I didn't know him well. He came to perform at the White House, and he was extraordinary, and creative, **really** original, and full of energy. So it's a remarkable loss." (*Newsmax*, Prince's death)

Adverbs: Downtoner

[76] See what a massive, Hillary-shaped bullet America **just** dodged? (*Breitbart*, Woman's March)

[77] The N.R.A. disagrees, saying the Washington measure will strip gun owners of their rights **merely** on the say-so of someone else. (*The New York Times*, gun control)

Adjectives: Predicative

[78] Twenty seventeen has exposed many monsters, and that is a positive thing. Breathless reports of a looming monster epidemic, however, seem greatly **exaggerated**. (*National Review*, #MeToo Movement)

[79] Trump's steel and aluminum tariffs are absolutely **horrible** for American workers. People who have money on the line are **fearful** about what this move means for the country as a whole. (*AlterNet*, Trump's tariffs)

All three of these features were marked in fake news texts in Grieves and Woodfield (2023) so it is interesting that they are also associated with extreme extents of bias in newspaper writing. As Grieves and Woodfield (2023) point out, adjectives and adverbs are often avoided in news writing because "they tend to add inconsequential information while risking weakening or

editorializing statements.” (p. 54). Indeed, this appears to fit with the other key features that have emerged as key for this group of texts suggesting that features related to functions of expressing opinion are associated with extreme extents of bias.

The key feature analysis comparing articles rated as no extent to those rated as extreme extent demonstrates that there are grammatical differences based on extent of bias. The key features for extreme extent include features marked for epistemic and attitudinal stance, clausal and phrasal elaboration, description, emphasis, and evaluation while key features for texts rated as no extent include features related to reporting what happened (who, what, when, and where). Perhaps unsurprisingly, many features in texts rated as extreme extent are related to stance whether directly or indirectly and none of the key features in texts rated as no extent were related to stance. It seems logical that a heavier reliance on stance would be associated with more extreme bias, however, to the best of the researcher’s knowledge, this connection has not been explicitly or empirically made. This again demonstrates how telling a key feature analysis can be in an intra-register examination.

8.6 Conclusions

Importantly, in both key feature comparisons, we see that there are differences in lexicogrammatical features based on direction and extent of bias, and that these features can be attributed to functional and stylistic characteristics of each group. Specifically, in left versus right rated texts we saw that right-rated texts were involved at the individual level and inexplicitly expressed stance through verb complement clauses while left-rated texts were involved at the group level and inexplicitly expressed stance through non-finite clauses.

Additionally, left-rated texts demonstrate a tendency towards formality while right-rated texts are more colloquial in nature. The key features for extreme extent texts include features marked for epistemic and attitudinal stance, clausal and phrasal elaboration, description, emphasis, and evaluation while key features for texts rated as no bias include features related to reporting what happened (who, what, when, and where). Importantly, most of the features that were key for extreme extent texts functionally express opinion and were similar to those previously connected to fake news. We cannot explicitly state that these differences are the reason that raters perceive these texts as left/right or biased/unbiased, but we would be remiss not to at least consider the differences as a contributing factor.

CHAPTER 9. KEYWORD ANALYSIS

9.1 Introduction

The purpose of this chapter is to present the methods and results of the keyword analysis. The impetus behind looking at both grammatical and lexical variation across divisions of bias is that we can get different information from each examination. As we saw above, grammatical variation can highlight differences that are functional in nature (e.g., stance), but lexical information can tell us the “aboutness” of a group of texts (i.e., topics of concern), and further, evaluation is often expressed lexically (Biber & Zhang, 2018); thus, it is important to examine both to accurately identify patterns that may be linked to bias, and a keyword analysis is a useful way to comparatively analyze lexical features. While grammatical features are important for identifying functions, lexical features can provide some indication of sentiment or “aboutness” that readers may perceive as bias (McEnery, 2015). This type of analysis is considered particularly useful here as the goal in this dissertation is to determine whether identifiable

linguistic patterns exist across clines of political leaning and extent and keyword analyses are “[used] to identify the words that are especially characteristic of the texts in a target discourse domain” Egbert and Biber (2019, p. 77).

A keyword analysis works in a similar fashion to the key feature analysis described above. However, instead of Cohen’s *d*, a keyword analysis depends on tests of log-likelihood to determine whether words are statistically more or less likely to appear in a target corpus (based either on word frequency or text dispersion). In addition, they differ in terms of unit of observation. In a keyword analysis, the unit of observation is a text whereas in a key feature analysis the unit of observation is an entire corpus or subsection of a corpus.

Two methods of keyness were used in the present analysis. The first, corpus frequency keyness, relies on word frequency to calculate keywords, and the second, text-dispersion keyness, uses the text, rather than the corpus, as the unit of observation (Egbert and Biber, 2019). Keywords from both methods were compared and only words that emerged as key for both methods were examined further. This allowed for both dispersion and frequency to be considered and made for a smaller selection of keywords for each grouping. Dispersion is important as it has been shown as a valid method to capture the “aboutness” of a corpus (Egbert & Biber, 2019). The frequency method was also considered important as “how much” an entity or actor is discussed in a text is an important factor of coverage bias.

Further impetus behind using two methods of keyness is related to the fact that each method contains limitations that the other method accounts for, and currently, there is no consensus on one perfect solution (Gries, 2021). Specifically, corpus frequency keyness does not consider dispersion which can lead to keywords from very few texts misrepresenting the entire group if these texts contain repeated use of a frequent word; however, considering frequency can

highlight content-distinctive words that are key in the target domain. Conversely, attempts to consider dispersion have typically relied on an analyst arbitrarily defining a threshold range, and while the results of dispersion methods are content-generalizable, in a small data set, dispersion may remove insights to content-distinctive words (see Baker, 2004 or Millar & Budgell, 2008 for examples, or Gries, 2008 for a comprehensive overview). Egbert & Biber (2019) responded to the issue in keyword analyses by developing the text-dispersion method which completely disregards word frequency under the hypothesis that a word will likely have a moderate frequency if it occurs across multiple texts. Given the lack of consensus and the fact that there is valuable information to be obtained from both frequency and dispersion methods (how much a word is used and in how many texts), results from both methods will be relied upon as a means of extracting the most valuable keywords. The dataset in the present analysis is certainly small and results do show that not considering results of the frequency method would have excluded words that are content-distinctive. An example of a word that I feel was important to understand in terms of both dispersion and frequency was *Trump*. This word occurred in over 100 texts and over 2,200 times. I think that understanding both pieces of information are important to highlighting the way bias can be expressed linguistically. At the very least, it does not hurt to consider both pieces of analysis and makes the resulting keyword lists much more manageable.

9.2 Methods of Keyword Analysis

9.2.1 Corpus Frequency Keyness

The first keyword analysis, here referred to as the corpus frequency method after Egbert & Biber (2019), was run with AntConc which relies on frequency to calculate keywords. The comparison was done statistically using log-likelihood ratio, and a “keyness” score that uses the

combined score of both the target (here one grouping of biased texts) and reference (here the grouping of texts rated as biased in the opposite direction) type score. The log-likelihood (LL) test is a null-hypothesis significance test that indicates the likelihood of observing the data under the null-hypothesis; LL is typically used in keyword analyses to determine whether the difference in relative frequencies of a word observed in two corpora is due to chance or not. Simply put, “the smaller the p-value, the more likely that the word’s strong presence in one of the sub-corpora isn’t due to chance but a result of the author’s (conscious or subconscious) choice to use that word repeatedly” (Baker, 2006b, p. 125).

Log-likelihood is calculated by creating a contingency table, as in table 10 below (<http://ucrel.lancs.ac.uk/llwizard.html>), for every word type in the target corpus.

Table 9.1

Contingency Table for Log-likelihood: Corpus Frequency

	<u>Corpus 1</u>	<u>Corpus 2</u>	<u>Total</u>
Frequency of a word	a	b	a + b
Frequency of other words	c - a	d - b	c + d - a - b
Total	c	d	c + d

In this table, ‘a’ corresponds to the observed frequency of a word in the target corpus, ‘b’ corresponds to observed frequency in the comparative corpus (and collectively these are the observed values, or O), ‘c’ is the total number of words in the target corpus, and ‘d’ is the total number of words in the comparative corpus (or the N values). To calculate the LL, we need to compare the observed values (O) to the expected values (E) which can be calculated with the following formula (<http://ucrel.lancs.ac.uk/llwizard.html>):

$$E_i = \frac{N_i \sum_i O_i}{\sum_i N_i}$$

Where N_i is the total number of words in the corpus and O_i is the observed frequency of a word in the corpus. Log-likelihood can then be calculated with the following formula:

$$G^2 = 2 \sum_i O_i \ln \left(\frac{O_i}{E_i} \right)$$

Where O_i is the observed frequency of a word in the target and reference corpus and E_i is the expected frequency of the word occurs in the target and reference corpus. The higher the LL value, the greater the disparity in frequency scores. In other words, “high LL value or keyness suggests that a word has a more significant relative frequency difference between the two corpora.” (Pojanapunya & Todd, 2018 p. 5)

In this study, as with the key feature analysis, comparisons were conducted across texts rated as extreme extent to texts rated as no extent as well as from texts rated as left to those rated as right. The corpus frequency method of keyness is often relied upon to represent content-distinctive information about a corpus but has notable limitations achieving content-generalizability since it does not consider dispersion. That is to say that if a word occurs frequently, but in few or even one text, this method will mark the word as key for the entire group. To avoid this type of potential misrepresentation while still capturing words that are content-distinctive, an additional keyword analysis was conducted that relies on dispersion across texts; this method is discussed below.

9.2.2 Text Dispersion Keyness

Text dispersion keyness, uses the text, rather than the corpus, as the unit of observation (Egbert and Biber, 2019). This method disregards word frequency and instead generates keyword

lists based solely on word dispersion across texts. Because the goal is to identify lexical patterns that are characteristic of each bias grouping (left, right, extreme extent, and no extent), considering dispersion is essential as it will paint a clearer picture of lexical features that are content-generalizable to each group. As a result, this may provide more insight into qualities specific to political leanings as well as to extent of bias.

As with corpus frequency keyness, text dispersion keyness also relies on log-likelihood with the following changes to the contingency table:

Table 9.2
Contingency Table for Log-likelihood: Text Dispersion

	<u>Corpus 1</u>	<u>Corpus 2</u>	<u>Total</u>
Texts with word	a	b	a + b
Texts without word	c - a	d - b	c + d - a - b
Total	c	d	c + d

In this table, ‘a’ corresponds to the observed number of texts where the word occurs, ‘b’ corresponds to observed number of texts where the word occurs in the comparative corpus (and collectively these are the observed values, or O), ‘c’ is the total number of texts in the target corpus, and ‘d’ is the total number of texts in the comparative corpus (or the N values).

Additionally, the formula to calculate the expected values (E) and the LL are changed in the following ways:

$$E_i = \frac{N_i \sum_i O_i}{\sum_i N_i}$$

Where N_i is the total number of texts in the target and reference corpora and O_i is the observed number of texts where the word occurs in the target and reference corpus, and:

$$G^2 = 2 \sum_i O_i \ln \left(\frac{O_i}{E_i} \right)$$

Where O_i is the observed number of texts where the word occurs in the target and reference corpus and E_i is the expected number of texts where the word occurs in the target and reference corpus.

The first keyword analyses were run comparing texts rated as left to texts rated as right (presented in 9.3) and the second was done comparing texts rated as extreme to those rated as no extent (presented in 9.4). The corpus frequency approach was done using AntConc and the statistical threshold used for a cutoff point was $p < 0.05$ (+ Bonferroni). With this cutoff, a list of 44 keyword types was produced for left-rated texts, a list of 45 keyword types was produced for right-rated texts, a list of 39 keyword types emerged for extreme extent texts, and a list of 86 keyword types emerged for no extent texts.

9.2.3 Combining both Keyword Methods

The text-dispersion method was run using a Python code which writes each word from the corpus, a keyness value, and the number of texts the word occurs within to a CSV file. Because this is done for every word in each corpus, the text dispersion lists were necessarily very long; thus, every word that emerged as key on the shorter corpus frequency lists were compared to their respective text-dispersion lists, and words that appeared as key on both lists were taken as results in the present analysis. In other words, keywords from the frequency lists were searched for in the text dispersion lists, and words that were key on both lists were retained in this study. Words that were not key on both lists were dropped. In order to determine what counted as 'key' for the text-dispersion method, the same cutoff as the frequency method was

used at $p < 0.05$ meaning that the keyness value needed to be at 3.84 or above as this is the LLR-value denoting significance (Rayson, 2008b; 2013).

Taking words that met this cutoff on both lists produced results that are key for reasons of frequency and dispersion and did not require that I set any arbitrary threshold range for dispersion since the words that were key on both lists occurred in 10% or more of the texts for each group. Additionally, a stoplist was not necessary as most function words were eliminated by one method or the other. A total of 15 keywords for left-rated texts appeared on both lists, 14 keywords for right-rated texts, 18 keywords for extreme bias texts, and 10 for no bias texts. Results are presented and discussed below.

9.3 Results of the Keyword Analyses: Left Vs Right

9.3.1 Left -Rated Texts

After both methods of keyness were run, a total of 15 words remained key for the left leaning group. Due to the fact that two methods of dispersion were run, the tables in this section presenting results will contain information from both methods. Specifically, for the frequency method, columns detailing the word’s frequency as well as the frequency keyness value are presented, and for the dispersion-method, the number of texts the word was dispersed across and the dispersion keyness value are presented. The words that were key for the left-rated texts, as well as the keyness value for dispersion and frequency, the number of texts the word is dispersed across, and the word’s total frequency are presented in table 9.3 below.

Table 9.3

Keyword List for Texts Rated Left

Word	# of Texts	Frequency	Dispersion	Word
Frequency		Keyness Value	Keyness Value	
2232	107	91.17	10.07	Trump

211	44	68.15	15.64	Russia
119	66	43.13	14.30	anti
147	55	38.3	16.13	health
54	13	32.62	10.13	lgbtq
48	28	30.83	16.96	resistance
143	38	30.27	11.22	violence
186	40	26.5	10.91	nation
58	58	25.94	9.93	Muslims
210	89	25.94	9.88	power
199	54	24.87	9.23	war
193	46	24.75	10.01	gun
102	49	22.86	15.43	protest
116	64	22.55	11.17	act
160	74	22.3	9.53	rights

Table 9.4 lists the keywords for the left leaning group divided into three major semantic domains: social and political issues, people, and political unrest. These semantic domains provide a clear picture of the issues and actors that are significant in text rated as left leaning. The social and political issues that emerge are *health (care)*, *gun (control)*, *(human) rights*, and *LGBTQ*. The people that are focused on include *Trump* and *Muslims* while the places include *Russia* and *nation*. The next semantic domain that emerges relates to a theme of political unrest and suggests a rather aggressive sentiment with *anti*, *resistance*, *violence*, *power*, *war* *protest*, and *act*. Each of these semantic domains are explored in detail below.

Social and political issues	People	Political unrest
health	Trump	anti
LGBTQ	Muslims	resistance
gun	(LGBTQ)	violence
rights		power
nation		war
Russia		protest
		act

The Semantic Domain of ‘People’ in Left Leaning Texts

The first semantic grouping, people, demonstrates that texts rated as left leaning focus heavily on Trump. Given that the years of interest were 2016 – 2018, a focus on Trump indicates that bias reflected in lexical choices in news can be time-sensitive; however, this is not to say that the tendency to focus on actors of the opposition is not a general characteristic of news bias - indeed results below and previous studies suggest this may be a continuing tendency of media bias - but it is to say that the actors of opposition will necessarily change with time.

Trump's name occurs in 40% of the texts for this group which is a drastic overfocus compared to texts rated as right leaning. A look at some of the texts included in the left leaning group demonstrate a tendency to focus on Trump, even on topics that do not require a connection to him. For example, the text in sample [a] below is a text on Brexit, [b] is a text on the solar eclipse in 2017, [c] is on Prince Harry and Meghan Markle's wedding, [d] is on NYC bombings, and [e] is on the 2018 World Cup.

[a] Did Trump-Style Islamophobia Break Up the European Union? (Brexit)

[b] Trump Looked Directly at the Solar Eclipse Without Eye Protection, Because of Course He Did (Solar Eclipse)

[c] The British Government Is Really This Afraid of Upsetting Trump (Royal wedding)

[d] Donald Trump on New York Bombing: 'I Called It Before the News' (NYC bombings)

[e] FIFA Prepares for a Post-Trump World (World cup)

Considering that there are barely any drastic differences in the topic divisions between the left and right rated texts (refer back to table 7.4), the overfocus on Trump is surprising; however, it matches the common consensus and previous research that the media unfairly focused on Trump. A study conducted by Harvard's Shorenstein Center on Media, Politics, and

Public Policy showed that 41% of all news coverage was on Trump which is three times as much as coverage of previous presidents (Patterson, 2017). In addition, several journalists attest to their own overt focus on trump as can be seen in the quotes below:

“The mainstream media certainly gave Trump harsh and even overtly hostile coverage.”
(Chait, 2021)

“I criticized the press’s obsession with the former reality-TV star, yet I was caught up in it, too... I certainly was aware that if I wrote a column with Trump’s name in the headline, it probably would find a passionate audience: thousands of comments and retweets, hundreds of emails, requests to talk on TV. And because I wrote about the news media, and Trump never stopped using the news media as a foil, there was so much to say.” (Coll, 2022: p. 3)

Interestingly, two of the headlines in the left leaning group point out the focus on Trump which suggests that the media were aware of this while it was happening:

[f] Sucked in Again: With One Stupid Tweet, Donald Trump Grabs Much of Our Attention

[g] All Those Donald Trump Headlines Are Distracting You from the Really Scary News

This phenomenon has been referred to as the Trump Effect or the Trump Factor which explain that Trump’s reliance on inflammatory rhetoric and unusual behavior made him more ‘newsworthy’, hence his repeated appearance as a news headliner. At the same time, Trump actively attacked the press with claims of bias and fake news effectively causing a feedback loop in which his claims of a biased media were verified by the obsessive reporting leading to diminished trust in the media for many people (Coll, 2022).

Trump being key in left leaning texts is also related to several other words that were key in text dispersion but did not make the frequency cutoff. Because these words were also shown to be collocates of Trump, they will be considered in this analysis. The words that are collocated and key in only one of the keyword methods employed here include: twitter, liar, bully, and bullying. In fact, when searching for these words in the texts, it is clear that they co-occur and

that they suggest a clear negative sentiment of Trump – something that is often connected to a ‘left’ political leaning. Text excerpts 1-5 below offer some examples, keywords are presented in bold.

[1] Throughout last year’s campaign and into the early months of his presidency, the concern among Mr. **Trump**’s advisers was mainly political. Every time the president let loose with one of his **Twitter** blasts, it distracted from his agenda and touched off a media frenzy that could last for days. (*The New York Times*, Covfefe Tweet)

[2] All of that being said, this is a case more complicated than just calling out Nike for commodifying dissent. Kaepernick has spent the past two years being showered with hatred and death threats, vilified on **Twitter** and from the presidential **bully** pulpit. (*The Nation*, Nike hire Kaepernick)

[3] There are dozens of reasons why **Trump** won, but misogyny was a big part of it. And if you didn’t know women can be misogynistic, now you do. Trumpettes, you voted for a grotesque **liar**, bankrupt, and groper with no public-service experience. The **Trump** phenomenon was like a **Twitter** comment thread come to life: aggressive, **bullying**, ignorant, and contemptuous of women. (*The Nation*, 2016 Presidential Elections)

[4] Editor’s note: Donald **Trump** regularly incites political **violence** and is a serial **liar**, rampant xenophobe, racist, misogynist and birther who has repeatedly pledged to ban all **Muslims** - 1.6 billion members of an entire religion - from entering the U.S. (*HuffPost*, NYC bombings)

[5] If you believe the Comey statement, you must take away from it that **Trump** is a **liar**, a **bully** and a criminal. (*The New York Times*, Comey’s testimony)

In addition, *Muslims* emerged as key under the semantic domain of people. This shows that left leaning texts have a clear focus on this group of people which is further supported by the fact that *islamophobia* correlated highly with *Trump* for the frequency method of keyness and was key in its own right in the text dispersion method (but left out of the results since it was not key in both methods). One possibility could be that Trump’s ban on several middle eastern countries is commonly referred to as ‘*a ban on Muslims*’ in these texts and not in right leaning texts.

Another possibility is that the topics included many situations directly relevant to the Muslim community; namely, the travel ban, the Orlando shooting, the Kendall Jenner Pepsi ad, Brexit,

and ISIS activity; however, there is not a significant difference of occurrence in texts on these topics in the left rated group which strongly suggests that there is a choice in this focus.

Keywords presented are presented in bold.

[6] The ad failed to mention any of the issues that have troubled American **Muslims** over the past few months. Continuing religious-based discrimination and surveillance, President Donald **Trump**'s backdoor ban on **Muslims** and his resounding silence about attacks on mosques, the **bullying** of Muslim kids, the rise in prominence of white supremacist groups, the fight for black lives. (*HuffPost*, Pepsi ad)

[7] In a statement, Ruah said that the new executive order signed by **Trump** is "more of the same **Islamophobia** that targets **Muslims**." The **Trump** administration has denied that the ban targets **Muslims**, but critics see the order as a way for **Trump** to legally fulfill his campaign promise for a "complete shutdown" of **Muslims** entering the country. (*HuffPost*, Trump's travel ban)

[8] As **Muslims** all over America sent their sympathies to Orlando, **Trump** mocked his Democratic rival's insistence that "**Muslims** are peaceful and tolerant people." He accused **Muslims** of causing "death and destruction" by covering up terrorism in their midst. (*AlterNet*, Orlando shooting)

The Semantic Domain of 'Social and Political Issues' in Left Leaning Texts

The next semantic domain for the keywords in texts rated as left leaning relate to social and political issues. These issues include *health* (care), *gun* (control), (human) *rights*, and *LGBTQ*. Again, considering that the division of topics was not significantly different between left and right leaning groups, these issues clearly emerge as a way of framing the topics for the left leaning group. This is further supported by the fact that *affordable*, *care*, *human*, and *transgender* also emerged as key in the text dispersion method (but were only correlates in the frequency method). In addition, upon examination of these keywords in context, it is clear that a large chunk of these words occur in approximation to *abortion* and *birth control* suggesting that reproductive issues are also relevant in the left leaning group.

Considering that these issues tend to be connected to liberal views, these keywords are unsurprising. Particularly, issues important to the ‘progressive left’ include women’s reproductive rights, racial and gender equality, laxer immigration policy, and social welfare (Dimick et al., 2014) all of which fit with the issues identified by this keyword analysis. Text excerpts [9]-[14] exemplify the use of *rights*, *health (care)*, and *LGBTQ* which are all closely related and [15] – [17] exemplify the use of *gun (control)*. Keywords are presented in bold while related terms of interest are underlined.

[9] Donald **Trump** is appointing judges who are #BadForWomen to lifetime positions. Many have spent their careers advancing policies that would limit a woman's **right** to **health care** & opposing **LGBTQ rights**, among other terrible things. (*HuffPost*, Trump’s judicial picks)

[10] **Anti-abortion** legislators claim the law, which is a set of restrictions folded into a single House bill called H.B. 2, is necessary to protect women’s **health**, while abortion providers and other **health care** experts say they are unnecessary. (*AlterNet*, Scalia’s death)

[11] We marched so that women have access to quality **healthcare**, **affordable birth control** and to protect women’s **rights** to abortion. (*AlterNet*, 2017 Women’s march)

[12] Despite local and federal fair-housing regulations, about an eighth of respondents reported experiencing housing discrimination such as “landlords who refused to rent to” or mistreated them on the basis of **LGBTQ** identity. About one in 10 trans respondents reported they could not get the **health care** that they needed because they are **transgender**. (*The Nation*, Gay Marriage)

[13] "Every child deserves to attend school in a safe, supportive environment where they can learn, thrive, and grow," DeVos writes. Gupta was credited with nearly the same quotes in a May 2016 press release on ensuring the civil **rights** of **transgender** students. (*CNN*, Betsy DeVos appointment as Secretary of Education)

[14] Bush, in the end, bowed to the same extremists Reagan did when it came to abortion, AIDS and **LGBTQ rights**. (*HuffPost*, Bush death)

[15] Let’s Make **Gun Control** a Gay Issue. It’s time to glitter bomb the NRA. Who’s in? (*The Nation*, Orlando shooting)

[16] After nearly two decades of expanding legal access to firearms, a succession of horrific shootings like Mr. Houser’s have refocused attention on **gun control**. Since the

2012 massacre of 26 elementary school children and teachers in Newtown, Conn., **gun control** advocates have scored some significant victories in state legislatures. (*The New York Times*, gun control)

[17] A handful of well-known celebrities and political leaders have recorded YouTube videos encouraging the public to use a Brave New Films documentary on rampant **gun violence** to press for new **gun control** laws across America. (*AlterNet*, gun control)

The Semantic Domain of ‘Political Unrest’ in Left Leaning Texts

The last semantic domain that emerges in the keywords for the left leaning group is political unrest. This domain demonstrates that texts rated as left maintain an air of resistance or aggression against the government with words like *anti*, *protest*, *violence*, *war*, *resistance*, *power*, and *act*. There were also several words that were key in the text dispersion model (but not the frequency model) that add to this sentiment including *march* and *voices*.

An examination of these keywords in context shows that *anti* most often occurs as a prefix with its most frequent collocates being *Trump*, *gay*, *LGBTQ*, *Semitic*, *Muslim*, *abortion*, *discrimination*, *contraception*, *slavery*, *immigrant*, *black*, and *choice*. The collocates for *violence* include *police*, *gun*, and *political* and the collocates for *act* include *affordable care (act)*, *violence against women (act)*, *voting rights (act)*, *(act) of violence*, and *(act) of protest*. Many of these concepts directly relate to the social and political issues identified above solidifying that issues of reproduction and equality are a major focus of left leaning texts. In addition, it suggests that the political unrest identified in this semantic domain is due to issues of equality. The text excerpts below demonstrate this sentiment with keywords presented in bold.

[18] Gutting the Voting **Rights Act**, affirmative action, **gun control**, the **Violence Against Women Act**, and - one last middle finger raised against **human** progress - EPA regulations to control global warming. (*The Nation*, Scalia’s death)

[19] Did that **war** against men behaving badly feed into the larger battle for women's equality? In many ways, yes: Susan B. Anthony herself began as a temperance organizer. (*The New York Times*, #MeToo)

[20] "As **Muslims**, we're united in our outrage over this senseless **act of violence**." Because it is inconceivable that the people of such an advanced **nation** will tolerate an ever-worsening state of armed **violence** and insurrection. (*CNN*, gun control)

[21] Clinton supporters grew numb as the night unfolded and the presidency increasingly grew within reach of Donald **Trump**, the Republican who models racism, misogyny, white supremacy and gratuitous **violence**, and marches lockstep to a far-right social agenda. (*AlterNet*, Trump elected president)

[22] **Trump** and members of his administration have shown a reckless disregard for democratic norms. They are embroiled in scandals. They abuse **power** and obstruct justice, operate the White House as a personal racket and a tool for lining their pockets with public money and have likely been manipulated by a foreign **power**. (*AlterNet*, Covfefe)

[23] As Chaudhry pointed out, the **voices** of marginalized groups like women are important because **power** does not see itself or its own abuses: "The only way to see what patriarchy is and how it works is to have diverse **voices** that include those marginalized by it." To galvanize the **Muslim** #MeToo revolution, we must amplify the **voices** of victims regardless of whether the perpetrators are our beloved preachers or scholars, and regardless of **Islamophobia**. (*The Atlantic*, #MeToo)

[24] These are people who are rising up and will take to the streets. They will become part of the **Resistance**. And in so doing, they will convince countless others to join them - to take the critical step from passive opposition to active **resistance**. (*The Nation*, 2017 Women's March)

[26] For Democrats, heading to the polls on Tuesday was framed as an **act of resistance** against **Trump**, both symbolically and substantially (*NPR*, 2018 midterm elections)

Taken together these keywords show that left-rated texts demonstrate a tendency to concentrate on enemy #1 – Trump – and to publish stories in a way that highlight the issues that are typically associated with a liberal agenda. Namely those related to healthcare, gun control, and human rights. In addition, there appears to be a sentiment across these texts that indicates an unhappiness not only with the Trump administration, but with how the aforementioned issues are being handled. All in all, these findings align well with previous research on fake and biased news in that a tendency to ingroup/outgroup and overly focus on groups one finds opposing. In

addition, this analysis has revealed a coverage bias in the concentration on Trump and a presentation bias via the tendency to ingroup and outgroup and express negative sentiment.

9.3.2 Right -Rated Texts

A total of 14 words remained key for the right leaning group after both keyness methods were conducted. The words, as well as the keyness value for dispersion and frequency, the number of texts the word is dispersed across, and the word's total frequency are presented in table 9.5 below.

Table 9.5
Keyword List for Texts Rated Right

Frequency	# of Texts	Frequency Keyness Value	Dispersion Keyness Value	Word
111	25	90.34	12.03	wall
80	25	73.19	4.55	illegal
288	35	55.78	6.06	national
289	79	51.57	3.98	Obama
164	68	41.31	4.53	left
58	26	40.76	10.59	liberals
190	61	35.8	4.63	Republicans
75	26	30.89	3.91	emails
206	35	29.67	7.54	Bush
48	24	26.86	5.53	terror
50	23	26.34	3.84	Reagan
53	30	25.71	3.86	terrorism
78	23	24.74	6.84	immigration
323	58	22.38	3.85	Clinton
44	29	22.3	7.31	Freedom

The keywords for right leaning texts fit into the major semantic domains of political actors, especially former presidents, political issues, and political groups (see table 9.6). Like with the left rated texts, these semantic domains help paint a clear picture of the issues and actors that are prominent in the right leaning group. The issues that emerge include the *wall*, *illegal (aliens)*, *Hillary Clinton's emails*, *terror and terrorism*, *freedom*, and *immigration*. The actors

that emerge as key make two of the semantic domains here: (1) the political groups of (the) *left*, *liberals*, and *republicans*, and (2) political actors including Hillary *Clinton* and three former presidents: *Obama*, *Bush*, and *Reagan*. Each semantic domain for right leaning texts is discussed in further detail below.

Political actors (former presidents)	Political issues	Political groups	Other
Obama	wall	left	national
Bush	illegal	liberals	
Reagan	emails	republicans	
Clinton	terrorism		
	Freedom		
	Immigration		
	terror		

The Semantic Domain of ‘Political Issues’ in Right Leaning Texts

The political issues relevant in right leaning texts are made clear by the keywords identified in this analysis. Considering there are no stark differences in the topics that were rated left versus right, the issues of the *wall*, *illegal (aliens)*, *immigration*, *terrorism*, *freedom*, and *Hillary Clinton’s emails* are emphasized in the right leaning group. As with the left leaning texts, some of these words indicate topics that emerge due to time rather than explicit connection to a bias. Specifically, in this case Hillary Clinton’s emails are directly related to the years of interest (2016 – 2018). However, like with Trump, this still represents a tendency to focus on actors of the opposition – especially considering that texts on Hillary Clinton’s emails were exactly evenly split between the left and right leaning groups. Words that add to the political issues important in right leaning texts, but only emerged as key in one of the keyness methods, are *border*, *alien*, and *aliens*. These three keywords are relevant to the issues of *illegal*, *immigration*, *terrorism*, and the

wall as demonstrated in text excerpts [27] - [32] below. As shown in excerpts [33], [34], and [35] below, *freedom* comes up in reference to the first amendment and collocates strongly with speech in right leaning texts.

[27] Whatever President Trump has done right or wrong in this scenario, the **liberal** impulse to blame America first and leave **illegal aliens** blameless has the **left** deeply soiled. . . . **Immigration** and Customs Enforcement is a “**terrorist** organization,” according to Cynthia Nixon, the far-**left** actress now running for New York’s Democratic gubernatorial nomination. (*National Review*, child immigrants separated at border)

[28] Unfortunately, many Democrats see **illegal immigration** as a voter registration drive. They aid, abet and encourage **illegal aliens** to enter and stay in America in the guise of charity, but their actions are all about politics. They see **illegal immigrants** as their wards and dependents. Once the Democrats come to their aid, they prey upon the **illegal aliens** by making them dependent on the government for public assistance and more. (*Fox*, 2018 midterm elections)

[29] The Trump administration also demanded \$25 billion to fully fund the border wall, sticking by Trump’s 2016 campaign promises. The border wall is essential. We cannot solve the problem of immigrants crossing our border illegally until we secure the border. (*Fox*, child immigrants separated at border)

[30] I hate to say I told you so, but I have been warning about this for years. Here’s me in November expressing concern at mass Muslim **immigration**. “Well @nero, you called Radical Islam coming for gays in the West next... 50 dead, Islamic **terrorism** tie eyed in Orlando gay bar shooting... These are dirty secrets the progressive **Left** and social justice warriors would rather you didn’t know. (*Breitbart*, Orlando shooting)

[31] “Like every nation, the United States has a right to control who enters our country, and to keep out those who would do us harm,” said Attorney General Jeff Sessions, pointing out that three of the countries were state sponsors of **terrorism** and the other three were safe havens from **terrorism**. (*Breitbart*, Trump’s travel ban)

[32] What a crock. The **left**, for some odd reason, has a thing for **terrorists**. Or for open borders. Or both. How can it be discriminatory to want to keep out citizens of countries that are known hot spots for **terror**? (*The Washington Times*, Trump’s travel ban)

[33] Unrivaled since the last days of the Roman Empire, this **illegal** invasion has Americans pleading for a security **wall** on their border. (*NewsMax*, Trump’s tax return)

[34] Hollywood’s crassness today is the product of a decades-long progression of the **left**’s selective regard for the cherished pillars of our society’s **freedom**. **Freedom** of speech today is nearly unrecognizable from what the Founders of our country had

envisioned. From the **liberal** perspective, **freedom** of speech is protected - so long as it conforms to the **liberal** orthodoxy. (*The Washington Times*, Kathy Griffin fired)

[35] Is the First Amendment a great thing or what? **Freedom** of speech. **Freedom** of assembly. **Freedom** of religion. Americans of all stripes exercising their constitutional rights. What a reason to celebrate - no matter how you feel about President Trump. (*Fox*, 2017 women's march)

[36] We have the **freedom** as Americans to speak out on matters that deeply concern each of us. But in an ever-polarized society, there is simply no excuse for disrespecting our country, our **National** Anthem, or the sacrifice made by so many brave veterans. (*Breitbart*, Take a knee)

The Semantic Domain of 'Political Actors' and 'Political Groups' in Right Leaning Texts

Because two of the major semantic domains in right leaning texts focus on actors, these semantic domains will be analyzed together. The political actors that emerge as key in this group include *Bush*, *Reagan*, *Obama*, and *Clinton* which refers mostly to Hillary but does include former president Bill Clinton as well. It is interesting that there is such a focus on former presidents in right leaning texts. We can see that the tendency to focus on the opposition or 'enemy' emerges prevalently in both groups with comparatively little focus on actors of their own political orientation. This reflects a very clear coverage bias suggesting that right leaning texts tend to point the finger more commonly at political actors of opposition. As some of the excerpts below demonstrate, it appears that there is a common tendency to compare presidents or make reference to something a former president did (see [37] – [41]). In addition, Clinton comes up as key due to her emails being a common focus in these right rated texts ([42] – [43]). The division of texts on this topic in the left leaning group versus the right leaning group is exactly even suggesting that right leaning texts are bringing up Clinton's emails in other situations and possibly more than strictly necessary.

In addition to former presidents, the terms *left*, *liberals*, and *republicans* came out as key for right leaning texts. The term *liberal* was also key in the text dispersion method. As

demonstrated by text excerpts [44] – [49], these terms are clearly being used to create an ingroup and an outgroup or a sort of ‘us’ vs ‘them’ sentiment further supported by the fact that *liberal* is often used to demonstrate negative evaluation.

[37] The Senate has confirmed 16 judicial nominations from President Trump this year. Of modern two-term presidents, at this point Ronald **Reagan** had 33, Bill **Clinton** had 28, George W. **Bush** had 21, and Barack **Obama** had 11. All this, despite unprecedented obstruction tactics from Senate Minority Leader Chuck Schumer (*Breitbart*, Trump’s judicial picks)

[38] Trump’s approach would seem to echo that of President Bill **Clinton**, who reacted to the midterm election loss (of both the House and the Senate) in 1994 by reaching out to **Republicans**. **Clinton** struck deals on welfare reform and a balanced budget that stand as the great achievements of his presidency. (*Breitbart*, 2018 midterm election)

[39] Trump Must Channel **Reagan** on **National Security Post**. Thirty-five years ago, the tenure of a **national** security adviser to the president was prematurely truncated by a scandal. Ronald **Reagan**’s choice of Richard Allen’s successor proved to be one of the most consequential of his presidency. Donald Trump is now facing a similar moment and must make a similar choice. (*NewsMax*, Trump on national security)

[40] It was a time to recognize, in **Bush**’s words, a “world in which nations recognize the shared responsibility for **freedom** and justice.” (*Breitbart*, Brexit)

[41] The **Obama** campaign did not even bother to refute the claim. Instead, it circulated examples of lines that it said **Clinton** herself had borrowed from **Obama**. The **left** media defended **Obama**, saying that he had not committed plagiarism, but merely, at worst, “poor footnoting.” (*NewsMax*, Melania’s plagiarism)

[42] What is happening now with the FBI and DOJ on Hillary **Clinton**’s **emails** and **illegal** server, including her many lies and her lies to Congress are worse than what took place in the administration of Richard Nixon, and far more **illegal**. (*Breitbart*, Trump’s tax returns)

[43] Later that evening (at 9:41 P.M.), Mills e-mailed Podesta, informing him: “We need to clean this up - he has **emails** from her - they do not say state.gov.” To be clear, Mills was affirming that **Obama** and **Clinton** had e-mailed each other, and that it was done via **Clinton**’s private e-mail address. (*National Review*, Clinton’s emails)

[44] The leader of the Democrats, Rep. Nancy Pelosi, is very unpopular. The voters prefer the **Republicans** and President Trump and the **Republicans** on most issues over the Democrats. (*NewsMax*, 2018 midterm elections)

[45] They can promote the notion that **Republicans** are racist by saying ‘look they are abusing kids because they are black and brown.’ Yes it is perfectly clear that **liberals** are very concerned about mothers when it comes to family separation on the southern border. (*NewsMax*, child immigrants separated at border)

[46] Democrats are in an awful position just now. Hillary Rodham **Clinton** was beaten by Donald Trump; **Republicans** control the Senate; **Republicans** control the House; **Republicans** are about to put an Antonin Scalia style constitutionalist on the Supreme Court, a development made possible by the Democrats’ weak position in the Senate; **Republicans** control 34 of 50 governorships; **Republicans** control the great majority of state legislative houses. What, exactly, are the Democrats up to? Dressing up as vaginas and inviting Madonna to rile up the rubes with empty speeches in D.C. (*National Review*, Betsy DeVos as Secretary of Education)

[47] If you say this tide was driven by white, bigoted, racist voters - your analysis is too shallow. It's because of Mr. **Obama's** radical **liberal** agenda. It's his building a coalition on the disparate interests of minority groups and the political global elite, and ignoring middle-class America. (*The Washington Times*, Trump elected)

[48] "President **Obama's** record of ruling through executive action and regulatory mandates has shown he is willing to circumvent Congress and bypass the will of the people," Sen. David Perdue, Georgia **Republican** and a member of the Senate Judiciary Committee, said Thursday. "We cannot allow the president to continue to use the judicial system to achieve his **liberal** agenda." (*The Washington Times*, Scalia’s death)

[49] **Liberals** talk about unity, but President **Obama** has been the most divisive president in modern American history. In their post-election speeches, Hillary **Clinton** and President **Obama** both called for unity and inclusiveness. These are very nice-sounding words. But "inclusiveness," like so many other words of the **left**, is pregnant with implications and accusations. In reality, it has been not an appeal for unity, except in the most superficial sense, but a battle cry to **liberals** and an indictment of conservatives.

The keywords of right rated texts demonstrate a tendency to focus on former presidents and issues that are typically important to conservatives including illegal immigration, terrorism, and protecting constitutional freedoms. In addition, labels are employed to create an ‘us vs them’ which has been associated with fake and biased news in previous studies. As with the right-rated texts, these keywords provide clear examples of both coverage bias with an intentional focus on political actors of opposing political affiliations and presentation bias with labels that ‘other’ and demonstrate a negative sentiment.

9.4 Results of the Keyword Analyses: Extreme V No

9.4.1 Extreme Extent Texts

A total of 18 words remained key for the extreme extent group after both keyness methods were conducted. The words, as well as the keyness value for dispersion and frequency, the number of texts the word is dispersed across, and the word's total frequency are presented in table 9.7 below.

Table 9.7

Keyword List for Texts Rated Extreme Bias

Frequency	# of Texts	Frequency Keyness Value	Dispersion Keyness value	Word
328	63	119.88	3.85	Clinton
1260	66	89.16	10.07	Trump
132	28	76.51	7.54	Bush
115	34	53.06	3.86	FBI
198	68	44.77	4.84	campaign
168	26	44.15	4.23	gun
123	32	35.73	3.93	children
127	45	34.14	6.21	republicans
66	19	30.97	3.84	illegal
128	31	28.01	7.81	Whites
172	53	26.46	6.32	Democrats
63	24	25.91	4.23	votes
54	23	25.63	8.46	immigration
62	24	25.23	6.91	border
126	38	25.21	5.45	party
43	19	23.46	10.13	LGBTQ
31	16	21.95	3.87	Biden
49	17	19.27	20.60	right-wing

The extreme extent group contains 71% overlap with texts from the right leaning and left leaning group, so naturally several of the keywords in the extreme extent group overlap with the keywords from those groups. Because there are clear differences in the topics that were rated as extreme extent versus no extent, many of the keywords here suggest that articles written on specific topics lend themselves much easier to the label of extreme in their bias. The topics that

emerge in this keyword analysis are mostly political in nature and include: the presidential and midterm elections (*Campaign, votes, and party*), gun control (*gun*), immigration (*immigration, children, illegal, border*), and the topics of Hillary Clinton’s emails, Comey’s Testimony, and Trump’s numerous legal issues (*FBI, Clinton*). A look back at table 7.4 demonstrates that there are indeed more texts on these topics included in this group.

Table 9.8 below lists the words that overlap as well as those that are unique to the extreme extent group before listing the major semantic domains of political actors, elections, legal issues, and social issues in table 9.9. These semantic domains are discussed in greater detail and exemplified with text excerpts below.

Table 9.8
Overlap of Keywords between Directional Bias Groups and Extreme Bias

Overlap with Left keywords	Overlap with right keywords	Unique keywords
Trump	Clinton	FBI
gun	Bush	Campaign
LGBTQ	republicans	children
	illegal	democrats
	immigration	votes
		border
		party
		Right-wing
		Biden
		whites

Table 9.9
Semantic Domains for Extreme Bias Keywords

Political actors	Elections	Legal issues	Social & political issues
Trump	Campaign	FBI	gun
Clinton	Votes	(illegal)	LGBTQ
Bush	Whites		immigration
Biden	(democrats)		children
Democrats	(republicans)		illegal

Republicans	(Clinton)	border
Right-wing	(Trump)	
Party		

The Semantic Domain of ‘Political Actors’ in Extreme Extent Texts

The political actors that emerge as key in extreme extent texts include *Clinton*, *Bush*, *Trump*, *Biden*, *democrats*, *republicans*, *right-wing*, and (political) *party*. This suggests that these political actors are associated heavily with views of extreme bias. A closer examination of the texts in this group reveals that, indeed, the majority of these actors are used in these texts to push an agenda or a sentiment – usually a negative sentiment against the opposing political group. This is done in a number of interesting ways.

For example, **Trump** correlates with *@realDonaldTrump* because many of the texts in this group include references to Trump’s tweets – some to make him look bad ([50]) and some to make him look good ([51]). Additionally, *Biden* is referenced in many of the texts in this group as a comparison to Trump and usually as a way to try and shift blame (see [52] – [55]). *Party*, *republicans*, *democrats*, and *right-wing* are all used to create an ingroup/outgroup sentiment by publications whose leaning is most severe. Text excerpts [56] & [58] are clear examples of negative portrayals of democrats while [57] & [59] are examples of republicans being presented negatively and [60]- [61] demonstrate the term *right-wing* being used to express negative sentiment (much as the terms liberals was for the right extent group). Unsurprisingly, texts from *AlterNet*, *The Nation*, *HuffPost*, *Breitbart*, *NewsMax*, and *National Review* provide the most support for keywords in the extreme extent group.

[50] Not content with simply denying the allegations against him, **Trump** also regularly uses his Twitter account to point fingers at others. Monday was Vice President Joe **Biden**’s turn, and comes a day after **Biden** said **Trump** suffers from an “instinctive abuse of power.”

@realDonaldTrump: @MarkSimoneNY: Watch Joe **Biden's** Long History Of Grabbing, Kissing and Groping Women Who Are Cringing: <http://710wor.iheart.com/onair/mark-simone-52176/watch-joe-bidens-long-history-of-13133819/#ixzz4NJ5E8GFG>" (*HuffPost*, Trump's sexual assault)

[51] Needless to say, the Left is silent on the subject: @realDonaldTrump **Is** President Obama going to finally mention the words radical Islamic terrorism? If he doesn't he should immediately resign in disgrace!

A lot of people laughed at Donald **Trump** when he suggested a temporary ban on Muslims entering the country. No doubt they thought he was further harming his chances in the general election. (*Breitbart*, Orlando shooting)

[52] Joe **Biden**, Groper-in-Chief, Accuses Donald **Trump** of Sexual Assault. On the campaign trail in 2012, **Biden** groped a woman in Ohio - and her companions were not amused. (*Breitbart*, Trump sexual assault)

[53] **Trump's** Morning: Attack **Biden**, female accusers, fellow **republicans**, and electoral integrity - all before noon. (*HuffPost*, Trump sexual assault)

[54] **Biden** was found to have borrowed heavily from the oratory - and the biography - of British Labour Party leader Neal Kinnock, without attribution. (*Breitbart*, Melania's plagiarism)

[55] "Just watch Vice President **Biden** learn about Weiner's involvement for the very first time," Oliver told viewers. **Biden** actually found out about Weiner's involvement on-air, and containing his exasperation required immense composure. (*AlterNet*, Weiner's sexting)

[56] Here is one possibility: The Democratic party in reality is the cartoon version of the Republican party stood on its head, with cold-eyed self-serving economic interests using the so-called social issues to stir up the rubes while they go about seeing to their own paydays and pensions. (*National Review*, DeVos as Secretary of Education)

[57] In its place stands **Trump** and the modern Republican Party, which together represent the most abhorrent forms of leadership and have an incredibly destructive political agenda. (*AlterNet*, Trump elected)

[58] The Republican Party still offers the best and only path to prosperity and freedom. America has experienced eight years of Obama and the **Democrats**, and the negative results are obvious. (*NewsMax*, McCain's healthcare bill)

[59] This, of course, is completely unacceptable to **Republicans**, but that's because they don't know the real history of the Second Amendment, and the real history of the Second Amendment is as ugly as it gets. (*AlterNet*, Las Vegas shooting)

[60] Thanks to corporate media's unquestioning regurgitation of **right-wing** talking points, most Americans think that Second Amendment is in the Constitution to protect the rights of individual **gun** owners from the government. (*AlterNet*, Las Vegas shooting)

[61] And for decades, we propped up these **right-wing** authoritarian regimes, which only perpetuated the bad governance and grinding poverty. (*The Nation*, refugee crisis)

The Semantic Domain of 'Elections' in Extreme Extent Texts

The next semantic group that emerged as key for extreme extent texts relate to elections and also include words from the semantic domains described above. Namely, *republicans*, *democrats*, *party*, *Clinton*, and *Trump* are also in this semantic domain as they correlate with *campaign* and *votes*. *Whites*, though seemingly odd, is included here as an examination of the texts demonstrates that it is often referring to 'whites' as a voting demographic. These keywords show that texts directly related to elections are prominently connected to extreme bias. The text excerpts below provide examples of this semantic domain in context with keywords presented in bold and related words underlined.

[62] Kushner, who has been assigned a rather large array of tasks as a White House senior adviser, has faced plenty of scrutiny in the multiple investigations into whether **Trump's campaign** colluded with Russia last year. (*HuffPost*, Fire and Fury book release)

[63] "In theory, everything Mueller is investigating the Russians for could be applied to the whole Fusion GPS/**Clinton campaign** DNC operation," said Tom Fitton, who directs Judicial Watch (*The Washington Times*, Russia probe)

[64] **Republicans** carried the state in six consecutive elections, from 1968 to 1988. But its non-white population grew steadily throughout. By 1992 it was essentially split 50\50 between minorities and **whites**. **Democrats** won five of the next six elections **Republicans** have failed for the past half century to attract non-white **votes**. Richard Nixon won 32 percent of non-**white's** **votes** in 1960. (*National Review*, 2016 presidential elections)

[65] That apparently was the case in Florida, and also in Georgia, where exit polls found 85 percent of black and Latino **votes** were for **Clinton**, while 80 percent of **whites** voted for Trump. (*AlterNet*, Trump elected)

The Semantic Domain of 'Legal Issues' in Extreme Extent Texts

The semantic domain of legal issues contains only one word: *FBI* and a secondary consideration of the word *illegal*. This domain likely emerged as key in extreme extent texts as many of the topics relate to the myriad investigations that were occurring between 2016 and 2018. Namely, the Russia investigation, the investigation into Clinton’s emails, the investigation of Anthony Weiner, the investigations into Trump, the firing of Robert Mueller and James Comey, and the appointment of several justices (including Kavanaugh). Thus as was noted above, some of the results are more related to time than explicit bias. The text excerpts below provide examples of this semantic domain in context with keywords presented in bold and related words underlined.

[66] “The New York Times, like establishment media in general, is an extension of the **Clinton Campaign**, the Democratic Party and their global special interests,” the **Trump campaign** said in a statement about the publishing of the documents. “What is happening now with the **FBI** and DOJ on Hillary **Clinton**’s emails and **illegal** server, including her many lies and her lies to Congress are worse than what took place in the administration of Richard Nixon, and far more **illegal**.” (*Breitbart*, Trump’s tax return)

[67] I could pick dozens of examples, but to me the scariest was the day **Trump** went after the **FBI** director he’d just fired, James Comey, and warned him not to leak damaging information about their conversations because there might be “tapes.” (*The Nation*, Covfefe)

[68] #HidingHillary has been avoiding unscripted questions and appearances that aren’t carefully controlled by her **campaign**. Yesterday’s release of information by the **FBI** on their investigation into her use of a private email server and handling of classified information while she served as Secretary of State has raised many questions. (*Breitbart*, Clinton’s emails)

The Semantic Domain of ‘Social Issues’ in Extreme Extent Texts

The last semantic domain that emerged for extreme extent texts relate to social issues. Unsurprisingly these are a mix of the social issues that were key in the left leaning and right leaning groups with *LGBTQ*, and *gun (control)* being from left rated texts, and *illegal immigration* being from the right rated texts. These issues likely emerge as key for the extreme

extent group as these issues garner heavy emotional opinion and can easily lead to attacking the opposition. Interestingly, *gun* correlates to *control*, *controlled*, and *ban* suggesting that right leaning texts are adding to this emerging as key for this group. Indeed, excerpts [71] & [72] demonstrate a clear right bias on the *gun* debate while [73] – [75] demonstrate a clear left bias. Excerpt [76] – [79] exemplify issues related to *immigration* while [69] and [70] exemplify *LGBTQ* concerns. Keywords are presented in bold while related words are underlined.

[69] Among other things, Grasz, 56, pushed to amend the Omaha City Charter in 2013 to let employers discriminate against **LGBTQ** people, and he served on a nonprofit board that backed so-called conversion therapy for **LGBTQ** kids. (*HuffPost*, Trump’s judicial picks)

[70] who, as governor of Indiana, pushed through a bill permitting discrimination against **LGBTQ** people; who tried to force women to hold funerals for aborted or miscarried embryos. (*The Nation*, 2016 presidential election)

[71] If Hillary cares so much about the issue, she might take ten minutes to learn something about it, but **gun**-controllers tend to be low-information advocates. (National Review, as Vegas shooting)

[72] While many **Democrats** long have been labeled as anti-**gun**, firearms advocates say Mrs. **Clinton**'s position is especially disingenuous. "This is an outright and outrageous lie that [Mrs. **Clinton**] doesn't want to take people's **guns**," said Dudley Brown, president of the National Association for **Gun** Rights. (*The Washington Times*, gun control)

[73] The "Second Amendment" as we know it today is a legal fiction invented by the **gun** industry and their buddies on the Supreme Court and sold to Americans by an expensive multi-decade-long PR campaign. Despite what you might hear on Fox So-Called news, there actually was no “individual right to own a **gun**” until 2008, (*AlterNet*, Las Vegas shooting)

[74] The gun-control movement needs a left flank that can take to the streets, resists calls for polite dialogue and demands real change. (*The Nation*, Orlando shooting)

[75] Let’s get 1,000 LGBTQ folks to the NRA headquarters in Fairfax, Virginia, and wrap it in a huge Gilbert Baker rainbow flag and with glitter and feathers shut that building down. And do it again and again. Let’s get waves of **LGBTQ** folks to take rainbow flags and sew shut the doors of offices of senators and congresspeople who have voted over and over against sane limits on **guns** in the hands of everyday Americans. (*The Nation*, Orlando shooting)

[76] Unfortunately, many **Democrats** see **illegal immigration** as a voter registration drive. They aid, abet and encourage **illegal** aliens to enter and stay in America in the guise of charity, but their actions are all about politics. (*Fox*, 2018 midterm elections)

[77] Not one **illegal**-alien child has reached the U.S.–Mexican **border** because Donald J. **Trump** ordered him there. (*National Review*, child immigrants separated at border)

[78] Obama Administration placed **border children** with human smugglers The Obama administration failed to protect thousands of Central American **children** who have flooded across the U.S. **border** since 2011, leaving them vulnerable to traffickers and to abuses at the hands of government-approved caretakers, a Senate investigation has found. (*The Washington Times*, child immigrants separated at border)

[79] **Republicans** are breaking ranks with the president because the optics of crying children and mothers isn't selling well with the majority of Americans. This is precisely why the **Democrats** have suddenly decided they don't want to discuss **immigration** at all. (*NewsMax*, child immigrants separated at border)

Taken together, keywords for extreme extent texts demonstrate that issues relevant to both right and left political parties emerge as well as labels and political actors specific to each. As was the trend in the left leaning and right leaning texts, the labels are clearly being used to refer negatively to the opposing political party and the actors focused on are those from opposing political party showing that these texts in this maintain clear coverage and presentation biases. In addition, as was noticed above, some of the keywords emerge as a product of the years of interest in this study, specifically legal issues and elections, which suggests that in examinations of lexical bias, time must be considered.

9.4.2 No Extent Texts

A total of 10 words remained key for the no extent group after both keyness methods were conducted. The words, as well as the keyness value for dispersion and frequency, the number of texts the word is dispersed across, and the word's total frequency are presented in table 9.10 below.

Table 9.10

Keyword List for Texts Rated as No Bias

Frequency	# of Texts	Frequency Keyness Value	Dispersion Keyness Value	Word
115	31	84.36	5.14	fire
60	19	59.72	4.99	California
37	23	42.71	8.76	residents
69	25	38.77	3.86	Sunday
62	41	35.01	5.50	South
41	17	29.25	4.16	athletes
38	17	22.5	4.87	expected
71	30	21.83	3.92	team
33	23	21.35	6.32	authorities
67	23	21.15	5.36	online

The major semantic domains for the no extent group are perhaps the most clear as they relate directly to who, what, when, and where or the ‘5 Ws of journalism’ (minus the why). Importantly, the four Ws that emerge here are the four most directly connected to factual reporting and are the least likely to require interpretation or analysis (as the why category might have done). Thus, the keywords identified in the no bias extent are unsurprising. Table 9.11 lists the keywords under each semantic domain, and each is discussed in more detail below.

Table 9.11
Semantic Domains for No Bias Keywords

Who	What	When	Where
Residents	fire	Sunday	California
Athletes	expected		south
Team			online
Authorities (fire)			

The Semantic Domain of ‘Who’ in No Bias Texts

The first semantic domain that emerges as key for no extent texts is who which includes: *residents, authorities, teams, athletes*, and a secondary consideration of *fire* as it was a premodifier for many groups referenced in these texts (e.g., **fire** department, **fire teams**, etc.). As shown in the text excerpts below, these keywords exemplify that no bias texts often focus on who the newsworthy event is affecting and/or who is connected to the event. In addition, it

exemplifies who journalists are collecting information from in their reporting. Keywords are presented in bold.

[80] President Donald Trump issued a warning to **residents** of North and **South** Carolina on Monday, warning them about the looming hurricane off the East Coast. (*Breitbart*, Hurricane Florence)

[81] A coalition of law firms calling itself Northern **California Fire** Lawyers filed a negligence lawsuit Tuesday on behalf of more than a dozen **residents** affected by the blaze... (*Wall Street Journal*, CA Fires)

[82] While the program rode an initial wave of popularity across Rio, recent years have seen a rise in violence, corruption schemes and deaths of suspects, law enforcement and **residents** caught in between. (*USA Today*, Rio Olympics)

[83] Wyman who was recently diagnosed with ALS and traveled 11 hours to see the eclipse, and Charlottesville **residents** Beverly Ingram and Doreen Bonnet who said the eclipse brought people in their community together. (*NPR*, Solar eclipse)

[84] During a press conference, officers also encouraged anyone with cellphone footage of the shooting to turn it in to the **authorities**. (*The Atlantic*, Las Vegas shooting)

[85] **Authorities** are focused on the possibility that the attack was connected to a bombing that took place 11 hours earlier in New Jersey. (*NewsMax*, NYC bombing)

[86] The incident is being investigated internally as well as by the local district attorney, and federal **authorities** have also agreed to investigate (*NPR*, Black Lives Matter protests)

[87] Police and **fire teams** forced their way into some rooms to free the **residents**, who were temporarily taken to the town's center for public safety. (*USA Today*, Hurricane Florence)

[88] **California** Highway Patrol Chief Brent Newman asked for the public's patience as **teams** clear affected areas. As part of those efforts, CHP teams removed 60 abandoned cars and set up barriers around dangerous areas. (*CNN*, CA fires)

[88] creating a Refugee Olympic **Team** that would be “treated ... like all the other [national] **teams**,” in having those **athletes** march into the Opening Ceremony right ahead of host country Brazil... (*The Atlantic*, Rio Olympics)

[89] The U.S. **team** includes 10 African-Americans, 11 Asian-Americans and its first two openly gay **athletes**. Four of the five **athletes** on the U.S. women's bobsled **team** are people of color, led by Elana Meyers Taylor. (*NPR*, Winter Olympics)

[90] The International Olympic Committee delivered a reprimand and the Anti-Defamation League was among groups who issued statements regarding hostilities toward Israeli **athletes** at the Olympic games in Rio, (*NewsMax*, Rio Olympics)

[91] The president's attack on **athletes** turned the anthems - usually sung during commercials - into must-watch television shown live by the networks and Yahoo!, which streamed the game in London. (*AP*, Take a knee)

[92] The New York City **Fire** Department examined the damaged buildings and determined they were safe. (*Wall Street Journal*, NYC bombing)

The Semantic Domain of 'What' in No Extent Texts

The next semantic domain that came out as key in no extent texts relate to what the stories are reporting on. The words include *fire* and *expected*. *Fire* came out as key due to the articles on the CA fires, the Fire and Fury book release, and many of the articles on the several shootings ending up in the no extent group. *Expected* was included in the 'what' category as its collocates are *is* and *was* indicating that *expected* is always connected to the reporting of an event – either one that predictions are being made about as in [96] – [99] or an event that happened that either matched expectations or doesn't match expectations as in excerpts [100] – [103].

[93] A gunman opened **fire** during a country-music show at the Mandalay Bay Resort and Casino in Las Vegas, killing at least 58 people and injuring hundreds more. (*The Atlantic*, Las Vegas shooting)

[94] Citi analyst Praful Mehta estimated the total damage caused by the current **fire** could reach \$15 billion, comparable to the 2017 fire damages, though PG&E wouldn't necessarily bear those costs in full if found liable for the blaze. (*Wall Street Journal*, CA fire)

[95] Hours after Omar Mateen opened **fire** June 12 on a packed gay nightclub in Orlando, the Westboro Baptist Church, a well-known anti-LGBTQ group, tweeted "GOD SENT THE SHOOTER to #Pulse in Orlando! The murderer is in hell! (*USA Today*, Orlando shooting)

[96] The Camp **Fire** is **expected** to exacerbate housing problems in Butte County, where the flames decimated the housing stock. (*NPR*, CA fires)

[97] Later this week, life-threatening storm surges -- up to 13 feet -- are **expected** along the coasts, and up to 35 inches of rain could fall through early next week over parts of the Carolinas and Mid-Atlantic states. (*CNN*, Hurricane Florence)

[98] Florence is **expected** to turn toward the north and go through the Ohio Valley by Monday, according to the weather service. (*NPR*, Hurricane Florence)

[99] About 100 people are still unaccounted for statewide, authorities said Monday. Officials are hoping those people are safe elsewhere, but with wildfires still burning, the death toll is **expected** to rise. (*CNN*, CA fires)

[100] Pompeo left Pyongyang for Japan early **Sunday** afternoon, but it wasn't immediately clear if he met with Kim Jong Un, as had been **expected**. (*USA Today*, Trump and North Korean disarmament)

[101] Voter tallies for the House races, as calculated by The New York Times, show that more Democrats voted than Republicans. That was **expected**, given Democratic outrage at the Trump administration. (*NPR*, 2018 midterm elections)

[102] Investors were largely expecting a victory for Mrs. Clinton and noted that a market shock--similar to the gains the day after the U.K. voted to leave the European Union--was **expected** if Republican nominee Donald Trump takes the White House. (*The Wall Street Journal*, 2016 presidential elections)

[103] It has a complication of microcephaly (a birth defect where the baby's head is much smaller than **expected**). (*USA Today*, Zika virus)

The Semantic Domain of 'When' in No Extent Texts

The semantic domain of 'when' emerged as key for no extent texts with only the one keyword of *Sunday*. It is unclear why *Sunday* is the day that emerged as key for the no extent group when there are many other days reported on in these texts; however, at the very least it demonstrates that in comparison to extreme extent texts, no extent texts report more about events on Sundays. Text excerpt [104] – [108] below demonstrate how no extent texts use Sunday – which is presented in bold.

[104] A second device, possibly a pressure cooker, was found four blocks away on West 27th Street and was safely removed early **Sunday**, according to the New York Police Department. (*USA Today*, NYC bombing)

[105] Mrs. Trump and Barron, 11, accompanied the president from the Trump National Golf Club in Bedminster, N.J., where he spent the weekend, arriving via Marine One on

the White House lawn shortly after 8 p.m. **Sunday**, according to the presidential pool report. (*USA Today*, Melania moves to White House)

[106] Most NFL players on **Sunday** locked arms with their **teammates** - some standing, others kneeling - in a show of solidarity. (*AP*, Take a knee)

[107] An epic Washington political battle took shape on **Sunday** after the death of Justice Antonin Scalia as Senate Republicans dug in and said they would refuse to act on any Supreme Court nomination by President Obama. But the White House vowed to select a nominee within weeks. (*The New York Times*, Scalia's Death)

[108] Deadline reported Monday that the final tally for the film was \$103.1 million, even higher than the initial **Sunday** estimates. (*NPR*, Wonder Woman movie)

The Semantic Domain of 'Where' in No Extent Texts

The last semantic domain for keywords of no extent texts is 'where' which is composed of the keywords *California, online, and south*. As the text excerpts below demonstrate, these words emerged as key considering the topics that were rated as no extent include: the California fires, Hurricane Florence, net neutrality, the Winter and Rio Olympics, Fire and Fury book release, Aretha Franklin's death, Amazon buys Whole Foods, etc. These locations are directly relevant to these topics. The text excerpts below exemplify these keywords in context with keywords presented in bold.

[109] The "Queen of Soul," who died last week at 76, "instilled hope, uplifted generations, and changed the lives of millions," according to the bill introduced Tuesday by Democratic Sen. Kamala Harris of **California** (*NewsMax*, Aretha Franklin death)

[110] He added that Zervos had emailed Trump's office this past April 14 to ask if he could visit her restaurant in **California**. (*Fox*, Trump sexual assault)

[111] Because of the severity of the Camp **Fire**, the **California** Department of Forestry and **Fire** Protection has produced an **online** structure status map of Paradise and surrounding areas (*NPR*, CA fires)

[112] But in an article titled "The Withering **California** Dream, by the Numbers," Chetty found that 30-year old Californians were less likely in every decade to achieve the American Dream (*Breitbart*, Millennials and the economy)

[113] Several states, including New Jersey, Washington, Oregon and **California**, have gone so far as to push legislation to enforce the principles of net neutrality within their borders. (*CNN*, net neutrality)

[114] The storm has caused more than 770,000 power outages in North Carolina, with more than 150,000 reported in **South** Carolina... (*NPR*, Hurricane Florence)

[115] The declaration comes as China, the region's other superpower, continues to make neighbors nervous by unilaterally developing what most analysts agree are military installations on disputed reefs in the **South** China Sea, one of the world's busiest shipping channels. (*CNN*, China as a superpower)

[116] North Korea has been showering the United States and **South** Korea with goodwill gestures in recent months, including the shutdown of its main nuclear testing site and the release of three American detainees. (*AP*, Trump and North Korean disarmament)

[117] When Korean athletes from both states marched under the same flag — a “unification” flag — this was highly meaningful to the **South** Koreans watching. (*National Review*, Winter Olympics)

[118] Walmart, which has the largest share of the U.S. food market, has already been pushing harder into e-commerce to build on strength in its stores and groceries. It announced Friday that it’s buying **online** men’s clothing retailer Bonobos for \$310 million, following a string of **online** acquisitions. (*The Washington Times*, Amazon buys Whole Foods)

[119] Ms. Griffin went **online** on Tuesday night to express regret. “I sincerely apologize,” she said in a video. (*The New York Times*, Kathy Griffin fired)

The keywords identified for the no extent texts are those one would expect of news. And while there were differences in text topics for no extent and extreme extent texts, they are still considered as variation in extent of bias. The reason for this is that the differences identified in topic, publication, and even article type (refer back to chapter 6) are directly related to impartiality and the features that emerge here are those one would expect from an impartial news report.

9.5 Conclusions

The keyword analysis identifies some important differences in lexical patterns in left leaning versus right leaning texts as well as extreme extent versus no extent texts. For left

leaning versus right leaning, a clear difference emerges in that left leaning texts are more concerned with social issues while right leaning texts are more concerned with political issues. Specifically, left leaning texts demonstrate a higher concern for issues related to equality while right leaning texts demonstrate a higher concern for illegal immigration, terrorism, and loss of freedom. In addition, both left and right leaning texts contain semantic domains related to people or actors, but left leaning texts are concerned with Trump and Muslims while right leaning texts are concerned with past presidents and creating an ‘us vs. them’ attitude with their group labels. Lastly, left leaning texts contain a number of keywords demonstrating their aggression and general unhappiness about the social issues related to equality. All in All, the words that emerge as key for both left and right leaning groups match the common consensus of left and right (or liberal and conservative) concerns.

Keywords for extreme extent vs no extent texts were also very telling. The no extent group contained keywords that were clearly connected to who, what, when, and where which were all very clearly connected to facts of events rather than analysis. Extreme extent texts contained 71% overlap with left leaning and right leaning texts so social and political issues, particularly those related to immigration, gun control, and LGBTQ rights, emerged as key in this group. In addition, previous presidents as well as several labels for political groups demonstrated a tendency for extreme extent groups to display an “us versus them” sentiment or to shift blame. Legal issues related to FBI investigations as well as elections were also key for the extreme extent group, but in most texts, there was a very clear democrat versus republican sentiment. This was not at all present in the no extent group.

This analysis demonstrates what a keyword analysis can show us about news bias, and specifically demonstrates which actors and issues are focused on – something that will be

analyzed further in relation to coverage and presentation bias in the next chapter. Something important to note here is that keywords for extreme vs no extent are very specific to time suggesting that examinations of lexical bias need to consider and control for time. While some words are likely to continuously be connected to bias (e.g., right-wing or left-wing), some words will be connected to bias during a specific time period (e.g., Trump, Clinton's emails, the wall).

CHAPTER 10. SYNTHESIS AND CONCLUSIONS

10.1 Introduction

The purpose of this chapter will be to synthesize the results into a succinct description and clearly provide answers to the research questions (section 10.3). It will also connect the results here back to the three frameworks of media bias identified in chapter 2 (gatekeeping, coverage, and presentation) and explain how the methods of this study can be used to examine the three types of bias (section 10.2). Section 10.4 will discuss limitations and future research directions and section 10.5 will provide conclusions.

10.2 Orientation to Media Bias Studies

10.2.1 Gatekeeping Bias

In this study, gatekeeping bias was examined per publication via a presence and absence of texts collected on each topic. As described in chapter 4, 25 trending and newsworthy topics were collected per year of observation for a total of 75 topics. Publications were systematically searched for articles on each topic and when the search yielded no eligible results, it was recorded as an absence. Occasionally, a publication posted an article taken from a different source (e.g., *The Nation* posts an article taken directly from *Tomsdispatch.com*). These articles

were collected as a presence when they were the only eligible text on a topic as they are displayed to readers and thus are still representative of the publication of interest.

Figure 10.1 below is a visual illustration of how gatekeeping bias was measured. The total row on the righthand side represents the total amount of texts collected on that topic while the total row at the bottom represents the total amount of texts collected per publication. A 1 indicates a presence of that article topic while a 0 indicates an absence.

The publications that published the least were *The Nation* (48), *National Review* (63), *AlterNet* (67), and *NPR* (67) while the publications that published on every topic include: *HuffPost* and *The New York Times*. The topics that were ignored the most were China as a superpower, Melania's move to the White House, the Royal wedding, Kanye West on politics, gay marriage, and the 2017 Russia investigation while the topics published on the most include: Black Lives Matter, ISIS, Orlando shooting, 2016 presidential elections, refugee crisis, Rio Olympics, Scalia death, Trump (sexual assault), Trump elected, Turkey coup, Zika Virus, Comey testimony, Covfefe, DeVos, Las Vegas shooting, McCain healthcare vote, Robert Mueller, Taking a knee, Trump's travel bans, Trump's national security strategy, Women's March, Wonder Woman movie, Black Panther (movie), Bush death, child immigrants, Fire and Fury book, Aretha Franklin Death, Hurricane Florence, Kavanaugh Supreme Court, March for our Lives, 2018 Midterm Elections, Nike & Kaepernick, Roseann Barr, Stormy Daniels, Trump 'very stable genius', Trump and NK disarmament, Trump's tariffs, Winter Olympics, and the World Cup.

	AlterNet	The Nation	HuffPost	CNN	The Atlantic	NYT	USA Today	NPR	Associated	WSJ	Fox	TWT	Breitbart	NewsMax	National Re	Total
Anthony Weiner texts	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Black Lives Matter	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Brexit	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
China as a Superpower	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	9
Clinton Emails	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	13
Gay Marriage	1	1	1	0	1	1	1	1	1	0	1	1	1	0	0	11
Global Warming hoax?	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Gun Control debates	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	13
Harriet Tubman on \$20	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
ISIS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Melania Trump's Plagiarism	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
NYC Bombings	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Orlando Shooting	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Penn El Chapo Interview	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Presidential Elections	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Prince's Death	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Refugee Crisis	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Rio Olympics	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Saudi Arabia 9/11 Bill	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	13
Scalia Death	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Trump (sexual assault)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Trump Elected	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Trump's Taxes	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Turkey Coup	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Zika Virus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Amazon Buys Whole Foods	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Comey Testimony	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Covfefe	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
DeVos	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Kathy Griffin	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Kennedy Center Honors	0	0	1	1	0	1	1	1	1	1	1	1	1	1	0	11
Las Vegas Shooting	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
McCain Healthcare vote	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Melania moves to WH	0	0	1	1	0	1	0	1	0	1	0	0	0	1	0	7
MeToo	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	14
Millennials and the economy	1	0	1	1	0	1	0	0	1	1	1	1	1	1	1	11
Pence Mideast Visit	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	14
Pepsi Commercial	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	12
Robert Mueller	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Royal Wedding/ Markle	1	0	1	1	1	1	0	0	0	1	1	1	1	1	0	10
Russian investigation	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
SNL	1	0	1	1	1	1	1	0	1	0	1	1	1	1	0	11
Solar Eclipse	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Taking a Knee	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Travel Bans	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Trump's judicial picks	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Trump's National Security Strategy	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
United drags man off Plane	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Women's March	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Wonder Woman movie	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Black Panther (movie)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Bush death	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
CA Wildfires	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	13
Child Immigrants	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Fire and Fury book	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Franklin Death	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Hurricane Florence	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Kanye West Politics	1	0	1	1	1	1	0	1	0	1	1	1	1	1	0	11
Kavanaugh Supreme Court	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
March for our Lives	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Midterm Elections	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Mira Ricardel Firing	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	13
Musk's Tweets	1	0	1	1	1	1	1	1	1	1	1	0	1	0	0	11
Net Neutrality	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	12
Nike & Kaepernick	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Roseann Barr	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Russian Inquiry	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	12
Serena Williams Tennis sexism	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	13
Stormy Daniels	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Taylor Swift Politics	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Trump 'Very Stable Genius':	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Trump and NK disarmament	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Trump's Tariffs	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Winter Olympics	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
World Cup	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Total	67	48	75	74	68	75	73	67	72	69	74	72	72	71	63	

Figure 10.1 Gatekeeping bias per publication

In terms of gatekeeping bias, there are no overt patterns clearly related to leaning or extent of bias. There is a definite tendency for all publications to publish on topics that are political in nature or extremely divisive adding to the idea that news tends to feed into sensationalism (Ravenelle et al., 2021). This likely relates more to business interests and attracting an audience than it does to political ideology. The publication with the least publications, *The Nation*, is due to the fact of monthly rather than weekly or daily publishing while the publications with the most publications on these topics are certainly the most visible.

10.2.2 Coverage Bias

Coverage bias, or how much a topic or entity is discussed, was examined via the keyword analysis particularly the frequency method for publication and the dispersion method for the entire group. For left leaning texts, there was a clear focus on Trump, Muslims, and social issues related to health, gun control, and equality (LGBTQ, human rights, and reproductive rights). In addition, there was a clear focus on political unrest due to the aforementioned issues. In right rated texts, issues, actors, and groups that were political in nature were focused on much more. Particularly illegal immigration, terrorism, a loss of freedoms, and the actions of former presidents were of major concern.

In extreme extent texts, an amalgamation of political and social issues from right and left leaning groups emerged as key; specifically, social issues of LGBTQ rights and the political issues of gun control, gun banning, and illegal immigration were focused on more than other issues. In addition, legal issues related to FBI investigations as well as election issues emerged as key. This demonstrates that articles related to these topics lend themselves easier to further extents of bias. The no extent group showed a clear focus on who, what, when, and where in

their reporting suggesting that the extreme extent group relied less heavily on this type of basic information that is more generally connected to factual reporting.

Because the keyword analyses relied upon here statistically highlight words that are both content generalizable and content distinctive for each bias group, they are good ways to examine coverage bias. By this method, no manual measurement or decisions about specified issues or people are required. As long as the collection of texts is fair and representative, a keyword analysis will demonstrate whether certain entities, actors, or issues are being focused on more or less than another.

10.2.3 Presentation Bias

Presentation bias, or how a topic is discussed, was examined via the key feature and keyword analyses. The keyword analyses do not only indicate which topics or entities are focused on in a group of texts but can also highlight evaluations and sentiment – or how something is talked about. The keyword analyses in the present study demonstrated a clearly aggressive sentiment in left leaning texts in the political unrest semantic domain which suggested a clear unhappiness with Trump and the general status of social issues in the US. The political group labels used in right leaning texts demonstrate a negative evaluation towards ‘them’ who in this case were liberals or the left. The extreme extent keyword analysis showed a clear division between republican and democrat ideology with each side contributing texts depicting a negative assessment of the opposing side, and the no bias group demonstrated an absence of sentiment or evaluation.

The key feature analysis clearly demonstrates that there are differences in how both left and right leaning texts use grammatical structures that slyly reveal their opinions with the right relying on communication verbs and reported speech and the left relying on non-finite clauses

that make the statements less explicit. Both are grammatical ways that allow the author to avoid accepting responsibility for the expressed sentiment. Right leaning and left leaning text also differ grammatically in their tone with the right being much more colloquial and reliant on quoted material and the left being more formal and employing a wider range of vocabulary. This is further supported by the fact that the key grammatical features on the right display more verbal communication, shorter prose, and higher reliance on coordination while the key grammatical features on the left are more nominal, verbose, and express stance with non-finite clauses. Lastly the right and left leaning groups differ in their interaction and involvement in that the right leaning group is involved at the individual level (I, me, myself, etc.) and the left leaning group is involved at the group level (We, us, our, etc.). This fits well with the finding that left leaning texts are concerned more with social issues while right leaning texts are concerned more with political issues/fears.

The key feature analysis for extreme bias texts reveal that features of epistemic and attitudinal stance, clausal and phrasal elaboration, description, emphasis, and evaluation are key while key features for texts rated as no bias include features related to reporting what happened (who, what, when, and where). Perhaps unsurprisingly, many features in texts rated as extreme extent are related to stance whether directly or indirectly and no features in texts rated as no extent were related to stance. This indicates that extent of bias functionally explains these key features well as features of description, elaboration, emphasis, and evaluation have all been identified in previous literature as ways to reveal bias.

As demonstrated, the keyword and key feature analyses taken together can reveal a lot about presentation bias. Negative sentiment, evaluation, and demonstrations of stance will reveal themselves if they're occurring at a statistically higher rate in one group over another. This

method again limits the need for manual coding and allows the words and features examined for bias to be revealed by the data rather than predetermined by the researcher.

10.3 Synthesis of Results & Answers to RQs

As a reminder, the research questions that guided this analysis are:

- (3) Are there systematic linguistic differences among texts rated as biased in different directions?
 - a) Are there differences in key words?
 - b) Are there differences in key grammatical features?
- (4) Are there systematic linguistic differences among texts rated as biased to different extents?
 - a) Are there differences in key words?
 - b) Are there differences in key grammatical features?

To answer these questions, *SPIN* was created to carefully represent news texts across a political cline of left and right. Next, raters, who represent the average online newsreader, were asked to read the texts and rate them on a scale of left leaning to right leaning as well as a scale of extreme extent to no extent. Once texts were rated, ratings were aggregated into one mean score per bias type (extent and direction), and texts were separated into groups of: left leaning, right leaning, extreme extent, and no extent. Next a key feature analysis was run to examine the grammatical differences in left leaning and right leaning as well as no extent and extreme extent texts. Then a keyword analysis was run on the same groupings to examine lexical differences among groups. Finally, results were analyzed for functional and evaluative characteristics to determine what they suggest about each type of bias.

The analysis, summarized together in tables 10.1 and 10.2 below, demonstrates that there are indeed systematic linguistic differences, both lexical and grammatical, among texts rated as bias in different directions and to different extents. Key features of left leaning texts show that this group maintains a formal tone, involvement at the group level, and covertly reveals stance

via non-finite clauses while the keywords reveal a concern for social issues, a sentiment of aggression and political unrest, and an overfocus on Trump. Conversely, right leaning texts are colloquial in tone, involved at the individual level, and covertly reveal stance via reported speech while the keywords demonstrate a concern for political issues, the actions of former presidents, and a negative evaluation of the ‘liberal left’. Interestingly, the use of ‘othering’ through lexical and pronominal use and negative sentiment appears in both of these groups. These two features were identified repeatedly in linguistic and NLP studies as functional displays of political bias and Fake News.

Extreme extent texts include features marked for epistemic and attitudinal stance, clausal and phrasal elaboration, description, emphasis, and evaluation and the keywords reveal a concern for both political and social issues, political actors, and opposing negative evaluations of both democrats and republicans. Key features and keywords for texts rated as no extent include features related to reporting what happened (who, what, when, and where). The differences are, for the most part, attributable to a functional or evaluative difference related to direction or extent of bias. The keyword and key feature analyses in this study are clearly complimentary and reveal much about the bias in each group. To avoid repetition, the details of this will be discussed in the next section.

Table 10.1
Combined Results for Left and Right Bias Groups

	<u>Key Features</u>	<u>Keywords</u>	<u>Semantic Domains</u>
Left Bias	To' Complement Cl: Stance N	Trump	Social and political issues
	Verbs: Att	Russia	People
	Type Token Ratio	anti	Political unrest
	Word Count	health	
	1st ProN (pl)	LGBTQ	
		resistance	
		violence	
		nation	
		Muslims	

		power war gun protest act rights	
Right Bias	That' Complement Cl: Comm Verb That' Complement Cl: Stance Verb Conjunctions: Adverbial 1st ProN (sing.)	wall illegal national Obama left liberals Republicans emails Bush terror Reagan terrorism immigration Clinton freedom	Political issues Political actors (former presidents) Political groups

Table 10.2

Combined Results for Extreme and No Bias Groups

	<u>Key Features</u>	<u>Keywords</u>	<u>Semantic Domains</u>
Extreme Bias	WH' Compl. Clause: Verb of Likelihood Subordinating Conjunction: Conditional Modals of Necessity Verb of Likelihood: Other 'That' Complement Clause: Att N Coordinating Conjunction: and Adverbs: Downtoner Adjectives: Predicative Attitudinal Conjunctions Adverbs: Likelihood Nouns: Cognitive Adjective: Predicative Epistemic Adjective: Predicative Emphatics Nouns: New Stance Word Count	Clinton Trump Bush FBI campaign gun children republicans illegal Whites Democrats votes immigration border party LGBTQ Biden right-wing Obama	Political actors Elections Legal issues Social and political issues
No Bias	Adverbs of Place Adjectives of Time Verbs: Public	fire California residents	Who What When

Verbs: Past Tense
Nominalizations

Sunday
South
athletes
expected
team
authorities
online

In sum, the answer to the question of can we linguistically identify a biased news article, the answer is yes, we can. All in All, the words that emerge as key for both left and right leaning groups match the common consensus of left and right (or liberal and conservative) concerns. Keywords for extreme extent vs no extent texts were also very telling. The no extent group contained keywords that were clearly connected to who, what, when, and where which were all very clearly connected to facts of events rather than analysis. Extreme extent texts contained 71% overlap with left leaning and right leaning texts so social and political issues, particularly those related to immigration, gun control, and LGBTQ rights, emerged as key in this group. In addition, previous presidents as well as several labels for political groups demonstrated a tendency for extreme extent groups to display an “us versus them” sentiment or to shift blame. Importantly, in both key feature comparisons, we see that there are differences in lexicogrammatical features based on direction and extent of bias, and that these features can be attributed to functional and stylistic characteristics of each group.

The results of this study will help contribute to our knowledge of the linguistic nature of political bias in media, and hopefully, facilitate determinations about deceptive and impartial news practices. This study further demonstrates that media bias, namely gatekeeping, coverage, and presentation bias can be examined via corpus linguistic methods and that results are functionally and lexically attributable to direction and extent of bias. Importantly, it also

demonstrates ways in which a researcher can avoid making subjective decisions about bias by relying on exploratory methods of register variation.

10.4 Limitations and Future Research

This study, though done to the best of the researcher's ability, is not without its limitations. First and foremost, as many studies of media bias are, it is limited by its definition and operationalization of the term *bias*. Further research needs to be conducted to determine how best to go about collecting perceptions of this term. In this study, two constructs were used for bias: direction and extent. Results show that while there is immense overlap between right, left, and extreme bias, they are separate senses of the term. This needs to be examined further in future studies. Additionally, perceptions of the leaning of the texts in the present study align very closely with perceptions from *Media Bias Fact Check*, *Ad Fontes Media*, and *AllSides* which suggests a universal perception of 'left' and 'right' bias. This also requires further examination.

This study was also limited by its examination of only one medium of news: online newspapers. While this was essential for an empirical comparison, an examination including a wider range and differing mediums of news would be interesting. For example, examinations and comparison of TV news or news aggregators could reveal something different about bias.

In addition, this study was limited by the size of *SPIN*. While the corpus was meticulously designed to control for bias and to be as representative as possible, the final product was less than the desired 100 texts per publication and less than 1,000,000 words. This further affected the examination of bias which was only conducted on 66% of the texts from *SPIN*. In an ideal world, there would be many more texts and many more publications. However, texts were manually collected to ensure that the collection of topic and time were meticulously controlled for which brings us to the next limitation. If the collection of texts were automated, the corpus

could have been much larger and the bias groups likely more evenly comparable. Thus, time spent writing a code for topic-controlled and time-controlled text collection would be well worth it.

Future research on this topic could take the key features identified here and apply them to another set of texts to see if the features continually emerge along these same clines of bias. If they do, then an automatic tagger can be trained. This could also be done for keywords, however, as this study demonstrated, many of the words were relevant to the time in which texts were collected. This suggests that ‘lists of biased words’ relied on by other media bias studies may be seriously flawed in their thinking that these lists are evergreen. If a similar study were conducted on another time, different words would likely emerge as key for bias suggesting that researchers need to be vigilant in their continuous examinations of ‘bias words’. Perhaps an exploratory study such as this one should always be an initial step. Another study could also be conducted examining the middle ground texts that were left out of this study to determine whether they demonstrate any patterns relative to ‘middle ground’.

10.5 Conclusions

This dissertation aimed to determine what linguistic characteristics can reveal about bias in news and whether variation in linguistic characteristics, namely grammatical and lexical features, happen systematically across a cline of political leaning and extent. Through an examination of key features and keywords in news texts divided into left, right, extreme, and no bias groups, this aim was accomplished. This study further demonstrates that media bias, namely gatekeeping, coverage, and presentation bias can be examined via corpus linguistic methods and that results are functionally and lexically attributable to direction and extent of bias. Importantly,

it also demonstrates ways in which a researcher can avoid making subjective decisions about bias by relying on exploratory methods of register variation.

Figure 10.2 below illustrates the process of media bias. This study began at the bottommost level of this image by presenting the finished product to consumers, asking them to identify extent and direction of bias, and then examining each type of bias (green boxes) via linguistic analyses.

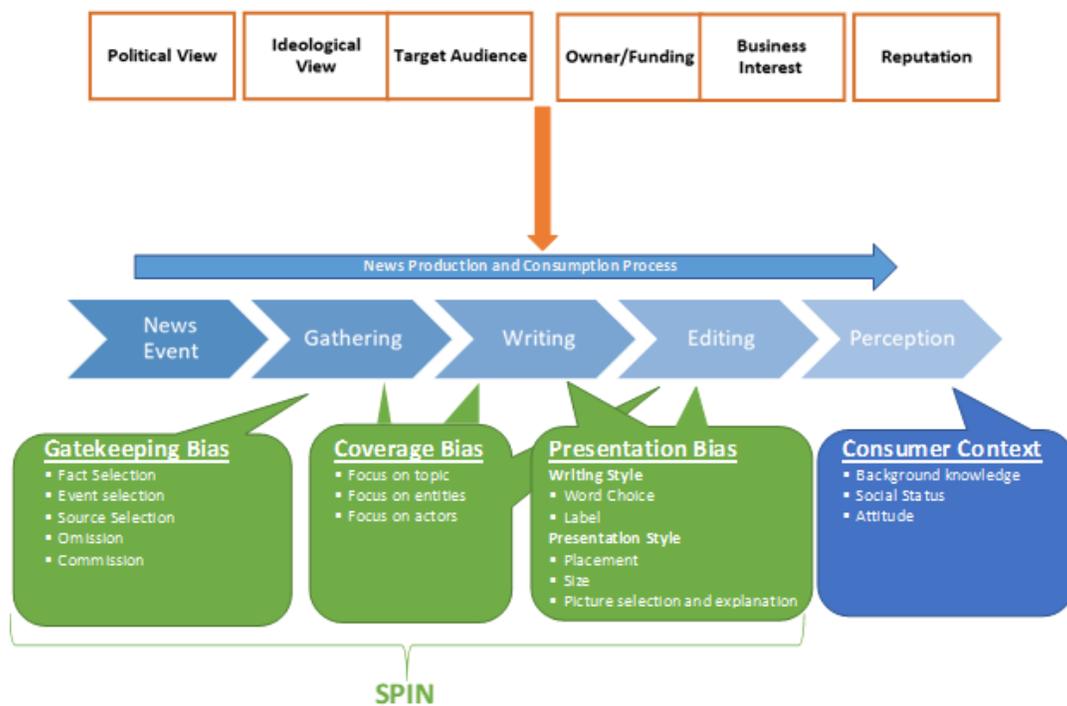


Figure 10.2 *Illustrated representation of media bias (adapted from Hamborg, Donnay, and Gipp, 2019).*

These analyses showed a clear connection back to political and ideological views with the left leaning, right leaning, and extreme extent comparisons. Each of the key features and keywords were attributable to some evaluative or functional reason related to political leaning or extent. It is also very likely that target audience, reputation, and business interest play a role as political leaning appears to be the way publications choose to communicate with their audience.

This is further supported by the fact that the most extremely rated publications (i.e., *AlterNet*, *HuffPost*, *Breitbart*, *NewsMax*, etc.) contributed the most texts to the extreme extent group and the left leaning and right leaning groups were composed of texts from left and right publications. Additionally, the no extent group tended to stick to presenting information as demonstrated by its key features and keywords, which is what an audience looking for unbiased news would expect. Further research is needed, but the results provided here are at least a good start in contributing to our knowledge of the linguistic nature of bias in news media.

REFERENCES

- Al-Gamde, A., & Tenbrink, T. (2020). Media Bias: A Corpus-based Linguistic Analysis of Online Iranian Coverage of the Syrian Revolution. *Open Linguistics*, 6(1), 584-600.
- Al-Hejin, B., 2015. Covering Muslim women: Semantic macrostructures in BBC news. *Discourse & Communication*, 9(1), pp.19-46.
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–36.
- Allsides.com (2019, May 28). Balanced news via media bias ratings for an unbiased news perspective. Retrieved July 10, 2019, from <https://www.allsides.com/unbiased-balanced-news>
- Baker, P. 2004. 'Querying keywords: questions in difference, frequency, and sense in keyword analysis, *Journal of English Linguistics* 32 (4), pp. 346–59.

- Baker, P., Gabrielatos, C. and McEnery, T., 2012. Sketching Muslims: A corpus driven analysis of representations around the word 'Muslim' in the British press 1998-2009. *Applied Linguistics*, 34(3), pp.255-278.
- Baker, P., Gabrielatos, C., Khosravini, M., Krzyżanowski, M., McEnery, T. and Wodak, R., 2008. A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), pp.273-306.
- Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., Van Der Goot, E., Halkia, M., ... & Belyaeva, J. (2013). Sentiment analysis in the news. *arXiv preprint arXiv:1309.6202*.
- Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting factuality of reporting and bias of news media sources. *arXiv preprint arXiv:1810.01765*.
- Bandy, J. (2020). Detecting Politically Biased Phrases from U.S. Senators with Natural Language Processing (Tutorial). Towardsdatascience .com
<https://towardsdatascience.com/detecting-politically-biased-phrases-from-u-s-senators-with-natural-language-processing-tutorial-d6273211d331>
- Barak, G., Leighton, P., & Flavin, J. (2007). Class, race, gender, and crime. Lanham, MD: Rowman and Littlefield.
- Baumer, E., Elovic, E., Qin, Y., Polletta, F., & Gay, G. (2015). Testing and comparing computational approaches for identifying the language of framing in political news. In *Proceedings of the 2015 conference of the North American chapter of the Association for Computational Linguistics: human language technologies* (pp. 1472-1482).

- Bednarek, M., & Caple, H. (2014). Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society*, 25(2), 135-158.
- Bell, A. (1983). Telling it like it isn't: Inaccuracy in editing international news. *Gazette (Leiden, Netherlands)*, 31(3), 185-203.
- Bell, A. (1994). Media (mis) communication on the science of climate change. *Public understanding of science*, 3(3), 259.
- Bell, A. (1995). News time. *Time & Society*, 4(3), 305-328.
- Bennett, W.L. and Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of communication*, 58(4), pp.707-731.
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political analysis*, 20(3), 351-368.
- Biber, D. (1993). 'Representativeness in corpus design'. *Literary and Linguistic Computing* 8/4: 243-57.
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2), pp.97-116.
- Biber, D. (2014). Using multi-dimensional analysis to explore cross-linguistic universals of register variation. *Languages in contrast*, 14(1), pp.7-34.
- Biber, D., & Egbert, J. (2018). *Register variation online*. Cambridge University Press.

- Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, 9(1), 93–124.
- Biber, D., & Conrad, S. (2019). *Register, Genre, and Style*. Cambridge University Press.
- Biber, D., & Zhang, M. (2018). Expressing evaluation without grammatical stance: informational persuasion on the web. *Corpora*, 13(1), pp.97-123.
- Biber, D., Egbert, J., & Keller, D. (2020). Reconceptualizing register in a continuous situational space. *Corpus Linguistics and Linguistic Theory*, 16(3), 581-616.
- Brady, D., & Ma, J. (2003, 16 November). Spot the Difference. *Wall Street Journal*, <https://www.wsj.com/articles/SB106860350224258900>
- Bryanov, K., & Vziatysheva, V. (2021). Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news. *PLoS One*, 16(6), e0253717.
- Brouwer, J., van der Woude, M. and van der Leun, J., 2017. Framing migration and the process of crimmigration: A systematic analysis of the media representation of unauthorized immigrants in the Netherlands. *European journal of criminology*, 14(1), pp.100-119.
- Carver, T. and Pikalo, J., 2008. *Political Language and Metaphor: Interpreting and Changing the World*.
- Chait, J. (2021, August 21). Why the Media Is Worse for Biden Than Trump. *Ney York Magazine*. <https://nymag.com/intelligencer/2021/08/why-the-media-is-worse-for-biden-than-trump.html>

- Chen, W. F., Wachsmuth, H., Al Khatib, K., & Stein, B. (2018, November). Learning to flip the bias of news headlines. In *Proceedings of the 11th International conference on natural language generation* (pp. 79-88).
- Chilton, P., 2004. *Analysing political discourse: Theory and practice*. Routledge.
- Choi, J. H., Watt, J. H., & Lynch, M. (2006). Perceptions of news credibility about the war in Iraq: Why war opponents perceived the Internet as the most credible medium. *Journal of computer-mediated communication*, 12(1), 209-229.
- Clifford, S., Jewell, R. M., & Waggoner, P. D. (2015). Are samples drawn from Mechanical Turk valid for research on political ideology?. *Research & Politics*, 2(4), 2053168015622072.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.
- Coll, S. (2022). *NEWSROOM CONFIDENTIAL Lessons (and Worries) From an Ink-Stained Life*. Macmillan publishing.
- Conboy, M. (2007). *The language of the news* (pp. x+-229). London: Routledge.
- Conrad, A. and Wiebe, J., 2012, July. Recognizing arguing subjectivity and argument tags. In: *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics* (pp. 80-88). Association for Computational Linguistics.
- D'Alessio, D., & Allen, M. (2000). Media bias in presidential elections: A meta-analysis. *Journal of communication*, 50(4), 133-156.
- Daly, K. (1997). Different ways of conceptualizing sex/gender in feminist theory and their implications for criminology. *Theoretical Criminology*, 1, 25-51

- Davis, J. (1990). News: beyond the myth of objectivity. *Center for media literacy*.
- Dean, M. (2011). *Democracy under attack: How the media distort policy and politics*. Policy Press.
- Deane, J. (2005). Media, democracy and the public sphere. *Media & Global Change: Rethinking Communication for Development*, 177-192.
- DeMarzo, P. M., Vayanos, D., & Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly journal of economics*, 118(3), 909-968.
- Dimock, M., Doherty, C., Kiley, J., & Krishnamurthy, V. (2014). Beyond red vs. blue: The political typology. *Pew Research Center*.
- Ditton, J. and Duffy, J., 1983. Bias in the newspaper reporting of crime news. *Brit. J. Criminology*, 23, p.159.
- Dooley, K.J., and S.R. Corman. 2002. "The Dynamics of Electronic Media Coverage." In *Communication and Terrorism: Public and Media Responses to 9/11*, ed. B.S. Greenberg, 121–35. Cresskill, NJ: Hampton Press.
- Druckman, J. N., & Parkin, M. (2005). The impact of media bias: How editorial slant affects voters. *The Journal of Politics*, 67(4), 1030-1049.
- Durant, A., & Lambrou, M. (2009). Language and media. *London/New York*.
- Eberl, J. M., Boomgaarden, H. G., & Wagner, M. (2017). One bias fits all? Three types of media bias and their effects on party preferences. *Communication Research*, 44(8), 1125-1148.

- Egbert, J. & Staples, S., 2017. Doing Multi-Dimensional analysis in SAS, SPSS, and R. In: Tony Berber-Sardinha and Marcia Veirano Pinto (Eds.), *Multidimensional Analysis*, London: Bloomsbury.
- Egbert, J., & Biber, D. (2019). Incorporating text dispersion into keyword analyses. *Corpora*, 14(1), 77-104.
- Eisinger, R. M., Veenstra, L. R., & Koehn, J. P. (2007). What media bias? Conservative and liberal labeling in major US newspapers. *Harvard International Journal of Press/Politics*, 12(1), 17-36.
- Entman, R. M. (2010). Media framing biases and political power: Explaining slant in news of Campaign 2008. *Journalism*, 11(4), 389-408.
- Fairclough, N. , 2000. *New Labour, New Language*. London: Routledge
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1), 298-320.
- Fowler, R. (2013). *Language in the News: Discourse and Ideology in the Press*. Routledge.
- Fung, Y. R., Huang, K. H., Nakov, P., & Ji, H. (2022, August). The battlefield of combating misinformation and coping with media bias. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 4790-4791).
- Gangula, R. R. R., Duggenpudi, S. R., & Mamidi, R. (2019, August). Detecting political bias in news articles using headline attention. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* (pp. 77-84).

- Gans, H. J. (1979). *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. New York: Pantheon Books.
- Geis, M.L., 2012. *The language of politics*. Springer Science & Business Media.
- Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of political Economy*, 114(2), 280-316.
- Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35-71.
- Gentzkow, M., Shapiro, J. M., & Stone, D. F. (2015). Media bias in the marketplace: Theory. In *Handbook of media economics* (Vol. 1, pp. 623-645). North-Holland.
- Gladstone, B., (2015). Trust in the news media has come to mean affirmation. *The New York Times*, December issue.
- Gligorić, K., Lifchits, G., West, R., & Anderson, A. (2021). Linguistic effects on news headline success: Evidence from thousands of online field experiments (Registered Report Protocol). *Plos one*, 16(9), e0257091.
- Godbole, N., Srinivasaiah, M. and Skiena, S., 2007. Large-Scale Sentiment Analysis for News and Blogs. *Icwsn*, 7(21), pp.219-222.
- Goulart, L. (2022). *Communicative Text Types in University Writing* (Doctoral dissertation, Northern Arizona University).
- Graber, D.A., 1981. Political languages. *Handbook of political communication*, pp.195-224.
- Gray, B., & Biber, D. (2012). Current conceptions of stance. In *Stance and voice in written academic genres* (pp. 15-33). Palgrave Macmillan, London.

- Greene, S. and Resnik, P., 2009, May. More than words: Syntactic packaging and implicit sentiment. In: *Proceedings of human language technologies: The 2009 annual conference of the north american chapter of the association for computational linguistics* (pp. 503-511). Association for Computational Linguistics.
- Grefenstette, G., Qu, Y., Shanahan, J. G., & Evans, D. A. (2004). Coupling niche browsers and affect analysis for an opinion mining application. *Proceedings of Recherche d'Information Assistée par Ordinateur (RIAO)*.
- Gries, S. T. (2021). A new approach to (key) keywords analysis: Using frequency, and now also dispersion. *Research in Corpus Linguistics*, 9(2), 1-33.
- Gries, Stefan Th. 2008. Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics* 13/4: 403–437.
- Grieve, J., & Woodfield, H. (2023). *The Language of Fake News* (Elements in Forensic Linguistics). Cambridge: Cambridge University Press.
- Groeling, T. (2013). Media bias by the numbers: Challenges and opportunities in the empirical study of partisan news. *Annual Review of Political Science*, 16, 129-151.
- Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4), 1191-1237.
- Gruenewald, J., Pizarro, J., & Chermak, S. M. (2009). Race, gender, and the newsworthiness of homicide incidents. *Journal of criminal justice*, 37(3), 262-272.
- Guo, X., Ma, W., & Vosoughi, S. (2022, May). Measuring Media Bias via Masked Language Modeling. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 16, pp. 1404-1408).

- Hackett, R.A., 1984. Decline of a paradigm? Bias and objectivity in news media studies. *Critical Studies in Media Communication*, 1(3), pp.229-259.
- Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital journalism*, 6(3), 330-343.
- Hamborg, F. (2020, July). Media bias, the social sciences, and NLP: automating frame analyses to identify bias by word choice and labeling. In *Proceedings of the 58th annual meeting of the association for computational linguistics: student research workshop* (pp. 79-87).
- Hamborg, F., Meuschke, N., Aizawa, A., & Gipp, B. (2017). *Identification and analysis of media bias in news articles*. Humboldt-Universität zu Berlin.
- Hamborg, F., Zhukova, A., Donnay, K., & Gipp, B. (2020, August). Newsalyze: enabling news consumers to understand media bias. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020* (pp. 455-456).
- Hamborg, F., Zhukova, A., & Gipp, B. (2019). Automated identification of media bias by word choice and labeling in news articles. In *Proceedings of the 19th ACM/IEEE-CS joint conference on digital libraries, urbana-champaign, Illinois, USA*, <http://dx.doi.org/10.1109/JCDL.2019.00036>.
- Harcup, T., & O'Neill, D. (2001). What is news? Galtung and Ruge revisited. *Journalism studies*, 2(2), 261-280.
- Hofstetter, C. R. (1978). News bias in 1972: A cross-media comparison. *Journalism Monographs*, 58.

- Hofstetter, C. R., & Buss, T. F. (1978). Bias in television news coverage of political events: A methodological analysis. *Journal of Broadcasting & Electronic Media*, 22(4), 517-530.
- Horne, B., & Adali, S. (2017, May). This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In *Proceedings of the international AAAI conference on web and social media* (Vol. 11, No. 1, pp. 759-766).
- Holsanova, J., Rahm, H., & Holmqvist, K. (2006). Entry points and reading paths on newspaper spreads: comparing a semiotic analysis with eye-tracking measurements. *Visual Communication*, 5(1), 65-93.
- Horne, B., Khedr, S., & Adali, S. (2018). Sampling the news producers: A large news and feature data set for the study of the complex media landscape. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 12, No. 1).
- Hyland, K., & Guinda, C. S. (Eds.). (2012). *Stance and voice in written academic genres*. Basingstoke: Palgrave Macmillan.
- Iyyer, M., Enns, P., Boyd-Graber, J. and Resnik, P., 2014. Political ideology detection using recursive neural networks. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (Vol. 1, pp. 1113-1122).
- Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-6

- Kaya, M., Fidan, G., & Toroslu, I. H. (2012, December). Sentiment analysis of Turkish political news. In *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology* (Vol. 1, pp. 174-180). IEEE.
- Kahn, K. F., & Kenney, P. J. (2002). The slant of the news: How editorial endorsements influence campaign coverage and citizens' views of candidates. *American political science review*, *96*(2), 381-394.
- Kenney, K., & Simpson, C. (1993). Was coverage of the 1988 presidential race by Washington's two major dailies biased? *Journalism Quarterly*, *70*, 345-355.
- Kim, K.H., (2014). Examining US news media discourses about North Korea: A corpus-based critical discourse analysis. *Discourse & Society*, *25*(2), pp.221-244.
- Kiouis, S. (2001). Public trust or mistrust? Perceptions of media credibility in the information age. *Mass communication & society*, *4*(4), 381-403.
- Koivunen, A., Kanner, A., Janicki, M., Harju, A., Hokkanen, J., & Mäkelä, E. (2021). Emotive, evaluative, epistemic: a linguistic analysis of affectivity in news journalism. *Journalism*, *22*(5), 1190-1206.
- Krestel, R., Wall, A., & Nejd, W. (2012, April). Treehugger or petrolhead?: identifying bias by comparing online news articles with political speeches. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 547-548). ACM.
- Krieger, J. D., Spinde, T., Ruas, T., Kulshrestha, J., & Gipp, B. (2022). A domain-adaptive pre-training approach for language bias detection in news. *arXiv preprint arXiv:2205.10773*.

- Kulas, J. T., Stachowski, A. A., & Haynes, B. A. (2008). Middle response functioning in Likert-responses to personality items. *Journal of Business and Psychology*, 22, 251-259.
- Kweon, J. (2017). Media bias in us presidential elections: A case study of major cable news stations in 2016.
- Lazaridou, K., & Krestel, R. (2016). Identifying political bias in news articles. *Bulletin of the IEEE TCDD*, 12.
- Lazaridou, K., Krestel, R., & Naumann, F. (2017, November). Identifying media bias by analyzing reported speech. In *2017 IEEE International Conference on Data Mining (ICDM)* (pp. 943-948). IEEE.
- Lee, T. T. (2005). The liberal media myth revisited: An examination of factors influencing perceptions of media bias. *Journal of Broadcasting & Electronic Media*, 49(1), 43-64.
- Lim, S., Jatowt, A., Färber, M., & Yoshikawa, M. (2020). Annotating and analyzing biased sentences in news articles using crowdsourcing, LREC.
- Lim, S., Jatowt, A., & Yoshikawa, M. (2018). Understanding characteristics of biased sentences in news articles. In *CIKM workshops*.
- Lin, C. W. (2011). The Study of Political Language: A Brief Overview of Recent Research. *嘉南學報(人文類)*, (37), 471-485.
- Lin, W.H., Wilson, T., Wiebe, J. and Hauptmann, A., 2006, June. Which side are you on?: identifying perspectives at the document and sentence levels. In: *Proceedings of the tenth conference on computational natural language learning* (pp. 109-116). Association for Computational Linguistics.

- Lin, Y., Hoover, J., Portillo-Wightman, G., Park, C., Dehghani, M., & Ji, H. (2018, August). Acquiring background knowledge to improve moral value prediction. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 552-559). IEEE.
- Lundman, R. J. (2003). The newsworthiness and selection bias in news about murder: Comparative and relative effects of novelty and race and gender typifications on newspaper coverage of homicide. *Sociological Forum*, 18, 357–386.
- Lutz, A: These 6 corporations control 90% of the media in America. *Bus. Insider* (2014). <http://www.businessinsider.com/these-6-corporations-control-90-of-the-media-in-america-2012-6>. Accessed 12 Jan 2017
- MacKinnon, R. (2009). China’s censorship 2.0: How companies censor bloggers. *First Monday*, 14(2).
- McCarthy, J. D., McPhail, C., & Smith, J. (1996). Images of protest: Dimensions of selection bias in media coverage of Washington demonstrations, 1982 and 1991. *American sociological review*, 478-499.
- McCombs, M. E., & Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *The Public Opinion Quarterly*, 36(2), 176-87.
- McQuail, D. (1997). Accountability of media to society: Principles and means. *European journal of communication*, 12(4), 511-529.
- Mediabiasfactcheck.com (2015). Search and Learn the Bias of News Media. Retrieved July 10, 2019, from <https://mediabiasfactcheck.com/>

- Mitrović, J. (2013). Crowdsourcing and its application. *INFOtheca*, 14(1), 37–46.
- Millar, N. and B. Budgell. 2008. ‘The language of public health – a corpus based analysis’, *Journal of Public Health* 16 (5), pp. 369–74.
- Mladenović, M., Mitrović, J., & Krstev, C. (2016). A language-independent model for introducing a new semantic relation between adjectives and nouns in a WordNet. In *Proceedings of 8th global wordnet conference* (pp. 218–225)
- Morris, J.S., 2007. Slanted objectivity? Perceived media bias, cable news exposure, and political attitudes. *Social Science Quarterly*, 88(3), pp.707-728.
- Mullainathan, S., & Shleifer, A. (2002). Media bias.
- Mukherjee, S., & Weikum, G. (2015, October). Leveraging joint interactions for credibility analysis in news communities. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (pp. 353-362).
- Nakov, P., & Da San Martino, G. (2021, August). Fake news, disinformation, propaganda, media bias, and flattening the curve of the COVID-19 infodemic. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 4054-4055).
- Newton, K. (2006). May the weak force be with you: The power of the mass media in modern politics. *European Journal of Political Research*, 45(2), 209-234.
- Niven, D. (2002). *Tilt?: The search for media bias*. Greenwood Publishing Group.
- Niven, D. (2003). Objective evidence on media bias: Newspaper coverage of congressional party switchers. *Journalism & Mass Communication Quarterly*, 80(2), 311-326.
- Niven, D. (2004). A fair test of media bias: Party, race, and gender in coverage of the 1992 house banking scandal. *Polity*, 36(4), 637-649.

- Obeng, S. G. (1997). Language and politics: Indirectness in political discourse. *Discourse & Society*, 8(1), 49-83.
- Oelke, D., Geißelmann, B., & Keim, D. A. (2012). Visual Analysis of Explicit Opinion and News Bias in German Soccer Articles. In *EuroVA@ EuroVis*.
- Oliver, P. E., & Maney, G. M. (2000). Political processes and local newspaper coverage of protest events: From selection bias to triadic interactions. *American Journal of Sociology*, 106(2), 463-505.
- Oxford Languages. (n.d.). Bias. In *languages.oup.com*. Retrieved July 6, 2021, <https://languages.oup.com/google-dictionary-en/bias>
- Page, B. I., & Shapiro, R. Y. (2010). *The rational public: Fifty years of trends in Americans' policy preferences*. University of Chicago Press.
- Pang, B.; and Lee, L. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. In Proceedings of the ACL.
- Papacharissi, Z., & de Fatima Oliveira, M. (2008). News frames terrorism: A comparative analysis of frames employed in terrorism coverage in US and UK newspapers. *The international journal of press/politics*, 13(1), 52-74.
- Park, S., Ko, M., Kim, J., Liu, Y., & Song, J. (2011, March). The politics of comments: predicting political orientation of news stories with commenters' sentiment patterns. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 113-122).
- Patterson, T. E. (2017). News coverage of Donald Trump's first 100 days.

- Paul, R., & Elder, L. (2006). How to detect media bias and propaganda. *Dillon Beach, CA: Foundation for Critical Thinking.*
- Pereira, A., Harris, E., & Van Bavel, J. J. (2023). Identity concerns drive belief: The impact of partisan identity on the belief and dissemination of true and false news. *Group Processes & Intergroup Relations*, 26(1), 24-47.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in cognitive sciences*, 25(5), 388-402.
- Plonsky, L. (2015). *Advancing quantitative methods in second language research*. Routledge.
- Pojanapunya, P., & Todd, R. W. (2018). Log-likelihood and odds ratio: Keynes statistics for different purposes of keyword analysis. *Corpus Linguistics and Linguistic Theory*, 14(1), 133-167.
- Popat, K., Mukherjee, S., Yates, A., & Weikum, G. (2018). Declare: Debunking fake news and false claims using evidence-aware deep learning. *arXiv preprint arXiv:1809.06416*.
- Pothast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2017). A stylometric inquiry into hyperpartisan and fake news. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), 231–240.
- Puglisi, R., & Snyder Jr, J. M. (2015). Empirical studies of media bias. In *Handbook of media economics* (Vol. 1, pp. 647-667). North-Holland.
- Raeijmaekers, D. and Maesele, P., 2017. In objectivity we trust? Pluralism, consensus, and ideology in journalism studies. *Journalism*, 18(6), pp.647-663.

- Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2931–7.
- Ravenelle, A. J., Newell, A., & Kowalski, K. C. (2021). “The Looming, Crazy Stalker Coronavirus”: Fear Mongering, Fake News, and the Diffusion of Distrust. *Socius*, 7, 23780231211024776.
- Raymond, C., & Taylor, S. (2021). “Tell all the truth, but tell it slant”: Documenting media bias. *Journal of Economic Behavior & Organization*, 184, 670-691.
- Rayson, Paul. 2008a. From key words to key semantic domains. *International Journal of Corpus Linguistics* 13(4). 519–149.
- Rayson, Paul. 2008b. Log-likelihood and effect size calculator.
<http://ucrel.lancs.ac.uk/llwizard.html> (accessed 27 December 2022).
- Rayson, Paul. 2013. Corpus analysis of key words. In Carol A. Chapelle (ed.), *The encyclopaedia of applied linguistics*, 1–7. Oxford: Wiley-Blackwell.
- Recasens, M., Danescu-Niculescu-Mizil, C. and Jurafsky, D., 2013. Linguistic models for analyzing and detecting biased language. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (Vol. 1, pp. 1650-1659).
- Rich, M. D. (2018). *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*. Rand Corporation.
- Richardson, J. (Ed.). (2013). *Language and journalism*. Routledge.

Rubin, V. L. (2019). Disinformation and misinformation triangle: A conceptual model for ‘fake news’ epidemic, causal factors and interventions. *Journal of Documentation*, 75(5), 1013–34.

Rubin, V. L., Conroy, N., Chen, Y., & Cornwell, S. (2016, June). Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection* (pp. 7-17).

Rubin, V. L., Chen, Y., & Conroy, N. K. (2015). Deception detection for news: Three types of fakes. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–4.

Saez-Trumper, D., Castillo, C., & Lalmas, M. (2013, October). Social media news communities: gatekeeping, coverage, and statement bias. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management* (pp. 1679-1684). ACM.

Sales, A., Zehe, A., Marinho, L. B., Veloso, A., Hotho, A., & Omeliyanenko, J. (2021, May). Assessing Media Bias in Cross-Linguistic and Cross-National Populations. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 15, pp. 561-572).

SANTOS-D’AMORIM, K., & de Oliveira Miranda, M. K. F. (2021). Misinformation, disinformation, and malinformation: clarifying the definitions and examples in disinfodemic times. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 26.

Schaefer, T., & Fordan, R. (2014). Balance and bias in network evening news coverage of presidential state of the union addresses. *Electronic News*, 8(4), 274-289.

- Schiffer, A. J. (2021). Teaching Media Bias: The Case of the Trump Presidency. *Journal of Political Science Education*, 17(sup1), 955-972.
- Schudson, M., 1981. *Discovering the news: A social history of American newspapers*. Basic Books.
- Sheppard, S. (2007). *The partisan press: A history of media bias in the United States*. McFarland.
- Shoemaker, P. J., & Reese, S. D. (1991). *Mediating the message: Theories of influences on mass media content*. New York: Longman.
- Shultziner, D., & Stukalin, Y. (2021). Politicizing what's news: How partisan media bias occurs in news production. *Mass Communication and Society*, 24(3), 372-393.
- Simons, G., & Strovsky, D. (2006). Censorship in contemporary Russian journalism in the age of the war against terrorism: A historical perspective. *European Journal of Communication*, 21(2), 189-211.
- Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From protest to agenda building: Description bias in media coverage of protest events in Washington, DC. *Social Forces*, 79(4), 1397-1423.
- Snider, M., (2018). Americans prefer to watch news, on TV and online, rather than read it, study finds. *USA Today*, Retrieved from:
<https://www.usatoday.com/story/money/media/2018/12/03/most-americans-prefer-watch-news-not-read-study-says/2189310002/>

Society of Professional Journalists (US). (2014). Society of Professional Journalists: Code of Ethics. Society of Professional Journalists.

Spinde, T. (2021, December). An interdisciplinary approach for the automated detection and visualization of media bias in news articles. In *2021 International Conference on Data Mining Workshops (ICDMW)* (pp. 1096-1103). IEEE.

Spinde, T., Hamborg, F., Donnay, K., Becerra, A., & Gipp, B. (2020, August). Enabling news consumers to view and understand biased news coverage: a study on the perception and visualization of media bias. In *Proceedings of the ACM/IEEE joint conference on digital libraries in 2020* (pp. 389-392).

Spinde, T., Kreuter, C., Gaissmaier, W., Hamborg, F., Gipp, B., & Giese, H. (2021, September). Do you think it's biased? how to ask for the perception of media bias. In *2021 ACM/IEEE Joint Conference on Digital Libraries (JC DL)* (pp. 61-69). IEEE.

Spinde, T., Rudnitckaia, L., Mitrović, J., Hamborg, F., Granitzer, M., Gipp, B., & Donnay, K. (2021). Automated identification of bias inducing words in news articles using linguistic and context-oriented features. *Information Processing & Management*, 58(3), 102505.

Stahl, B. C. (2006). On the difference or equality of information, misinformation, and disinformation: A critical research perspective. *Informing Science: The International Journal of an Emerging Transdiscipline*, 9, 83–96.

Stempel III, G. H. (1969). The prestige press meets the third-party challenge. *Journalism Quarterly*, 46(4), 699-706.

- Strapparava, C., & Mihalcea, R. (2007, June). Semeval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)* (pp. 70-74).
- Sunstein, C. R. (2001). *Designing democracy: What constitutions do*. Oxford University Press.
- Swift, A., 2017. Six in 10 in US See Partisan Bias in News Media. *Gallup News*.
- Syverson, T., 2017. Semiotics in the age of trump: Linguistics can help conservatives and liberals agree on objective reality again. *Quartz*. <https://qz.com/916319/the-linguistics-of-political-language-can-help-conservatives-and-liberals-agree-on-objective-reality-again/>
- Taboada, M. (2021). Authentic language use in fake news. In *Beyond Disinformation from the Social Science and Research Council*. <https://items.ssrc.org/beyond-disinformation/authentic-language-in-fake-news/>
- Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining ‘fake news’ A typology of scholarly definitions. *Digital Journalism*, 6(2), 137–53.
- Tangri, K. (2021). Using natural language to predict bias and factuality in media with a study on rationalization (*Doctoral dissertation, Massachusetts Institute of Technology*).
- Torabi Asr, F. (2019). The language gives it away: How an algorithm can help us detect fake news. *The Conversation*.
- Uscinski, J. E., Klofstad, C., & Atkinson, M. D. (2016). What drives conspiratorial beliefs? The role of informational cues and predispositions. *Political Research Quarterly*, 69(1), 57-71.

- Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of personality and social psychology*, 49(3), 577.
- Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in cognitive sciences*, 22(3), 213-224.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.
- Vraga, E. K., Tully, M., & Rojas, H. (2009). Media literacy training reduces perception of bias. *Newspaper Research Journal*, 30(4), 68-81.
- Wenzel, A. (2019). To verify or to disengage: Coping with 'fake news' and ambiguity. *International Journal of Communication*, 13, 1977–95.
- White, D. M. (1950). The gate-keeper: A case study in the selection of news. *Journalism Quarterly*, 27, 383–390.
- Williams, A. (1975). Unbiased study of television news bias. *Journal of Communication*, 25(4), 190–199.
- Wilson, J. (1990). *Politically Speaking: The Pragmatic Analysis of Political Language*, ser. *Language in society* (Oxford, England). Oxford: Basil Blackwell. Wilson, J., 2003. 20 Political Discourse. *The handbook of discourse analysis*, 18, p.398.
- Wolton, S. (2019). Are biased media bad for democracy?. *American Journal of Political Science*, 63(3), 548-562.

- Xu, M., Luo, Z., Xu, H., & Wang, B. (2022). Media bias and factors affecting the impartiality of news agencies during COVID-19. *Behavioral Sciences, 12*(9), 313.
- Yano, T., Resnik, P. and Smith, N.A., 2010, June. Shedding (a thousand points of) light on biased language. In: *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 152-158). Association for Computational Linguistics.
- Yin, J. (2008). Beyond the four theories of the press: A new model for the Asian & the world press. *Journalism & Communication Monographs, 10*(1), 3-62.
- Zhou, Z. (2022, February). Discourse Analysis: Media Bias and Linguistic Patterns on News Reports. In *2021 International Conference on Education, Language and Art (ICELA 2021)* (pp. 271-277). Atlantis Press.
- Zelizer, B., Park, D., & Gudelunas, D. (2002). How bias shapes the news: Challenging The New York Times' status as a newspaper of record on the Middle East. *Journalism, 3*(3), 283-307.

Perceptual Ratings of News 1

Start of Block: Directions

Q1 For the survey, you will be asked to read one newspaper article and provide your opinions about the bias of the article. You will be asked to provide your opinions about the extent of bias (0=no bias to 5=extreme bias) as well as the direction of bias (from strongly liberal to strongly conservative). Please be as honest as possible and avoid Googling any information about the articles. It is important for the research that you provide your initial perceptions.

Answers will remain anonymous and will be used strictly for research purposes. By clicking 'agree' below, you are giving your consent to have your responses included in the research. In order to qualify for payment, you must agree and complete the entire survey.

Payment is contingent upon honest, legitimate completion of the survey. Thus, if answers are contradictory, results will not be accepted, and payment will not be sent.



Q2 Do you consent to have your answers used in the research (responses will remain anonymous).

- Agree (1)
- Do not agree (2)

End of Block: Directions

Start of Block: Background Questions

Q3 Please provide your user (worker) ID:

Q4 Please select your age range.

- 18-29 (1)
- 30-49 (2)
- 50-64 (3)
- 65+ (4)

Appendix A

Q5 Please select your sex.

- Male (1)
- Female (2)
- Other (3) _____
- Prefer not to state (4)

Q6 Please select your highest degree achieved.

- High school diploma/ GED (1)
- Some College (2)
- Associate's degree (3)
- Bachelor's degree (4)
- Master's degree (5)
- PhD or other terminal degree (6)

Q7 Please select your political leaning.

- Strongly liberal (1)
- Liberal (2)
- Moderate (3)
- Conservative (4)
- Strongly conservative (5)

End of Block: Background Questions

Start of Block: The articles

Q8 Please click the link to access the article. It will download as a word document. Please read the text in it entirety before answering the following questions.

Appendix A

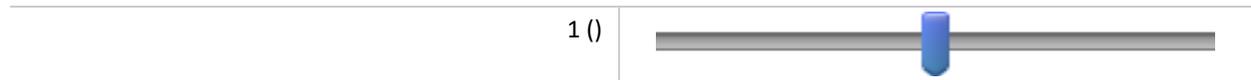
Q9 Please write the full number that appears in the link above. Example 1 (#).

End of Block: The articles

Start of Block: Questions about the reading

Q10 Do you agree with the perspective presented in the article?

Definitely yes	Mostly yes	Might or might not	Mostly not	Definitely not
1	2	3	4	5



Q11 Please select the extent to which you believe the article is biased.

Not at all biased	Slightly biased	Somewhat biased	Pretty biased	Extremely biased	
0	1	2	3	4	5



Q12 Please select the direction you believe this article leans.

Strongly Left/Liberal	Left/Liberal	Neutral (no obvious leaning)	Right/ Conservative	Strongly Right/ Conservative
1	2	3	4	5



Appendix A

Q13 Please provide any additional information you feel is important related to the bias of the article (optional).

End of Block: Questions about the reading

Start of Block: Random ID

Q14 Here is your ID: `{e://Field/Random%20ID}`

Copy and paste into MTurk

When you have copied this ID, please hit the next button to submit your survey

End of Block: Random I

Appendix B

Table B1

Articles Rated as Extremely Biased

Publication	Year	Topic	Article Title	Category
AlterNet	2016	Anthony Weiner sexts	“John Oliver Destroys Anthony Weiner Who Has Dangerously United Americans in Hatred of Him”	Media (Entertainment)
AlterNet	2016	Global Warming hoax	Donald Trump’s Global Warming Stance Denounced by China	Environment
AlterNet	2016	ISIS	ISIS Must Love Trump	Politics
AlterNet	2016	Orlando shooting	How the Corporate Media’s Spin on the Orlando Shooting Is Boosting America’s Dirty Wars in Muslim-Majority Countries	World
AlterNet	2016	Penn Chapo interview	When Hollywood Calls, Who Can Resist?	Media (Entertainment)
AlterNet	2016	2016 Presidential elections	10 Shocking 2016 Election Facts: Old Political Assumptions Are out the Window	Politics
AlterNet	2016	Trump elected	Trump Elected President by Wave of Angry White Voters Across Upper Midwest and South	Politics & elections
AlterNet	2016	Trump sexual assault	In Nazi-Like Speech, Trump Responds to Sexual Assault Claims with Broad Conspiracy Theory Designed to Foment Mayhem	Politics
AlterNet	2016	Trump’s taxes	Did Marla Maples Leak Donald Trump’s Tax Return?	Economy
AlterNet	2017	Las Vegas shooting	So Few Americans Understand What the Second Amendment Is Really About – or Its Dark History	Rights (U.S.)
AlterNet	2017	McCain’s healthcare bill	Let’s Not Ignore the Real Heroes Who Saved Our Health Care System from Total Ruin	Health
AlterNet	2017	#MeToo	6 Big Victories for Feminism in 2017, Despite Donald Trump and Other Toxic Men	Rights (Life)
AlterNet	2017	Solar eclipse	Trump Looked Directly at the Solar Eclipse Without Eye Protection, Because of Course He Did	Environment
AlterNet	2017	Trump’s travel ban	Donald Trump Rolls Out New Muslim Travel Ban	Economy
AlterNet	2017	Trump’s Supreme Court nominees	Trump Considered Rescinding His Neil Gorsuch Nomination Because He Wasn’t Sufficiently ‘Loyal’	Politics
AlterNet	2018	Kavanaugh Supreme Court	‘Out of Control’: Brett Kavanaugh Stuns Observers with Furious Opening Statement Revealing Himself as a Right-Wing Operative	The Right Wing (Politics)

Appendix B

AlterNet	2018	Nike hires Kaepernick	A Right-Wing Musician Cut Up His Nike Socks To Protest Kaepernick – And Was Immediately Mocked by the Internet	Media (Entertainment)
AlterNet	2018	Roseann Barr Ambien tweet	ABC's Roseanne Barr 'Now Leaving Twitter' After Apologizing for Racist Attack on Obama Adviser	Media (Entertainment)
AlterNet	2018	Trump's tariffs	Trump's Steel and Aluminum Tariffs Are Absolutely Horrible for American Workers	Economy
The Wall Street Journal	2016	Anthony Weiner sexts	FBI Reviewing Newly Discovered Emails in Clinton Server Probe; Emails surfaced during agency's investigation into former Rep. Anthony Weiner's alleged sexting	Politics
The Wall Street Journal	2016	Black Lives Matter	The Radical Democratic Party; Plus, the myths of Obama and Black Lives Matter	Opinion
The Wall Street Journal	2016	Brexit	Brexit: A Seismic Slap in the Face of Markets; Markets got Brexit dead wrong. Now comes the fallout	Economy
The Wall Street Journal	2016	China as a superpower	Trump Plans to Beat China at Its Own Game; Donald Trump's policies are strategically designed to beat the Chinese at their own game.	U.S.
The Wall Street Journal	2016	Clinton Emails	Oil Prices Sag on Concerns About OPEC Deal on Output Cut; New chapter in FBI probe of Hillary Clinton email server unsettles market	Economy
The Wall Street Journal	2016	Global Warming hoax	Donald Trump's Victory Injects Uncertainty into Climate Accord; President-elect has called climate change a hoax and said he would cancel Paris pact	Politics
The Wall Street Journal	2017	Solar eclipse	Marketers Cash in on Eclipse; Brands are piggybacking off the solar eclipse with marketing pitches ranging from funny to absurd to practical	Business/Economy
The Wall Street Journal	2017	Trump's Supreme Court nominees	Another Escalation in the Judicial War What will happen when a Senate majority votes in unison against every presidential nominee?	Opinion
The Wall Street Journal	2018	Net Neutrality	As Net Neutrality Ends, Activists Push Congress to Restore Obama-Era Rules	Politics
The Wall Street Journal	2018	Winter Olympics	Winter Olympics Prove Financial Headache for Some, Party for Others	World
Fox	2016	Melania's plagiarism	Melania Trump speechwriter admits 'mistake' in lifting Michelle Obama lines	Republican Convention (Politics)
Fox	2016	Orlando shooting	Hillary Clinton's response to Orlando proves she's too weak to be president	Opinion
Fox	2017	Betsy DeVos	Four reasons why Betsy DeVos will be a welcome remedy for our nation's education woes	Opinion
Fox	2017	#MeToo	Universities face #MeToo movement over sexual harassment	U.S.
Fox	2017	Saturday Night Live	EXCLUSIVE: Donald Trump blasts 'disgusting' Madonna and 'disgraced' ex-'SNL' writer on 'Hannity'	Entertainment

Appendix B

Fox	2017	Take a Knee	NFL protests: Stars should get off their knees and lead a constructive conversation on race	Opinion
Fox	2018	2018 Midterm elections	The 2018 midterm elections aren't just about candidates, they're also about respect, sovereignty, rule of law	Politics
Fox	2018	Mira Ricardel fired	Mira Ricardel to leave White House, after Melania Trump's office called for her dismissal	Politics
Fox	2018	Elon Musk social media	Elon Musk yanks SpaceX and Tesla's Facebook pages	Media (Entertainment)
Fox	2018	Net neutrality	Net neutrality nixed by FCC: What to know about the Internet regulations	Tech
Fox	2018	Taylor Swift politics	Stars like Taylor Swift don't always make a difference when pushing their politics: expert	Media (Entertainment)
The Washington Times	2016	Anthony Weiner sexting	Anthony Weiner's revenge on Clintons	Opinion
The Washington Times	2016	Black Lives Matter	Dan Patrick presses Obama to show more support for police	Politics
The Washington Times	2016	Gun Control	Hillary Clinton's gun platform would take away popular firearms	Politics
The Washington Times	2016	Harriet Tubman on \$20	Harriet Tubman to replace Andrew Jackson on \$20 bill after pushback against Alexander Hamilton plan	Politics
The Washington Times	2016	Penn Chapo interview	Sean Penn's 'El Chapo' interview stirs rumors of government role, reward for capture	U.S.
The Washington Times	2017	Betsy DeVos	Betsy DeVos wins confirmation as education secretary as VP Pence casts historic vote	Politics
The Washington Times	2017	Royal Wedding	Meghan Markle joins royal Christmas lunch at palace	People (Life)
The Washington Times	2017	Saturday Night Live	Alec Baldwin, who parodies Trump on SNL: All jokes aside, 'This country is lost'	Politics
The Washington Times	2017	Take a Knee	Once upon a knee at the old ball game	Opinion
The Washington Times	2017	Trump's national security	How Trump strengthens national security The executive order turns up the heat on known sponsors of terror	U.S.
The Washington Times	2018	Bush death	Bush's spirit of compassion, duty soared beyond confines of politics	U.S.
The Washington Times	2018	Child immigrants	Flashback: Obama Administration placed border children with human smugglers	National (U.S.)

Appendix B

The Washington Times	2018	2018 Midterm elections	Women win record number of congressional seats in 2018 midterm elections	Politics
The Washington Times	2018	Mira Ricardel Fired	'The new Melania': First lady stuns with public power play	Politics
The Washington Times	2018	Net neutrality	Your internet use could change as 'net neutrality' ends	Technology
The Washington Times	2018	Russia	Trump, conservatives demand probe into Clinton-Steele-Moscow links	World
The Washington Times	2018	Trump and North Korea	Trump's 'down the road' deadline to lift North Korea sanctions frustrates nations eager for trade	Politics
Breitbart	2016	Black Lives Matter	A Short List of Black Lives Matter's Cop-Killing Heroes	Politics
Breitbart	2016	Clinton's Emails	Hiding Hillary Day 273: Fallout from the FBI Release	Politics
Breitbart	2016	Gay marriage	More Than One Million March on Rome for 'Family Day' to Protest Gay Unions Bill	World
Breitbart	2016	Global Warming hoax	Report: With Global Warming, Expect 'Colder Winters'	Environment
Breitbart	2016	Gun control	Fail: Gun-Controlled Australia Admits Criminals Still Armed, Launches New Amnesty	World
Breitbart	2016	ISIS	Hillary Clinton's Policies Created Conditions for Rise of Islamic State	Politics
Breitbart	2016	Melania's plagiarism	Michelle Obama Copied Alinsky in Speech Melania Trump Allegedly Plagiarized	Politics
Breitbart	2016	Orlando shooting	The Left Chose Islam Over Gays. Now 100 People Are Dead Or Maimed In Orlando	Social Justice (U.S.)
Breitbart	2016	Penn Chapo interview	Sean Penn Interviews El Chapo for Rolling Stone... SHOCK: Magazine Gives Fugitive Drug Lord Editorial Control	Media (Entertainment)
Breitbart	2016	Rio Olympics	Five Crazy Crimes That Have Occurred at the Rio Olympics — Already	Sports
Breitbart	2016	Saudi Arabia 9/11 bill	Paul Sperry: Release the 28 Pages So We Can 'End This Suicidal Phony Alliance with Saudi Arabia'	Radio (Entertainment)
Breitbart	2016	Scalia's death	Antonin Scalia's Death Could Mark End of Constitution	U.S.
Breitbart	2016	Trump elected	5 Most Absurd Ways the Left Has Responded to the 2016 Election	Politics

Appendix B

Breitbart	2016	Trump sexual assault	Joe Biden, Groper-in-Chief, Accuses Donald Trump of ‘Sexual Assault’	Politics
Breitbart	2016	Trump’s taxes	Trump Campaign: NYT Published Tax Docs ‘Illegally Obtained’ After Editor Promised He Would Go to Jail over Trump Taxes	Politics
Breitbart	2016	Zika virus	Massive Study Casts Doubt on Zika As Cause of Microcephaly	Health
Breitbart	2017	Covfefe tweet	A Meme Is Born: Donald Trump ‘Covfefe’ Typo Breaks the Internet	Tech
Breitbart	2017	Kathy Griffin fired	Donald Trump on ‘Sick’ Beheading Photo: ‘Kathy Griffin Should Be Ashamed of Herself’	Entertainment
Breitbart	2017	Kennedy Center Honors	Donald Trump, First Lady Melania to Skip Kennedy Center Honors	U.S.
Breitbart	2017	Las Vegas shooting	‘Do Something’ After Las Vegas Shooting. OK — What?	Politics
Breitbart	2017	#MeToo	‘This Has Gone Too Far’: Tavis Smiley Challenges #MeToo Overreach	U.S.
Breitbart	2017	Take a Knee	Why Taking a Knee During the National Anthem Is Not Only Unpatriotic, It’s Ignorant	U.S.
Breitbart	2017	Trump’s national security	Michael Flynn Resigns as National Security Advisor	Politics
Breitbart	2017	United drags man off plane	United CEO: Man dragged off plane was ‘disruptive,’ ‘belligerent’	News (U.S.)
Breitbart	2017	Women’s March	Women’s March — See What a Massive, Hillary-Shaped Bullet America Just Dodged?	Politics
Breitbart	2018	Bush death	Left Attacks George H.W. Bush After His Death: ‘Blatant Racism and Toxic Masculinity,’ ‘F**k Him’	U.S.
Breitbart	2018	Child immigrants	13 Facts the Media ‘Pros’ Don’t Want You to Know About ‘Family Border Separation’	U.S.
Breitbart	2018	Kavanaugh Supreme Court	Every Witness Named by Kavanaugh’s Accusers Sides with Kavanaugh	Politics
Breitbart	2018	March for our Lives	March for Our Lives – This is What Demagoguery Looks Like	U.S.
Breitbart	2018	Net neutrality	Exclusive — Sen. Kennedy on Net Neutrality: Phone Companies Don’t Tell You What You Can Say, Neither Should Cable Companies	U.S.
Breitbart	2018	Nike hires Kaepernick	Nike Shares Drop as Kaepernick Ad Triggers Calls for Boycott	Economy

Appendix B

Breitbart	2018	Trump's tariffs	Promises Kept: Donald Trump Signs Steel and Aluminum Tariffs to Protect Workers	Economy
Newsmax	2016	Black Lives Matter	Trump Promises to Protect	Politics
Newsmax	2016	Brexit	Brexit Shows Us the Future and It's Great	Economy
Newsmax	2016	Prince death	The Donald Trump-Prince Connection	Entertainment
Newsmax	2016	Saudi Arabia 9/11 bill	Saudi's Pull Mafia Moves Over 9/11 Role	Politics
Newsmax	2016	Trump sexual assault	Trump: Why Don't Women Come Out Against Obama?	Politics
Newsmax	2017	Comey's testimony	Comey's Wounds Self-Inflicted, While Trump Unscathed	Politics
Newsmax	2017	Las Vegas shooting	Las Vegas Shooting Immediately Politicized	Newsfront (U.S.)
Newsmax	2017	McCain healthcare bill	Saying Republicans Can't Govern Is a Cheap Shot	Health
Newsmax	2017	Millennials	Millennials Would Give Up Sex, Coffee for Hobbies, Study Says	Newsfront (Life)
Newsmax	2017	Mueller firing	Dershowitz: 'Foolish' if Trump Fires Mueller; 'Really Foolish' to Seize Emails	Newsfront (U.S.)
Newsmax	2017	Saturday Night Live	Alec Baldwin, 'SNL's Donald Trump, to Host Show for Record 17th Time	TheWire (Entertainment)
Newsmax	2017	Take a Knee	National Anthem Knee-Takers Defacing the Ultimate Monument	America (U.S.)
Newsmax	2017	Trump's travel ban	Trump Travel Ban Faces Key Test in Multiple Courtrooms	Newsfront (U.S.)
Newsmax	2017	Trump's Supreme Court nominees	A Lasting Trump Future for the Federal Bench	Politics
Newsmax	2017	United drags man off plane	Dem Demands Hearing On United Passenger Dragged Off Plane	Newsfront (U.S.)
Newsmax	2017	Women's March	Protesters Litter Trump International, Trump Tower with Signs	Politics
Newsmax	2017	Wonder Woman movie	Movie Theater Doubles Down on Women-Only 'Wonder Woman' Showings	Entertainment

Appendix B

Newsmax	2018	Child immigrants	Obama Separated Immigrants Too, But Democrats Now See Midterm Issue	Opinion
Newsmax	2018	March for our Lives	At 'March for Our Lives,' the Young Lead the Foolish	Opinion
Newsmax	2018	Roseanne Barr Ambien tweet	Roseanne Barr, Samantha Bee and the Need for Consistency	Entertainment
Newsmax	2018	Stormy Daniels	Judge Rejects Trial Demand in Stormy Daniels Case	TheWire (U.S.)
Newsmax	2018	Trump's tariffs	World Threatens Retaliation on Trump Tariffs	Economy
The National Review	2016	Clinton's emails	'We Need to Clean This Up': More Evidence Obama Lied about Hillary's Private E-mails	Politics & policy
The National Review	2016	Gun control	Does Hillary Believe in an Individual Right to Bear Arms or Not?	Politics & policy
The National Review	2016	Orlando shooting	The Orlando Shooting Launches a War on Christianity	Politics & policy
The National Review	2016	Presidential election	The Blue Wave - Republicans must adapt to a diversifying electorate or lose their influence	Feature (Politics)
The National Review	2016	Trump sexual assault	Trump Is Validating the Left's Crudest Stereotypes of the Right	Culture (U.S.)
The National Review	2016	Trump's taxes	White House: No Plans to Release Trump Tax Returns After NYT Story	White House (Politics)
The National Review	2017	Betsy DeVos	What Is the Democratic Party?	Politics & Policy
The National Review	2017	Kathy Griffin fired	Kathy Griffin and Hillary Clinton: Wonder Women?	Culture (Life)
The National Review	2017	Las Vegas shooting	The Passionate Non-Sequiturs of the Gun Debate	Politics & policy
The National Review	2017	#MeToo	The Sexes after Weinstein	Culture (Life)
The National Review	2017	Take a Knee	I Understand Why They Knelt	Culture (Life)
The National Review	2017	Trump's travel ban	Trump's Revised Travel Ban is Just One of Many Good Developments in the War against Jihadists	Politics & policy
The National Review	2017	Trump's Supreme Court nominees	But Gorsuch ... and Other Excellent Judicial Picks ... and a Tax Cut ... and Major Deregulatory Actions ... and	Politics & policy

Appendix B

Immigration Enforcement ... and the End of the Individual Mandate ... and a Roll Back of the HHS Mandate ...

The National Review	2018	Bush death	Partisans Use 41 to Attack 45	White House (Politics)
The National Review	2018	Child immigrants	Bordering on Insanity	U.S.
The National Review	2018	Fire & Fury book release	Is Fire and Fury Fizzling?	Politics & policy
The National Review	2018	Mira Ricardel fired	Trump's Constantly Backfiring Loyalty Tests	White House (Politics)
The National Review	2018	Nike hires Kaepernick	The Party of Kaepernick	Elections (Politics)
The National Review	2018	Roseanne Barr Ambien tweet	Roseanne and the High Cost of Embracing Crazy	Entertainment
The National Review	2018	Taylor Swift politics	Taylor Swift Buys the Democrats' Marketing; Tennesseans Shouldn't	Culture (Life)
The National Review	2018	Trump's tariffs	Man of Steel	Economy & Business
The Nation	2016	Gun control	America's Refusal to Control Firearms Is Killing Teens at an Absurd Rate	U.S.
The Nation	2016	Orlando shooting	Let's Make Gun Control a Gay Issue It's time to glitter bomb the NRA. Who's in?	U.S.
The Nation	2016	2016 Presidential elections	An Unabashed Misogynist Is in Charge of Our Country. Now What?	Feature (Politics)
The Nation	2016	Refugee crisis	The US Has Its Own Refugee Crisis—One We Created, and Now Must Solve	Economy
The Nation	2016	Rio Olympics	Teachers and Students Occupy Schools in the Shadow of Olympic Rio	Events (Sports)
The Nation	2016	Scalia Death	America Changed. Antonin Scalia Never Did.	U.S.
The Nation	2017	Covfefe tweet	The Covfefe Is Often Worse Than the Crime	Politics
The Nation	2017	McCain healthcare bill	John McCain Isn't a Hero for Not Killing His Constituents	Health
The Nation	2017	#MeToo	From Bill O'Reilly to Al Franken: Katha Pollitt on #MeToo in 2017	Politics

Appendix B

The Nation	2017	Wonder Woman movie	The real heroines in this summer’s movies don’t need a CGI treatment.	Entertainment
The Nation	2018	Bush death	George H.W. Bush, Icon of the WASP Establishment—and of Brutal US Repression in the Third World Obituaries have transformed the terror that Bush inflicted, depicting it as heroism.	U.S.
The Nation	2018	Stormy Daniel’s	Stormy Daniels Going Public Is Pussy Grabbing Back Could Donald Trump become a victim of his own licentiousness?	U.S.
The Nation	2018	Trump and North Korea	Despite Anonymous Carping, US–North Korea Talks Continue	World
HuffPost	2016	Trump elected	Economic Insecurity, Racism And Sexism All Help Explain Trump’s Victory	Politics
HuffPost	2016	Trump sexual assault	Trump’s Morning: Attack Biden, Female Accusers, Fellow Republicans, Electoral Integrity All before noon.	Politics
HuffPost	2017	Amazon buys Whole Foods	Amazon Is Buying Whole Foods in A Quest to Beat Walmart with Luxury	Business (Economy)
HuffPost	2017	Kennedy Center honors	Donald Trump Will Skip Kennedy Center Honors to Avoid ‘Political Distraction’	Politics
HuffPost	2017	Mueller firing	Sen. Mark Warner Warns Trump: Firing Robert Mueller Would Be A ‘Gross Abuse of Power’	Politics
HuffPost	2017	Pepsi Commercial	Pepsi’s Portrayal of a Muslim Woman Was As Tone-Deaf As The Rest Of The Ad A headscarf is not a prop.	Women (Life)
HuffPost	2017	Take a Knee	Actor Michael Rapaport Takes A Knee, Unloads On ‘Dumb Motherf—ker’ Donald Trump	Entertainment
HuffPost	2017	Trump’s Supreme Court nominees	Trump’s Biggest Win In 2017? His Judicial Confirmations.	Politics
HuffPost	2017	Wonder Woman movie	Fox News Mulls Over Whether Wonder Woman Is ‘American’ Enough	Entertainment
HuffPost	2018	Fire and Fury book release	The Wildest Moments From ‘Fire And Fury,’ The Trump Book Everyone Is Talking About	Politics
HuffPost	2018	Net Neutrality	Net Neutrality Can Still Be Saved	Opinion
HuffPost	2018	Trump “very stable genius”	Mr. Trump and Mr. Ed: A very stable genius and a genius in the stable	Opinion
CNN	2016	Prince death	Obama on the death of Prince: 'It's a remarkable loss'	U.S.

Appendix B

CNN	2016	Refugee crisis	Refugee crisis: Why Scandinavian countries are trying to look bad	Opinion
CNN	2017	Russia probe	The Russia Investigation: Ex-Trump campaign adviser pleads guilty to making false statement	Politics
CNN	2018	Kavanaugh Supreme Court	How Kavanaugh maneuvered to win his confirmation fight	Politics
CNN	2018	March for our Lives	Student marchers call Washington's inaction on gun violence unacceptable	U.S.
The Atlantic	2016	Clinton emails	There's Simply No Comparison Between Clinton's Flaws and Trump's	Politics
The Atlantic	2016	Global Warming hoax	What President Trump Will Mean for Earth's Climate He could prove cataclysmic.	Science
The Atlantic	2016	ISIS	How ISIS Weaponized the Media After Orlando	Global (World)
The Atlantic	2016	Melania's plagiarism	Ten Reasons Why Melania Trump's Speech Will Have a Lasting Impact	Politics
The Atlantic	2017	Comey's testimony	James Comey's 'Shock and Awe' Testimony	Global (World)
The Atlantic	2017	Trump's national security	Trump's News-Conference Performance	News (U.S.)
The Atlantic	2018	Trump as a "stable genius"	How Actual Smart People Talk About Themselves Hint: not by discussing IQ	Science
The New York Times	2016	Gun control	What 130 of the Worst Shootings Say About Guns in America	U.S.
The New York Times	2016	Trump elected	New Yorkers Baffled and Anxious After Trump, One of Their Own, Clinches Election	New York (U.S.)
The New York Times	2016	Trump sexual assault	For Many Women, Trump's 'Locker Room Talk' Brings Memories of Abuse	Politics
The New York Times	2017	#MeToo	The Patriarchs Are Falling. The Patriarchy Is Stronger Than Ever.	Opinion
The New York Times	2017	Saturday Night Live	Donald Trump News Conference Gets the 'S.N.L.' Treatment	Arts (entertainment)
The New York Times	2017	Trump's national security	From Trump's Mar-a-Lago to Facebook, a National Security Crisis in the Open	Politics
The New York Times	2018	2018 Midterm elections	Weak Spots in Democrats' Strong Midterm Results Point to Challenges in 2020	The Upshot (Politics)

Appendix B

The New York Times	2018	Roseanne Barr Ambien tweets	Roseanne Barr's Tweets Didn't Come Out of Nowhere	Arts (Entertainment)
The New York Times	2018	Russia probe	Mueller Team Has 'Gone Absolutely Nuts,' Trump Says, Resuming Attacks on Russia Inquiry	Politics
The New York Times	2018	Serena Williams & sexism in tennis	Serena Williams Accuses Official of Sexism in U.S. Open Loss to Naomi Osaka	Sports
USA Today	2016	Black Lives Matter	Criticism of Black Lives Matter deceitful, activist says - Mckesson: Cops chase people 'for kicks'	News (U.S.)
USA Today	2016	Trump elected	Trump's surprise victory leaves a nation unnerved	Opinion
USA Today	2016	Trump sexual assault	Women say they were groped by Trump	Investigation (U.S.)
USA Today	2017	Pepsi commercial	How did Pepsi's ad even get off the drawing board?	Money (Economy)
USA Today	2017	Russia sanctions	Poll: Nearly half of Americans think Trump committed a crime related to Russian election meddling	Politics
USA Today	2018	Fire & Fury book release	'Fire and Fury' is No. 1 on USA TODAY's list as Trump book sets sales records	Books (Entertainment)
USA Today	2018	Aretha Franklin	Soul legend Aretha Franklin had pancreatic cancer, one of the most deadly forms	Nation (Entertainment)
NPR	2016	Anthony Weiner sexting	FBI Led Back to Clinton Email Server Case By Anthony Weiner Investigation	Politics
NPR	2016	Global Warming hoax	Has Trump's Stumping Already Frayed Transatlantic Ties?	Environment
NPR	2016	ISIS	Iraqi Forces Claim Victory Over ISIS in Fallujah	International (World)
NPR	2016	Scalia death	Scalia And Leonard Nimoy: Justice's Death Spurs Conspiracy Theories	Must Reads (U.S.)
NPR	2017	Mueller firing	3 Ways Trump or His Allies Might Try to Disrupt the Mueller Russia Probe	National security (U.S.)
AP	2018	Child immigrants	Immigrant kids seen held in fenced cages at border facility	U.S.

Appendix C

Table C1

Articles rated as no bias

Publication	Year	Topic	Article Title	Category
AlterNet	2017	Betsy DeVos	Protesters Block Education Secretary Betsy DeVos From Entering D.C. Middle School	Politics
AlterNet	2017	Millennials	Millennials and Capitalism Just Don't Mix	Economy
The Wall Street Journal	2016	NYC bombing	New York Bombs Filled with Explosives, Shrapnel;	U.S.
The Wall Street Journal	2016	Orlando shooting	After Orlando Shooting, Nightclubs Reevaluate Security; Police presence is ramped up in several big cities	U.S.
The Wall Street Journal	2016	2016 Presidential election	Gold Prices Reverse Gains Ahead of U.S. Presidential Election;	Business (Economy)
The Wall Street Journal	2017	Covfefe tweet	Trump's 'Covfefe' Tweet Remains a Mystery;	Politics
The Wall Street Journal	2017	Millennials	Why Millennials Crave Cold Coffee	Business (Economy)
The Wall Street Journal	2017	Trump's travel ban	Trump Travel Ban Cuts Demand for Flights to U.S.	Business (Economy)
The Wall Street Journal	2018	California wildfires	California's Largest Utility Pummeled by Wildfire Risks	Business (Economy)
The Wall Street Journal	2018	Aretha Franklin death	Aretha Franklin: Flexible Voice, Fierce Delivery	Life & Arts (Entertainment)
The Wall Street Journal	2018	Roseanne Barr Ambien tweet	Axing of 'Roseanne' Reveals Risks of Star-Driven Sitcoms	Media & Marketing (Entertainment)
The Wall Street Journal	2018	Taylor Swift politics	Cross Country: Marsha Blackburn's Tennessee Turnaround	Opinion
Fox	2016	Rio Olympics	Americans pull off historic sweep in 100m hurdles	Sports
Fox	2016	Trump sexual assault	Trump slams 'corrupt' media, as more women make sexual assault claims	Republicans (U.S.)
Fox	2016	Turkey Coup	Turkish coup attempt spawns conspiracy theories, bolsters Erdogan's clout	Middle East (World)
Fox	2017	Covfefe tweet	Trump tweets 'covfefe,' causes Twitter frenzy	White House (Politics)
Fox	2017	Melania Trump moves into White House	Melania Trump, son Barron move into the White House	First family (Life)
Fox	2017	Robert Mueller firing	Trump: I have no plans to fire Robert Mueller over Russia investigation	White House (Politics)
Fox	2018	Black Panther movie	'Black Panther' review: Marvel tackles politics, race with jaw-dropping new film	Review (Entertainment)
Fox	2018	California wildfires	California's Camp Fire now '100 percent contained,' as grim search by rescuers continues	Weather (Environment)
The Washington Times	2017	Amazon buys Whole Foods	Amazon deal for Whole Foods could bring retail experiments	Technology
The Washington Times	2017	United drags man off plane	Doctor dragged off United flight was convicted of multiple felony drug charges in 2004	U.S.

Appendix C

The Washington Times	2018	Aretha Franklin death	Trump pays tribute to Aretha Franklin: 'Great woman with a wonderful gift from God'	Politics
The Washington Times	2018	Trump as a "stable genius"	Trump, responding to allegations in book, says he's a 'stable genius'	Politics
Breitbart	2017	Millennials	NBER Study: Over Half of California Millennials Will Be Poorer than Parents	Economy
Breitbart	2018	Aretha Franklin death	Donald Trump Honors Aretha Franklin: Her Voice 'A Wonderful Gift from God'	Entertainment
Breitbart	2018	Hurricane Florence	Donald Trump Warns North and South Carolina Residents About Hurricane Florence	Weather (Environment)
Breitbart	2018	Elon Musk on social media	Elon Musk Deletes Facebook Pages of Tesla and SpaceX	Tech
Breitbart	2018	Roseanne Barr Ambien tweet	Roseanne Barr Says Fans Inspiring Her to Fight Back Against Show's Cancellation	Media (Entertainment)
Breitbart	2018	Taylor Swift politics	Donald Trump Slams Taylor Swift After Marsha Blackburn Smear	Politics
Newsmax	2016	NYC bombing	Bombs Were Filled with Shrapnel, Designed for Maximum Casualties	America (U.S.)
Newsmax	2016	Rio Olympics	Israeli Athletes in Rio Face 'Shocking' Hostility from Muslims	Newsfront (Sports)
Newsmax	2018	California wildfires	Trump Surveys California Fire Scars	Weather (Environment)
Newsmax	2018	Fire and Fury book release	'Fire and Fury' Reveals a Trump We Already Know	Media (Entertainment)
Newsmax	2018	Aretha Franklin death	Lawmakers Want Congressional Gold Medal for Aretha Franklin	News (Entertainment)
National Review	2016	Global Warming hoax	Trump the Climate-Slayer	Environment
National Review	2017	United drags man off plane	United Is Why People Hate Capitalism	Economy & Business
National Review	2018	Winter Olympics	Between the Koreas, a Dangerous and Delicate Dance	World
The Nation	2018	Aretha Franklin death	Aretha Franklin—Musical Genius, Truth Teller, Freedom Fighter	Media (Entertainment)
HuffPost	2016	Brexit	Bernie Sanders Explains What Democrats Can Learn From Brexit	Politics
HuffPost	2016	Penn Chapo interview	A U.S. Trial For El Chapo Could Expose More Than Some Americans Want	Media (U.S.)
HuffPost	2017	#MeToo	Topless Femen Activists Stage #MeToo Protest Against the Patriarchy At The Vatican	Women (Life)
HuffPost	2018	March for our Lives	The March for Our Lives Is Proof That Generation Z Can't Be Stopped	Politics
HuffPost	2018	Stormy Daniels	The Stormy Daniels Story Is A Rape Culture Story	Opinion
HuffPost	2018	Winter Olympics	Out LGBTQ Athletes Have Won Seven Medals At 2018 Olympics The 15-member team has two golds.	Queer Voices (Life)
CNN	2016	China as a Superpower	Defense Secretary Ash Carter: US to 'sharpen military edge' in Asia	Politics
CNN	2016	Melania's plagiarism	Melania admires Michelle? What Donald should learn from that	Opinion
CNN	2016	Penn Chapo interview	U.S. knew of actors' connections to Guzman before meeting, officials say	Politics

Appendix C

CNN	2016	Trump sexual assault	In first rally since lewd audio, Trump touts sports star accused of sexual assault	Politics
CNN	2017	Kennedy Center Honors	Trump to skip Kennedy Center Honors	Politics
CNN	2017	Las Vegas shooting	Weapons cache found at Las Vegas shooter's home	U.S.
CNN	2018	California wildfires	44 dead in California fires as the Camp Fire becomes the deadliest in state history	U.S.
CNN	2018	Aretha Franklin death	What it was like to work with Aretha Franklin	Entertainment
CNN	2018	Hurricane Florence	Hurricane Florence strengthens as 1 million people are told to flee US East Coast	U.S.
CNN	2018	Net Neutrality	Net neutrality rules are now repealed: What it means	Business (Economy)
CNN	2018	Roseanne Barr Ambien tweet	Roseanne Barr blames Ambien: 'I'm not a racist, just an idiot'	Media (Entertainment)
CNN	2018	Taylor Swift politics	Voter registration reportedly spikes after Taylor Swift post	Entertainment
The Atlantic	2016	Brexit	Trust in Government Is Collapsing Around the World	Global (World)
The Atlantic	2016	Rio Olympics	Rio 2016: Where Refugees Are Finally Being Recognized	Global (World)
The Atlantic	2017	Las Vegas shooting	The Worst Mass Shooting in Modern American History:	U.S.
The Atlantic	2017	Royal Wedding	Barack Obama Reflects on Leaving the Presidency	Politics
The Atlantic	2017	Black Panther movie	The Provocation and Power of Black Panther	Culture (Entertainment)
The Atlantic	2018	Fire and Fury book release	Fire and Fury Is a Strange Kind of Coup for Trump Abroad	Global (World)
The Atlantic	2018	Winter Olympics	The Out Olympics	Culture (Sports)
The New York Times	2016	NYC bombing	In a Fictional Pub, a Tale of a Bombing. A Block Away, a Very Real Explosion	New York (U.S.)
The New York Times	2016	Prince death	Hints of Prince's Failing Health Preceded a Sudden Death	Music (Entertainment)
The New York Times	2016	Saudi Arabia 9/11 bill	Saudi Arabia Warns of Economic Fallout if Congress Passes 9/11 Bill	Middle East (World)
The New York Times	2016	Scalia death	After Antonin Scalia's Death, Fierce Battle Lines Emerge	Politics
The New York Times	2016	Turkey Coup	Turkey's Coup That Wasn't	Opinion
The New York Times	2017	Kathy Griffin fired	Trump Lashes Out at Kathy Griffin and Democrats	Politics
The New York Times	2017	Wonder Woman movie	Wonder Woman, Weaponized	Olympics (Sports)
The New York Times	2018	Taylor Swift politics	Taylor Swift, Apolitical No More, Endorses Democratic Candidates in Tennessee	Music (Entertainment)
The New York Times	2018	Winter Olympics	Online, a Diverse Winter Olympics. But on the Ground?	Olympics (Sports)
USA Today	2016	Brexit	Brexit quake reaches the U.S. - Americans say Britain's split from the EU reflects anger that smolders in USA, too	News (World)
USA Today	2016	Gay marriage	Vying for the faithful in Iowa	Culture (Life)
USA Today	2016	NYC bombing	NYC police presence 'bigger than ever' - World leaders, including Obama, coming for annual U.N.	Nation (U.S.)

Appendix C

			meeting to city where explosion injured almost 30 people this weekend	
USA Today	2016	Orlando shooting	ISIL or not, shooting was hate crime against LGBT people	Opinion
USA Today	2016	Prince death	PURPLE REIGN - Remembering the iconic Prince	Life
USA Today	2016	Rio Olympics	Violence rises in Rio's poor communities - Shootings double during first week of Olympics	Sports
USA Today	2016	Zika virus	Zika vaccine waits on Congress - Virus is devastating to the small number of people getting infected during pregnancy	Health
USA Today	2017	Kathy Griffin fired	Did Kathy Griffin break the law with her photo of a decapitated Trump?	Nation (Entertainment)
USA Today	2017	McCain's healthcare bill	The Bubble: McCain no hero, conservatives and liberals say	Politics
USA Today	2017	Melania		
USA Today	2017	Trump moves into White House	First lady Melania Trump, son Barron officially move into the White House	People (Life)
USA Today	2018	Hurricane Florence	Florence marches into South Carolina, having killed at least 5 people despite dramatic rescues	News (Environment)
USA Today	2018	Elon Musk on social media	Who's deleting Facebook? Elon Musk joins the movement	Tech
USA Today	2018	Net Neutrality	Net neutrality's old rules are a thing of the past. Here's what could happen next	Tech
USA Today	2018	Trump and North Korea	North Korea slams talks with Secretary of State Pompeo as 'regrettable'	World
USA Today	2018	Winter Olympics	2018 Winter Olympics medal count: How Team USA is faring in Pyeongchang	Sports
NPR	2016	Black Lives Matter	Bodycam Video Shows Police Shooting of Unarmed 19-Year-Old In Fresno	America (U.S.)
NPR	2016	Harriet Tubman on \$20	Treasury Decides to Put Harriet Tubman on \$20 Bill	America (U.S.)
NPR	2016	Presidential election	As 'Blue Wall' Shows Signs of Strain, Here Are Battlegrounds To Watch For	America (U.S.)
NPR	2016	Refugee crisis	Diplomats Pledge Nearly \$11 Billion In Aid for Syria At Donor Conference	International (World)
NPR	2017	Kennedy Center Honors	President Trump to Skip Kennedy Center Honors, Highlighting Rift with Artists	America (U.S.)
NPR	2017	#MeToo	In the Wake Of #MeToo, More Victims Seek Help for Repressed Trauma	National (U.S.)
NPR	2017	Solar eclipse	Eclipse 2017: One Nation Under the Sun	Special series (Environment)
NPR	2017	Trump's national security	White House Press Secretary Says Trump Fired Flynn As National Security Adviser	Politics
NPR	2017	Women's March	Women's Marches Go Global: Postcards from Protests Around the World	America (U.S.)
NPR	2017	Wonder Woman movie	'Wonder Woman' Smashes Domestic Box Office Record for Female Directors	America (U.S.)
NPR	2018	Black Panther movie	'Black Panther' Is A Superhero Story You Haven't Seen Before — And It's Thrilling	Review (Entertainment)
NPR	2018	California wildfires	More Deaths Are Reported in California Fires	National (Environment)

Appendix C

NPR	2018	Fire and Fury book release	The New Year Rings in with 'Fire and Fury.' It Might Mean A Consequential 2018	Politics
NPR	2018	Aretha Franklin death	Aretha Franklin Was America's Truest Voice	Music (Entertainment)
NPR	2018	Hurricane Florence	Tropical Storm Florence, 'An Uninvited Brute,' Brings Floods, And At Least 5 Deaths	National (Environment)
NPR	2018	March for our Lives	Hundreds of Thousands March for Gun Control Across The U.S.	America (U.S.)
NPR	2018	2018 midterm elections	A Boatload of Ballots: Midterm Voter Turnout Hit 50-Year High	Elections (Politics)
NPR	2018	Roseanne Barr Ambien tweet	Roseanne Barr Says Ambien Played Role in Racist Tweet That Spiked Her Show's Reboot	America (U.S.)
NPR	2018	Winter Olympics	Bigger Than Ever, And More Diverse: Team USA At The 2018 Winter Olympics	Olympics (Sports)
AP	2016	Penn Chapo interview	Drug lord was interested in Mexican actress, not Sean Penn	Entertainment
AP	2016	Trump's taxes	Experts: No clear criminal case over Trump tax disclosure	Business (Economy)
AP	2017	Millennials	Candidate says New Mexico needs more millennials	National (U.S.)
AP	2017	Take a Knee	President's criticisms spark more protests at NFL games	Sports
AP	2017	Trump's travel ban	Trump signs new anti-terror travel ban _without new fanfare	Politics
AP	2017	Wonder Woman movie	Review: 'Wonder Woman' film and star live up to the name	Review (Entertainment)
AP	2018	Stormy Daniels	Trump seeks arbitration in Stormy Daniels case	Nation (U.S.)
AP	2018	Taylor Swift politics	Taylor Swift breaks political silence, backs Tennessee Dems	Entertainment
AP	2018	Trump and North Korea	Doubts on North Korean denuclearization shadow Pompeo visit	World

Appendix D

Table D1

Articles rated as Right leaning

Publication	Year	Topic	Article Title	Category
The Wall Street Journal	2016	Anthony Weiner sexts	FBI Reviewing Newly Discovered Emails in Clinton Server Probe;	Politics
The Wall Street Journal	2016	Black Lives Matter	The Radical Democratic Party; Plus, the myths of Obama and Black Lives Matter	Opinion
The Wall Street Journal	2016	China as a superpower	Trump Plans to Beat China at Its Own Game	Life & Work
The Wall Street Journal	2016	Clinton's emails	Oil Prices Sag on Concerns About OPEC Deal on Output Cut	Markets (Economy)
The Wall Street Journal	2016	Global Warming hoax	Donald Trump's Victory Injects Uncertainty into Climate Accord	Politics
The Wall Street Journal	2016	Gun control	How to Make a Gun at Home;	Books & Art (Entertainment)
The Wall Street Journal	2016	Harriet Tubman \$20	Change for a \$20;	Opinion
The Wall Street Journal	2016	Penn Chapo interview	Desire for Film Fame Led to 'El Chapo' Downfall, Officials Say;	World
The Wall Street Journal	2016	Rio Olympics	Judge Puts Rio Olympics Organizers' Funding Efforts on Hold	Sports
The Wall Street Journal	2016	Trump's taxes	Notes from the Underground;	Opinion
The Wall Street Journal	2016	Turkey Coup	Turkey Coup: A Test of Emerging Faith;	Markets (Economy)
The Wall Street Journal	2017	Betsy DeVos	How to Stop Betsy DeVos; A Congressman suggests the nuclear option for Education	Opinion
The Wall Street Journal	2017	Mueller firing	Mueller Probe Faces New Challenge from Trump Team	U.S.
The Wall Street Journal	2017	Pence visits Middle East	Pence Makes Surprise Trip to Afghanistan	Politics
The Wall Street Journal	2017	Russia sanctions	Bob Mueller's Sideshow; Nunes's Intelligence Committee plods on with the real Russia investigation	Opinion

Appendix D

The Wall Street Journal	2017	Solar eclipse	Marketers Cash in on Eclipse;	Business (Economy)
The Wall Street Journal	2017	United drags man off plane	United's Investors Face Different Risks Than Its Passengers	Markets (Economy)
The Wall Street Journal	2018	Bush death	George Bush's Leadership by Example	Opinion
The Wall Street Journal	2018	Kavanaugh Supreme Court	Confirm Brett Kavanaugh	Opinion
The Wall Street Journal	2018	Nike hires Kaepernick	Anthem Backlash Strikes Nike, Hurting Share Price	Markets (Economy)
The Wall Street Journal	2018	Trump and North Korea	U.S. Calls Out China, Russia on North Korea Energy Caps, Urges U.N. Action	World
The Wall Street Journal	2018	Trump's tariffs	Europe Fears Return of Steel Glut as U.S. Imposes Tariffs	Markets (Economy)
The Wall Street Journal	2018	Winter Olympics	Winter Olympics Prove Financial Headache for Some, Party for Others	World
FOX	2016	Black Lives Matter	Hey, Black Lives Matter, stop terrorizing our cities	Opinion
FOX	2016	Brexit	Singer-songwriter Carole King soothes Brexit-wracked Brits at London concert	Music (Entertainment)
FOX	2016	Clinton's emails	Trump, lawmakers praise FBI's decision to reopen Clinton case	Politics
FOX	2016	Melania's plagiarism	Melania Trump speechwriter admits 'mistake' in lifting Michelle Obama lines	Republican Convention (Politics)
FOX	2016	NYC bombing	Comey vows to review past warnings to FBI about alleged NY bomber	U.S.
FOX	2016	Orlando shooting	Hillary Clinton's response to Orlando proves she's too weak to be president	Opinion
FOX	2016	Prince death	Pop superstar Prince dead at 57	Music (Entertainment)
FOX	2016	Saudi Arabia 9/11 bill	Report: Saudis vow to sell US assets if Congress decides gov was involved in 9/11	Politics
FOX	2017	Amazon buys Whole Foods	Why Amazon buying Whole Foods does not mean the end of grocery stores	Economy
FOX	2017	Covfefe tweet	Trump tweets 'covfefe,' causes Twitter frenzy	White House (Politics)
FOX	2017	Betsy DeVos	Four reasons why Betsy DeVos will be a welcome remedy for our nation's education woes	Opinion
FOX	2017	Kennedy Center Honors	Trumps to skip Kennedy Center Honors to allow for celebration without 'political distraction'	Politics

Appendix D

FOX	2017	Saturday Night Live	EXCLUSIVE: Donald Trump blasts 'disgusting' Madonna and 'disgraced' ex-'SNL' writer on 'Hannity'	Entertainment
FOX	2017	Take a Knee	NFL protests: Stars should get off their knees and lead a constructive conversation on race	Opinion
FOX	2017	Trump's national security	Three candidates emerge to replace Flynn as national security adviser	National security/ Politics
FOX	2017	Trump's Supreme Court nominees	Trump judicial pick who drew ridicule at hearing withdraws	Law/ Politics
FOX	2017	Woman's March	What I learned at the Women's March	Opinion
FOX	2017	Wonder Woman movie	Alamo Drafthouse launches a 'women only' screening of 'Wonder Woman,' sparks outrage	Movies (Entertainment)
FOX	2018	Bush death	George Bush was a remarkable man (and the ultimate wingman)	Opinion
FOX	2018	Child immigrants	I was a childhood immigrant to the US -- Trump is acting wisely and compassionately on illegal immigration	Opinion
FOX	2018	Fire and Fury book release	Bannon expresses 'regret' for comments attributed to him in 'Fire and Fury'	Politics
FOX	2018	Hurricane Florence	Hurricane Florence moving slowly, but 'wreaking havoc' across Carolinas	Weather (Environment)
FOX	2018	March for our Lives	March for Our Lives: Gun control ideas sound good, but are deeply flawed and won't save lives	Opinion
FOX	2018	2018 Midterm elections	The 2018 midterm elections aren't just about candidates, they're also about respect, sovereignty, rule of law	Politics
FOX	2018	Mira Ricardel fired	Mira Ricardel to leave White House, after Melania Trump's office called for her dismissal	Politics
FOX	2018	Roseanne Barr Ambien tweet	ABC cancels 'Roseanne' after Barr's racist tweet	Entertainment
FOX	2018	Trump as a "very stable genius"	Trump says he's 'like, really smart,' 'a very stable genius'	Politics
FOX	2018	Trump and North Korea	The other North Korea threat -- that almost never gets talked about	Opinion
The Washington Times	2016	Anthony Weiner sexts	Anthony Weiner's revenge on Clintons	Opinion
The Washington Times	2016	Clinton's emails	Democrats turn against FBI director, accused of meddling in election	Politics
The Washington Times	2016	Gay marriage	Donald Trump: Supreme Court should have left gay marriage to the states	Politics
The Washington Times	2016	Harriet Tubman \$20	Harriet Tubman to replace Andrew Jackson on \$20 bill after pushback against Alexander Hamilton plan	Politics

Appendix D

The Washington Times	2016	ISIS	Obama tries to go full Orwellian on the Islamic threat	Opinion
The Washington Times	2016	Melania's plagiarism	Media cover Melania Trump's plagiarism scandal more than President Obama's	Culture (Life)
The Washington Times	2016	NYC bombing	Trump excoriated for calling out NYC bombing, but media ignores Clinton doing the same	Opinion
The Washington Times	2016	Prince death	Prince's music changed cultural landscape	National (U.S.)
The Washington Times	2016	Refugee crisis	Obama administration delivered illegal immigrant children to predators, lawmakers say	Politics
The Washington Times	2016	Rio Olympics	Kim Rhode and guns: Fans blame skimpy coverage of shooter's Olympic record on media bias	Olympics (Sports)
The Washington Times	2016	Saudi Arabia 9/11 bill	Sept. 11 families criticize Obama opposition to Saudi probe	National (U.S.)
The Washington Times	2016	Scalia death	Conservatives defend vulnerable senators vowing to block Obama Supreme Court nominee	Politics
The Washington Times	2016	Trump elected	Obama decimated Democratic Party; Hillary Clinton helped	Opinion
The Washington Times	2016	Trump's taxes	Breaking down the questions surrounding Donald Trump's tax return	Millennial Mindset (Life)
The Washington Times	2017	Betsy DeVos	Betsy DeVos wins confirmation as education secretary as VP Pence casts historic vote	Politics
The Washington Times	2017	Kathy Griffin fired	The fall of Kathy Griffin Beheads herself in gruesome Trump display	Communities (Life)
The Washington Times	2017	Kennedy Center Honors	The Latest: Kennedy Center awards to go on without Trump	News (U.S.)
The Washington Times	2017	Las Vegas shooting	Terrorism, race, religion: Defining the Las Vegas shooting	Politics
The Washington Times	2017	Pepsi commercial	Pepsi pulls widely mocked ad featuring Kendall Jenner	Entertainment
The Washington Times	2017	Royal Wedding	Meghan Markle joins royal Christmas lunch at palace	People (Life)

Appendix D

The Washington Times	2017	Saturday Night Live	Alec Baldwin, who parodies Trump on SNL: All jokes aside, 'This country is lost'	Politics
The Washington Times	2017	Solar eclipse	Total solar eclipse crosses lower 48, leaves Americans moonstruck	National (U.S.)
The Washington Times	2017	Take a Knee	Once upon a knee at the old ball game	Opinion
The Washington Times	2017	Trump's travel ban	Trump travel ban: It's not the Muslims. It's the terrorists	Politics
The Washington Times	2017	Trump's national security	How Trump strengthens national security	Opinion
The Washington Times	2018	Bush death	Bush's spirit of compassion, duty soared beyond confines of politics	U.S.
The Washington Times	2018	Child immigrants	Flashback: Obama Administration placed border children with human smugglers	National (U.S.)
The Washington Times	2018	Fire and Fury book release	Michael Wolff's spotty record raises questions about Trump tell-all	Politics
The Washington Times	2018	Aretha Franklin death	Trump pays tribute to Aretha Franklin: 'Great woman with a wonderful gift from God'	Politics
The Washington Times	2018	Hurricane Florence	How liberals' rush to blame Trump for hurricane activity is debunked by science	Weather (Environment)
The Washington Times	2018	Kanye West politics	'Blexit': Kanye West calls for blacks to abandon Democrats	Entertainment
The Washington Times	2018	Kavanaugh Supreme Court	Trump celebrates winning streak, hosts Kavanaugh swearing-in ceremony	Politics
The Washington Times	2018	March for our Lives	March for Our Lives organizers dramatically inflate crowd size	National (U.S.)
The Washington Times	2018	2018 Midterm elections	Women win record number of congressional seats in 2018 midterm elections	Politics
The Washington Times	2018	Net Neutrality	Your internet use could change as 'net neutrality' ends	Technology
The Washington Times	2018	Nike hires Kaepernick	Colin Kaepernick's Nike deal prompts flurry of debate	Football (Sports)

Appendix D

The Washington Times	2018	Russia probe	Trump, conservatives demand probe into Clinton-Steele-Moscow links	Politics
The Washington Times	2018	Serena Williams & sexism in tennis	What Trump and Serena Williams have in common	Sports
The Washington Times	2018	Stormy Daniels	Porn star describes threat over alleged Trump affair in '60 Minutes' interview	Entertainment
The Washington Times	2018	Taylor Swift politics	Trump says Taylor Swift hitting sour notes by endorsing Democrats	News (U.S.)
The Washington Times	2018	Trump and North Korea	Trump's 'down the road' deadline to lift North Korea sanctions frustrates nations eager for trade	World
The Washington Times	2018	Winter Olympics	Bitter cold at Winter Olympics chills global-warming hype	Culture (Sports)
Breitbart	2016	Black Lives Matter	A Short List of Black Lives Matter's Cop-Killing Heroes	Politics
Breitbart	2016	Brexit	BREXIT Signals the End of the New World Order	Europe (World)
Breitbart	2016	Clinton's emails	Hiding Hillary Day 273: Fallout from the FBI Release	Politics
Breitbart	2016	Gay marriage	More Than One Million March on Rome for 'Family Day' to Protest Gay Unions Bill	World
Breitbart	2016	Global Warming hoax	Report: With Global Warming, Expect 'Colder Winters'	Environment
Breitbart	2016	Gun control	Fail: Gun-Controlled Australia Admits Criminals Still Armed, Launches New Amnesty	World
Breitbart	2016	ISIS	Hillary Clinton's Policies Created Conditions for Rise of Islamic State	Politics
Breitbart	2016	Melania's plagiarism	Michelle Obama Copied Alinsky in Speech Melania Trump Allegedly Plagiarized	Politics
Breitbart	2016	NYC bombing	Hillary Calls NYC Explosion 'Bombings,' Slams Trump for 'Bomb'	The media (U.S.)
Breitbart	2016	Orlando shooting	The Left Chose Islam Over Gays. Now 100 People Are Dead Or Maimed In Orlando	Social Justice (U.S.)
Breitbart	2016	Penn Chapo interview	Sean Penn Interviews El Chapo for Rolling Stone... SHOCK: Magazine Gives Fugitive Drug Lord Editorial Control	Media (Entertainment)
Breitbart	2016	Prince death	Aretha Franklin: Was It the Zika Virus That Killed Prince?	Entertainment
Breitbart	2016	Rio Olympics	Five Crazy Crimes That Have Occurred at the Rio Olympics — Already	National security (Sports)
Breitbart	2016	Saudi Arabia 9/11 bill	Paul Sperry: Release the 28 Pages So We Can 'End This Suicidal Phony Alliance with Saudi Arabia'	Radio (Entertainment)

Appendix D

Breitbart	2016	Scalia death	Antonin Scalia's Death Could Mark End of Constitution	U.S.
Breitbart	2016	Trump elected	5 Most Absurd Ways the Left Has Responded to the 2016 Election	Politics
Breitbart	2016	Trump sexual assault	Joe Biden, Groper-in-Chief, Accuses Donald Trump of 'Sexual Assault'	Politics
Breitbart	2016	Trump's taxes	Trump Campaign: NYT Published Tax Docs 'Illegally Obtained' After Editor Promised He Would Go to Jail over Trump Taxes	Politics
Breitbart	2016	Turkey Coup	First the Turkey Coup, Now the Economic Crash	Local (U.S.)
Breitbart	2016	Zika virus	Massive Study Casts Doubt on Zika As Cause of Microcephaly	Health
Breitbart	2017	Covfefe tweet	A Meme Is Born: Donald Trump 'Covfefe' Typo Breaks the Internet	Tech
Breitbart	2017	Betsy DeVos	Reim: Harsh Betsy DeVos Opposition Showed Power of Teachers Unions over Dem Party	Radio (Entertainment)
Breitbart	2017	Kathy Griffin fired	Donald Trump on 'Sick' Beheading Photo: 'Kathy Griffin Should Be Ashamed of Herself'	Entertainment
Breitbart	2017	Kennedy Center Honors	Donald Trump, First Lady Melania to Skip Kennedy Center Honors	U.S.
Breitbart	2017	Las Vegas shooting	'Do Something' After Las Vegas Shooting. OK — What?	Politics
Breitbart	2017	McCain healthcare bill	Promise Broken: Sens. McCain, Murkowski, and Collins Tank Obamacare Repeal Bill	Politics
Breitbart	2017	#MeToo	'This Has Gone Too Far': Tavis Smiley Challenges #MeToo Overreach	U.S.
Breitbart	2017	Mueller firing	Laura Ingraham: Unless Investigated, Mueller's 'Irreparably Tainted' Russia Probe 'Will and Should' Collapse	Politics
Breitbart	2017	Pence visits Middle East	Pence delays Mideast visit: White House	News (U.S.)
Breitbart	2017	Pepsi commercial	Left Blasts Pepsi, Kendall Jenner for 'Appropriating' Anti-Trump 'Resistance' to Sell Soda (Update)	Entertainment
Breitbart	2017	Take a Knee	Why Taking a Knee During the National Anthem Is Not Only Unpatriotic, It's Ignorant	U.S.
Breitbart	2017	Trump's travel ban	Trump Officials Defend 'Common Sense' and 'Constitutional' Executive Order on Travel Restrictions	Politics
Breitbart	2017	Trump's national security	Michael Flynn Resigns as National Security Advisor	Politics
Breitbart	2017	Trump's Supreme Court nominees	Trump Posts Big Wins on Judges as Senate Committee Moves 10 More Judicial Nominees	Politics
Breitbart	2017	United drags man off plane	United CEO: Man dragged off plane was 'disruptive,' 'belligerent'	News (U.S.)
Breitbart	2017	Woman's March	Women's March — See What a Massive, Hillary-Shaped Bullet America Just Dodged?	Politics

Appendix D

Breitbart	2017	Wonder Woman movie	Ms. Magazine: When Will Wonder Woman Be a 'Fat Femme Woman of Color'?	Tech
Breitbart	2018	Black Panther movie	'Black Panther' Review: The Movie's Hero is Trump, the Villain is Black Lives Matter	Entertainment
Breitbart	2018	Bush death	Left Attacks George H.W. Bush After His Death: 'Blatant Racism and Toxic Masculinity,' 'F**k Him'	U.S.
Breitbart	2018	California fires	Jerry Brown Blames Climate Change for California Wildfires	Environment
Breitbart	2018	Child immigrants	13 Facts the Media 'Pros' Don't Want You to Know About 'Family Border Separation'	U.S.
Breitbart	2018	Fire and Fury book release	Report: 'Fire and Fury' Author Wolff Has 'Dozens of Hours' of Tape Recordings Confirming Quotes in Book	The media (Entertainment)
Breitbart	2018	Aretha Franklin death	Donald Trump Honors Aretha Franklin: Her Voice 'A Wonderful Gift from God'	Entertainment
Breitbart	2018	Kavanaugh Supreme Court	Every Witness Named by Kavanaugh's Accusers Sides with Kavanaugh	Politics
Breitbart	2018	2018 Midterm elections	President Trump Suggests Midterm Election Result Was Best Possible Outcome for Dealmaking	Politics
Breitbart	2018	Net Neutrality	Exclusive — Sen. Kennedy on Net Neutrality: Phone Companies Don't Tell You What You Can Say, Neither Should Cable Companies	Politics
Breitbart	2018	Nike hires Kaepernick	Nike Shares Drop as Kaepernick Ad Triggers Calls for Boycott	Economy
Breitbart	2018	Trump as a "very stable genius"	Donald Trump Defends Mental Health: I Am a 'Smart' and 'Very Stable Genius'	Politics
Breitbart	2018	Trump's tariffs	Promises Kept: Donald Trump Signs Steel and Aluminum Tariffs to Protect Workers	Economy
Newsmax	2016	Clinton's emails	Democrats Wage All Out Assault on James Comey	Opinion
Newsmax	2016	ISIS	Trump Tells the Truth on Islamist Terror	Politics & Policy
Newsmax	2016	Melania's plagiarism	Easy to Defend Melania's Speech	Politics
Newsmax	2016	2016 Presidential elections	Win or Lose, Trump's Message Resonates	Politics
Newsmax	2016	Refugee crisis	Rep. Katko: Obama's Failure on Visa Waiver Law Endangers US	Newsfront (U.S.)
Newsmax	2016	Saudi Arabia 9/11 bill	Saudi's Pull Mafia Moves Over 9/11 Role	Politics
Newsmax	2016	Trump elected	Liberals Need to Practice What They Preach on 'Inclusiveness'	Opinion
Newsmax	2016	Trump sexual assault	Trump: Why Don't Women Come Out Against Obama?	Politics
Newsmax	2016	Trump's taxes	No Scandal in Trump's Tax Dodge	Opinion
Newsmax	2016	Turkey Coup	Erdogan's Turkey: An Opportunity for the West	Global talk
Newsmax	2016	Zika virus	Senate Zika Funding Blocked Third Time by Democrats	The wire (U.S.)

Appendix D

Newsmax	2017	Comey's testimony	Comey's Wounds Self-Inflicted, While Trump Unscathed	Politics
Newsmax	2017	Betsy DeVos	Gingrich Rips 'Fascists' Protesting DeVos' Visit to DC Public School	Politics
Newsmax	2017	Las Vegas shooting	Las Vegas Shooting Immediately Politicized	Newsfront (U.S.)
Newsmax	2017	#MeToo	'The Rock' Joining Men in #MeToo Black at Golden Globes	The wire (U.S.)
Newsmax	2017	Millennials	Millennials Would Give Up Sex, Coffee for Hobbies, Study Says	Newsfront (Life)
Newsmax	2017	Mueller firing	Dershowitz: 'Foolish' if Trump Fires Mueller; 'Really Foolish' to Seize Emails	Newsfront (U.S.)
Newsmax	2017	Pepsi commercial	Pepsi's Kendall Jenner TV Ad Falling Flat on Its Pretty Face	TheWire (Entertainment)
Newsmax	2017	Saturday Night Live	Alec Baldwin, 'SNL's Donald Trump, to Host Show for Record 17th Time	TheWire (Entertainment)
Newsmax	2017	Take a Knee	National Anthem Knee-Takers Defacing the Ultimate Monument	America (U.S.)
Newsmax	2017	Trump's travel ban	Trump Travel Ban Faces Key Test in Multiple Courtrooms	Newsfront (U.S.)
Newsmax	2017	Trump's national security	Trump Must Channel Reagan on National Security Post	Opinion
Newsmax	2017	Trump's Supreme Court nominees	A Lasting Trump Future for the Federal Bench	Politics
Newsmax	2017	Woman's March	Protesters Litter Trump International, Trump Tower with Signs	Politics
Newsmax	2017	Wonder Woman movie	Movie Theater Doubles Down on Women-Only 'Wonder Woman' Showings	Entertainment
Newsmax	2018	Bush death	George H.W. Bush Will Be Ranked Among America's Greatest Presidents	America (U.S.)
Newsmax	2018	Child immigrants	Obama Separated Immigrants Too, But Democrats Now See Midterm Issue	Opiniom
Newsmax	2018	Hurricane Florence	Trump: 'We Will Be 100 Percent' With Hurricane Florence Victims	Hot topics (Environment)
Newsmax	2018	Kavanaugh Supreme Court	Kavanaugh and the Death and Decline of Democracy	Opinion
Newsmax	2018	March for our Lives	At 'March for Our Lives,' the Young Lead the Foolish	Opinion
Newsmax	2018	2018 Midterm elections	2018 Midterms: Opportunity Lost to Keep GOP Majority	Politics
Newsmax	2018	Net Neutrality	FCC Chair, Family Still Receive Death Threats Over Net Neutrality	Politics
Newsmax	2018	Nike hires Kaepernick	Trump Attention Helped Make Kaepernick Controversy What It Is Today	America (Sports)
Newsmax	2018	Roseanne Barr Ambien tweet	Roseanne Barr, Samantha Bee and the Need for Consistency	Entertainment
Newsmax	2018	Stormy Daniel's allegations	Judge Rejects Trial Demand in Stormy Daniels Case	TheWire (U.S.)

Appendix D

Newsmax	2018	Taylor Swift politics	Marsha Blackburn Responds to Taylor Swift	Politics
Newsmax	2018	World cup	Nike Won't Supply Iran With Cleats for FIFA World Cup	Newsfront (Sports)
National Review	2016	Anthony Weiner sexts	Wait, the New Emails Came from Anthony Weiner's Devices?	Politics & policy
National Review	2016	Black Lives Matter	Black Lives Matter: Radicals Using Moderates to Help Tear America Apart	Politics & policy
National Review	2016	Brexit	Brexit Has Nothing to Do with Donald Trump	Politics & policy
National Review	2016	Clinton's emails	'We Need to Clean This Up': More Evidence Obama Lied about Hillary's Private E-mails	Politics & policy
National Review	2016	Gun control	Does Hillary Believe in an Individual Right to Bear Arms or Not?	Politics & policy
National Review	2016	Harriet Tubman \$20	Harriet Tubman Is a Great Choice, Not a Politically Correct One	Culture (U.S.)
National Review	2016	Melania's plagiarism	No, Melania Trump's Plagiarism Was Not "Racist"	Politics & policy
National Review	2016	NYC bombing	Countering Islamist Terror Requires a Strategy, Not Denial	Politics & policy
National Review	2016	Orlando shooting	The Orlando Shooting Launches a War on Christianity	Politics & policy
National Review	2016	Penn Chapo interview	If Stupidity Were a Crime, Sean Penn Would Be the Fugitive	Culture (Entertainment)
National Review	2016	2016 Presidential elections	The Blue Wave - Republicans must adapt to a diversifying electorate or lose their influence	Feature (Politics)
National Review	2016	Prince death	Rain or Reign?	Culture (U.S.)
National Review	2016	Refugee crisis	Raul Castro Is Launching a 'Weapon of Mass Migration' against the U.S.	National Security & defense (U.S.)
National Review	2016	Rio Olympics	U.S. Has Done Fine with No Government Department of Sports	National Security & defense (U.S.)
National Review	2016	Saudi Arabia 9/11 bill	First, Let's Get the Facts on Saudi and Iranian Involvement in 9/11	Politics & policy
National Review	2016	Scalia death	Justices or Ayatollahs?	Politics & policy
National Review	2016	Trump elected	Trump Bulldozes Blue Wall, Wins White House	Politics & policy
National Review	2016	Trump sexual assault	Trump Is Validating the Left's Crudest Stereotypes of the Right	Culture (U.S.)
National Review	2016	Trump's taxes	White House: No Plans to Release Trump Tax Returns After NYT Story	White House (Politics)
National Review	2016	Zika virus	Federal Agencies' Rampant Incompetence Fuels Zika Outbreak	Politics & policy
National Review	2017	Betsy DeVos	What Is the Democratic Party?	Politics & policy

Appendix D

National Review	2017	Las Vegas shooting	The Passionate Non-Sequiturs of the Gun Debate	Politics & policy
National Review	2017	#MeToo	The Sexes after Weinstein	Culture (Life)
National Review	2017	Millennials	Social Security Makes Millennials Less Financially Responsible	Economy & business
National Review	2017	Mueller firing	Criticism of Mueller Isn't a Threat to the Rule of Law	Politics & policy
National Review	2017	Take a Knee	I Understand Why They Knelt	Culture (Life)
National Review	2017	Trump's travel ban	Trump's Revised Travel Ban is Just One of Many Good Developments in the War against Jihadists	Politics & policy
National Review	2017	Trump's Supreme Court nominees	But Gorsuch ... and Other Excellent Judicial Picks ... and a Tax Cut ... and Major Deregulatory Actions ... and Immigration Enforcement ... and the End of the Individual Mandate ... and a Roll Back of the HHS Mandate ...	Politics & policy
National Review	2017	Woman's March	The Top Five Worst Speeches at the Women's March on Washington	Politics & policy
National Review	2018	Black Panther movie	Black Panther's Circle of Hype	Film & TV (Entertainment)
National Review	2018	Bush death	Partisans Use 41 to Attack 45	White house (Politics)
National Review	2018	Child immigrants	Bordering on Insanity	U.S.
National Review	2018	Fire and Fury book release	Is Fire and Fury Fizzling?	Politics & policy
National Review	2018	Aretha Franklin death	Aretha Franklin, an American Original	Culture (Entertainment)
National Review	2018	Kavanaugh Supreme Court	Brett Kavanaugh Should Be Angry	Law & the courts (U.S.)
National Review	2018	March for our Lives	The Teenage Demagogues	Politics & policy
National Review	2018	2018 midterm elections	Some Election Implications	Elections (Politics)
National Review	2018	Nike hires Kaepernick	The Party of Kaepernick	Elections (Politics)
National Review	2018	Serena Williams & sexism in tennis	Serena Behaved Abysmally	Sports
National Review	2018	Taylor Swift politics	Taylor Swift Buys the Democrats' Marketing; Tennesseans Shouldn't	Culture (Life)
National Review	2018	Trump's tariffs	Man of Steel	Economy & business
HuffPost	2017	Mueller firing	Sen. Mark Warner Warns Trump: Firing Robert Mueller Would Be A 'Gross Abuse of Power'	Politics
CNN	2017	Pence visits Middle East	Pence and Abbas won't meet, vice president's office confirms	Politics
CNN	2018	Kavanaugh Supreme Court	How Kavanaugh maneuvered to win his confirmation fight	Politics

Appendix D

CNN	2018	Elon Musk social media	Elon Musk deletes Facebook accounts for Tesla and SpaceX	Business (Economy)
USA Today	2017	Take a Knee	Anthem kneeling isn't aimed at veterans, and other NFL protest misconceptions	Nation now (U.S.)
USA Today	2018	Bush death	George H.W. Bush, the 41st U.S. president and father of the 43rd, has died at age 94	Politics
USA Today	2018	Kanye West politics	Kanye West distances himself from Blexit and Trump: 'I've been used to spread messages'	Life & people
USA Today	2018	Taylor Swift politics	Taylor Swift's political statement is a redeeming moment for the star	Life & music
NPR	2016	Anthony Weiner sexts	FBI Led Back to Clinton Email Server Case By Anthony Weiner Investigation	Politics
NPR	2016	Brexit	After Brexit Vote, U.K. Sees A Wave of Hate Crimes And Racist Abuse	Politics & policy
NPR	2016	ISIS	Iraqi Forces Claim Victory Over ISIS in Fallujah	International (World)
NPR	2016	Prince death	No Signs That Prince's Death Was A Suicide, Sheriff Says	America (U.S.)
AP	2016	2016 Presidential elections	Competing claims of voter fraud, intimidation raise tensions	Elections (Politics)
AP	2016	Prince death	Beyond music, Prince's legacy includes black activism	Entertainment
AP	2017	Betsy DeVos	Protesters greet DeVos in her first visit to a public school	Politics

Appendix E

Table E1				
<i>Articles rated as left leaning</i>				
Publication	Year	Topic	Article title	Category
AlterNet	2016	Anthony Weiner sexts	John Oliver Destroys Anthony Weiner Who Has Dangerously United Americans in Hatred of Him	Leadership (Entertainment)
AlterNet	2016	Black Lives Matter	FBI Greenlights Crackdown on Black Lives Matter Protesters	Politics
AlterNet	2016	Brexit	Did Trump-Style Islamophobia Break Up the European Union?	Politics
AlterNet	2016	China as a superpower	American Global Power Is Being Challenged by Russia and China – What Does the Future Hold?	Politics
AlterNet	2016	Clinton's emails	10 Ways Trump Broke the Law and Got Away with It: Putting the Latest Clinton Email Media Frenzy in Perspective	Politics
AlterNet	2016	Global Warming hoax	Donald Trump's Global Warming Stance Denounced by China	Environment
AlterNet	2016	Gun control	Michael Douglas, Alec Baldwin and Tony Goldwyn Join Brave New Film's Gun Control Effort	Media
AlterNet	2016	Harriet Tubman \$20	Harriet Tubman to Share \$20 Bill with President Who Called for Some Abolitionists to 'Atone ... With Their Lives'	Economy
AlterNet	2016	ISIS	ISIS Must Love Trump	Politics
AlterNet	2016	Melania's plagiarism	Melania's Plagiarism Caps Really Weird Day One at RNC	Politics
AlterNet	2016	NYC bombing	NYC Bombing Suspect Taken into Custody After Shootout	News (U.S.)
AlterNet	2016	Orlando shooting	How the Corporate Media's Spin on the Orlando Shooting Is Boosting America's Dirty Wars in Muslim-Majority Countries	World
AlterNet	2016	Penn Chapo interview	When Hollywood Calls, Who Can Resist?	Media (Entertainment)
AlterNet	2016	2016 Presidential elections	10 Shocking 2016 Election Facts: Old Political Assumptions Are out the Window	Politics
AlterNet	2016	Prince death	What We Lost When We Lost Prince	Culture (Life)
AlterNet	2016	Refugee crisis	Is the EU Using Its Refugee Crisis as an Excuse to End Everyone's Freedom to Travel?	Economy
AlterNet	2016	Scalia death	How Scalia's Death Could Upend the Presidential Election and Spark the Biggest Political Battle of Our Times	Politics
AlterNet	2016	Trump elected	Trump Elected President by Wave of Angry White Voters Across Upper Midwest and South	Politics & elections
AlterNet	2016	Trump's sexual assault	In Nazi-Like Speech, Trump Responds to Sexual Assault Claims with Broad Conspiracy Theory Designed to Foment Mayhem	Politics
AlterNet	2016	Trump's taxes	Did Marla Maples Leak Donald Trump's Tax Return?	Economy

Appendix E

AlterNet	2016	Turkey coup	Erdogan Is Using the Failed Coup to Get Rid of the Last Vestiges of Secular Turkey	World
AlterNet	2016	Zika virus	How the Christian Right's Sex Hangups Turn Zika into a Bigger Crisis	Health
AlterNet	2017	Comey's testimony	Will Comey's Testimony Take Trump Down?	Politics
AlterNet	2017	Covfefe tweet	Sucked in Again: With One Stupid Tweet, Donald Trump Grabs Much of Our Attention	Politics
AlterNet	2017	Las Vegas shooting	So Few Americans Understand What the Second Amendment Is Really About – or Its Dark History	Rights (U.S.)
AlterNet	2017	McCain's Healthcare vote	Let's Not Ignore the Real Heroes Who Saved Our Health Care System from Total Ruin	Health
AlterNet	2017	#MeToo	6 Big Victories for Feminism in 2017, Despite Donald Trump and Other Toxic Men	Rights (Life)
AlterNet	2017	Mueller firing	Behind the Right-Wing Plot Against Mueller	The Right Wing (U.S.)
AlterNet	2017	Pence visits the Middle East	Team Trump Add Insult to Injury for the Palestinians	Politics
AlterNet	2017	Royal Wedding	The British Government Is Really This Afraid of Upsetting Trump	Politics
AlterNet	2017	Russia sanctions	Trump's Commerce Secretary Owns Stake in Russian Companies While He Oversees Potential Sanctions	Politics
AlterNet	2017	Saturday Night Live	Alec Baldwin Brutally and Hilariously Mocks Trump's Press Conference on SNL	Media (Entertainment)
AlterNet	2017	Solar eclipse	Trump Looked Directly at the Solar Eclipse Without Eye Protection, Because of Course He Did	Environment
AlterNet	2017	Take a Knee	Following Trump's Example, High School Violates Students' Constitutional Right by Banning Taking a Knee	Politics
AlterNet	2017	Trump's national security	5 Disturbing New Revelations About Trump's Dysfunctional National Security Council	Politics
AlterNet	2017	Trump's Supreme Court nominees	Trump Considered Rescinding His Neil Gorsuch Nomination Because He Wasn't Sufficiently 'Loyal': Report	Politics
AlterNet	2017	United drags man off plane	Jimmy Kimmel Has Only Two Words for United Airlines CEO After Its Viral Video Disgrace	Media (Entertainment)
AlterNet	2017	Women's March	The Women's March Heralds a Renaissance of Resistance	Rights (Life)
AlterNet	2017	Wonder Woman movie	Why Gal Gadot Is a Real-Life Wonder Woman	Culture (Entertainment)
AlterNet	2018	Black Panther movie	Did Anyone Else Pick Up on the Disturbing Messages in Marvel's 'Black Panther'?	Culture (Entertainment)

Appendix E

AlterNet	2018	Bush's death	George Herbert Walker Bush and the myth of the 'good' Gulf War	Politics
AlterNet	2018	Fire and Fury book release	Does 'Fire and Fury' Passage About Nikki Haley Imply She and Trump Are Having an Affair?	Media (Entertainment)
AlterNet	2018	Kanye West politics	Kanye West Just Turned His Back on Trump in the Most Public Way Possible: 'I've Been Used'	Politics
AlterNet	2018	Kavanaugh Supreme Court	'Out of Control': Brett Kavanaugh Stuns Observers with Furious Opening Statement Revealing Himself as a Right-Wing Operative	The Right Wing (Politics)
AlterNet	2018	2018 Midterm elections	The Blue Wave that Everyone Missed: Democrats Had a Huge Win in the Midterms – But We're Only Realizing It Now	Politics
AlterNet	2018	Nike hires Kaepernick	A Right-Wing Musician Cut Up His Nike Socks to Protest Kaepernick – And Was Immediately Mocked by the Internet	Media (Entertainment)
AlterNet	2018	Russia probe	'Pages of lies': Democratic lawmaker reveals how Devin Nunes has buried key evidence in the Russia probe	Politics
AlterNet	2018	Stormy Daniel's allegations	Trump Finally Addresses Stormy Daniels Story – Her Lawyer Is Thrilled	Politics
AlterNet	2018	Trump as a "very stable genius"	Trump, the 'Very Stable Genius,' Is Falling Apart as Mueller Seeks Interview	Politics
AlterNet	2018	Trump & North Korea	'Trump Got Played': Foreign Policy Expert Explains How the 'Fatally Naive' North Korea Summit Could Lead to War	Politics
AlterNet	2018	Trump's tariffs	Trump's Steel and Aluminum Tariffs Are Absolutely Horrible for American Workers	Economy
AlterNet	2018	Winter Olympics	Winter Olympics Threatened by Global Warming, but NBC Won't Acknowledge It	Environment
Fox	2017	Scalia death	Confusion over Scalia's death stirs sideshow debate, conspiracy theories	Politics
Fox	2017	Pepsi commercial	Kendall Jenner's Pepsi ad sparks social media outcry	Entertainment
Breitbart	2017	Saturday Night Live	Alec Baldwin to Revive Donald Trump Role on 'SNL' Following Inauguration	Entertainment
Breitbart	2018	Winter Olympics	Gay U.S. Olympian Ramps Up Attacks on Vice President Mike Pence	Politics
Newsmax	2017	Pence visits the Middle East	Pence Makes Surprise Visit to Afghanistan for war Meetings	Global talk
Newsmax	2018	Fire and Fury book release	'Fire and Fury' Reveals a Trump We Already Know	Media (Entertainment)
National Review	2017	Comey's testimony	Comey's Weak Case	Editorial (Politics)
The Nation	2016	Black Lives Matter	Black Lives Still Must Matter, Even After Dallas	Police and Law Enforcement (U.S.)

Appendix E

The Nation	2016	Clinton's emails	The Dangerous Cowardice of James Comey	Politics
The Nation	2016	Gay marriage	In a Gay-Friendly Age, Poor LGBT New Yorkers Still Struggle	Culture (Life)
The Nation	2016	Gun control	America's Refusal to Control Firearms Is Killing Teens at an Absurd Rate	U.S.
The Nation	2016	ISIS	ISIS Is a Symptom, Not the Cause, of the Middle East's Disintegration	World
The Nation	2016	Orlando shooting	Let's Make Gun Control a Gay Issue	Politics
The Nation	2016	2016 Presidential elections	An Unabashed Misogynist Is in Charge of Our Country. Now What?	Feature (Politics)
The Nation	2016	Refugee crisis	The US Has Its Own Refugee Crisis—One We Created, and Now Must Solve	Economy
The Nation	2016	Rio Olympics	Teachers and Students Occupy Schools in the Shadow of Olympic Rio	Events (Sports)
The Nation	2016	Scalia death	America Changed. Antonin Scalia Never Did.	U.S.
The Nation	2016	Turkey coup	To Fight the Coup, Erdogan Adopted Tools Pioneered by Protesters	World
The Nation	2017	Comey's testimony	Can Robert Mueller Really Hold Trump Accountable?	Politics
The Nation	2017	Covfefe tweet	The Covfefe Is Often Worse Than the Crime	Politics
The Nation	2017	Betsy DeVos	Betsy DeVos Has Been Confirmed. Now the Fight Really Begins.	Politics
The Nation	2017	Las Vegas shooting	The American Impulse to Equate Guns with Freedom and Masculinity with Violence Is Killing Us	Politics
The Nation	2017	McCain's Healthcare vote	John McCain Isn't a Hero for Not Killing His Constituents	Health
The Nation	2017	#MeToo	From Bill O'Reilly to Al Franken: #MeToo in 2017	Politics
The Nation	2017	Mueller firing	If Trump Fires Mueller, We Must Impeach	
The Nation	2017	Pence visits the Middle East	Palestine's First Intifada Is Still a Model for Grassroots Resistance	World
The Nation	2017	Take a Knee	Taking a Knee Is Not About Abstract Unity but Racial Justice	Culture (Sports)
The Nation	2017	Trump's travel ban	The Real Goal of Trump's Travel Ban Is to Make America White Again	Politics
The Nation	2017	Trump's national security	All Those Donald Trump Headlines Are Distracting You from the Really Scary News	Politics
The Nation	2017	Women's March	Protests Are Putting Trump on Notice: You're in for a Fight	Culture (Life)
The Nation	2018	Black Panther movie	Black Panther and Did You Wonder Who Fired the Gun?	Culture (Entertainment)
The Nation	2018	Bush's death	George H.W. Bush, Icon of the WASP Establishment—and of Brutal US Repression in the Third World	Politics

Appendix E

The Nation	2018	California fires	The Fire This Time California's Camp Fire is a reckoning, an alarm shrieking that our leaders won't save us from climate change—but perhaps we can	Events (Environment)
The Nation	2018	Child immigrants	Trump's Border: Gitmo for Kids	Politics
The Nation	2018	Fire and Fury book release	A President Who Attempts to Ban Books That Expose Him Should Be Impeached	Politics
The Nation	2018	Hurricane Florence	Hurricane Florence Will Be Made Worse by 'Stupid Political Decisions'	Events (Environment)
The Nation	2018	March for our Lives	Get Ready to 'March for Our Lives' and Vote Out the Gun Lobby	Politics
The Nation	2018	2018 Midterm elections	9 Lessons From the 2018 Midterms	Politics
The Nation	2018	Nike hires Kaepernick	On Colin Kaepernick's Nike Ad: Will the Revolution Be Branded?	Culture (Sports)
The Nation	2018	Roseanne Barr Ambien tweet	Roseanne Tried to Use 'Roseanne' to Prove that Trump Voters Aren't Racist. There Was Just 1 Problem.	Culture (Entertainment)
The Nation	2018	Serena Williams's and sexism in tennis	Serena Williams and the Policing of Anger	Events (Sports)
The Nation	2018	Stormy Daniel's allegations	Stormy Daniels Going Public Is Pussy Grabbing Back	Politics
The Nation	2018	Trump as a "very stable genius"	Donald Trump Is Not a Stable Genius, but You Should Be	Politics
The Nation	2018	Trump's tariffs	Don't Be Fooled by Trump's Tariffs. He's Running a Con.	Politics
The Nation	2018	Winter Olympics	Mike Pence and Japanese Leader Shinzo Abe Rain on South Korea's Olympics Parade	Events (Sports)
The Nation	2018	World Cup	FIFA's Weak Attempts to Fight Racism Are on Display at the World Cup in Russia	Events (Sports)
HuffPost	2016	Anthony Weiner sexts	Donald Trump Invokes 'Pervert' Anthony Weiner in Ad Attacking Hillary Clinton	Politics
HuffPost	2016	Black Lives Matter	Canadian Vocal Group Condemns Member For Changing National Anthem Lyric To 'All Lives Matter'	Black Voices (Life)
HuffPost	2016	Brexit	Bernie Sanders Explains What Democrats Can Learn From Brexit	Politics
HuffPost	2016	China as a superpower	Welcome To Our Dying Empire	Politics
HuffPost	2016	Clinton's emails	Sarah Paulson Gives Hillary Clinton Emails the Dramatic Twist Donald Trump Wanted	Comedy (Entertainment)

Appendix E

HuffPost	2016	Global Warming hoax	This Is VERY Bad for The Fight Against Climate Change	Politics
HuffPost	2016	Gun control	Gun Safety Group Releases 'Thoughts and Prayers' Book Mocking Lawmakers	Politics
HuffPost	2016	Harriet Tubman \$20	The Little-Known Reason Why Harriet Tubman on the \$20 Bill Is So Significant	Politics
HuffPost	2016	Melania's plagiarism	Trump Campaign Uses Every Excuse in The Book to Explain Melania's Plagiarism	Politics
HuffPost	2016	NYC bombing	Donald Trump on New York Bombing: 'I Called It Before The News'	Politics
HuffPost	2016	Orlando shooting	How Religious Leaders Are Responding to The Orlando Shooting	Religion (Life)
HuffPost	2016	Penn Chapo interview	A U.S. Trial For El Chapo Could Expose More Than Some Americans Want	Media (U.S.)
HuffPost	2016	Scalia death	What Scalia's Death Means for Abortion Rights in This Country	Women (Politics)
HuffPost	2016	Trump elected	Economic Insecurity, Racism and Sexism All Help Explain Trump's Victory	Politics
HuffPost	2016	Trump's sexual assault	Trump's Morning: Attack Biden, Female Accusers, Fellow Republicans, Electoral Integrity	Politics
HuffPost	2016	Trump's taxes	Seth Meyers Mocks 'Scam Artist' Donald Trump Over His Tax Returns	Comedy (Entertainment)
HuffPost	2016	Zika virus	Biden To GOP: If You Care About Children, Don't Let Them Get Zika	Politics
HuffPost	2017	Amazon buys Whole Foods	Amazon Is Buying Whole Foods in A Quest to Beat Walmart with Luxury	Business (Economy)
HuffPost	2017	Covfefe tweet	Hillary Clinton Serves Donald Trump's Covfefe Right Back at Him	Entertainment
HuffPost	2017	Betsy DeVos	Betsy DeVos Tries to Enter Public School, Gets Blocked by Protesters	Politics
HuffPost	2017	Kathy Griffin fired	Kathy Griffin Slams Trump Family For 'Trying to Ruin My Life'	Entertainment
HuffPost	2017	Kennedy Center Honors	Donald Trump Will Skip Kennedy Center Honors to Avoid 'Political Distraction'	Politics
HuffPost	2017	Las Vegas shooting	The Double Standard in How the Media Is Portraying The Las Vegas Shooter	Black voices (U.S.)
HuffPost	2017	McCain's Healthcare vote	George Takei Has a Revelation About John McCain and Severus Snape	Entertainment
HuffPost	2017	Melania moves into the White House	It's Official! Melania Trump Now Resides in The White House	Life
HuffPost	2017	#MeToo	Topless Femen Activists Stage #MeToo Protest Against the Patriarchy at The Vatican	Women (Life)
HuffPost	2017	Millennials	Author of The Millennial's Guide to Changing the World	Life

Appendix E

HuffPost	2017	Pepsi commercial	Pepsi's Portrayal of a Muslim Woman Was as Tone-Deaf as The Rest Of The Ad	Women (Life)
HuffPost	2017	Royal Wedding	Here's What to Expect at Prince Harry and Meghan Markle's Royal Wedding	Weddings (Life)
HuffPost	2017	Saturday Night Live	Alec Baldwin's Trump Takes on Obamacare and Golden Showers On 'SNL'	Comedy (Entertainment)
HuffPost	2017	Take a Knee	Actor Michael Rapaport Takes A Knee, Unloads On 'Dumb Motherf—ker' Donald Trump	Entertainment
HuffPost	2017	Trump's travel ban	Trump's Order Is Still a Muslim Ban, Faith Groups Say	Politics
HuffPost	2017	Trump's national security	Trump Administration Increasingly at Odds with U.S. Intelligence Community	Politics
HuffPost	2017	Trump's Supreme Court nominees	Trump's Biggest Win In 2017? His Judicial Confirmations.	Politics
HuffPost	2017	United drags man off plane	The 2017 Annual "Stupid Company" Award: United Wins!	Opinion
HuffPost	2017	Wonder Woman movie	Fox News Mulls Over Whether Wonder Woman Is 'American' Enough	Entertainment
HuffPost	2018	Bush's death	The Media Is Erasing George H.W. Bush's Catastrophic Harm to LGBTQ People	Opinion
HuffPost	2018	California fires	The Camp Fire Is Now the Deadliest Wildfire in California's History	U.S. News
HuffPost	2018	Child immigrants	Trump's 'Zero-Tolerance' Immigration Policy Is Tearing Families Like These Apart	Politics
HuffPost	2018	Fire and Fury book release	The Wildest Moments From 'Fire and Fury,' The Trump Book Everyone Is Talking About	Politics
HuffPost	2018	Kanye West politics	Kanye West Claims He's Done with Politics, Says He's 'Been Used'	Politics
HuffPost	2018	Kavanaugh Supreme Court	Confirming Brett Kavanaugh Now Would Destroy the Supreme Court as We Know it	Opinion
HuffPost	2018	2018 Midterm elections	The Midterm Elections Were a Stress Test for Democracy and It Did Not Go Well	Politics
HuffPost	2018	Net Neutrality	Net Neutrality Can Still Be Saved	Opinion
HuffPost	2018	Nike hires Kaepernick	Some White People Are Very Mad at Nike and Colin Kaepernick	Black voices (Life)
HuffPost	2018	Russia probe	Donald Trump Explains Reasons Behind Those Bizarre Russia Investigation Tweets	World news
HuffPost	2018	Taylor Swift politics	Conservatives Are Losing Their Minds Because Taylor Swift Backs a Democrat	Politics

Appendix E

HuffPost	2018	Trump as a "very stable genius"	Mr. Trump and Mr. Ed: A very stable genius and a genius in the stable	Opinion
HuffPost	2018	Trump & North Korea	Trump Is Reportedly Sending New Pal Kim Jong Un an Awkward Gift	News (U.S.)
HuffPost	2018	Trump's tariffs	Trump's Tariffs Could Make These Popular Products More Expensive	Politics
CNN	2016	Brexit	4 reasons why Brexit doesn't mean Donald Trump will win	Politics
CNN	2016	Clinton's emails	Trump reinvigorated by FBI Clinton probe	Politics
CNN	2016	Global Warming hoax	Climate change denialism is so last season	Opinion
CNN	2016	Gun control	Trump or Clinton, stricter gun control is inevitable	Opinion
CNN	2016	Harriet Tubman \$20	What Trump, Carson don't get about Harriet Tubman	Opinion
CNN	2016	Melania's plagiarism	Melania admires Michelle? What Donald should learn from that	Opinion
CNN	2016	NYC bombing	Trump says 'bomb went off in New York'	Politics
CNN	2016	Penn Chapo interview	U.S. knew of actors' connections to Guzman before meeting, officials say	Politics
CNN	2016	2016 Presidential elections	Election Day forecast: Rain possible for a few key states	Politics
CNN	2016	Prince death	Obama on the death of Prince: 'It's a remarkable loss'	U.S.
CNN	2016	Refugee crisis	Refugee crisis: Why Scandinavian countries are trying to look bad	Opinion
CNN	2016	Rio Olympics	Is this the most feminist Olympics ever?	Opinion
CNN	2016	Saudi Arabia 9/11 bill	White House threatens to veto 9/11 lawsuit bill	Politics
CNN	2016	Trump's sexual assault	In first rally since lewd audio, Trump touts sports star accused of sexual assault	Politics
CNN	2016	Trump's taxes	Trump's vision of capitalism is amoral	Opinion
CNN	2016	Zika virus	Congress' cynical Zika game threatens all Americans	Opinion
CNN	2017	Comey's testimony	James Comey just went nuclear on Donald Trump	Politics
CNN	2017	Covfefe tweet	'Covfefe' tells you all you need to know about Donald Trump	Politics
CNN	2017	Betsy DeVos	Betsy DeVos appears to have plagiarized quotes for Senate questionnaire	Politics
CNN	2017	Kathy Griffin fired	Kathy Griffin: Trump is 'trying to ruin my life' after photo scandal	Entertainment
CNN	2017	McCain's Healthcare vote	John McCain's maverick moment	Politics

Appendix E

CNN	2017	Mueller firing	How Fox News and President Trump create an anti-Mueller 'feedback loop'	Media (Entertainment)
CNN	2017	Saturday Night Live	Alec Baldwin, 'SNL' Donald Trump impersonator, will host show on February 11	Media (Entertainment)
CNN	2017	Take a Knee	There is no constitutional right to take a knee while you're at work	Opinion
CNN	2017	Trump's Supreme Court nominees	Exclusive: Grassley urges Trump to reconsider controversial judicial picks	Politics
CNN	2017	Women's March	Women's March pits new activists with a 'nasty' message against Trump	Politics
CNN	2017	Wonder Woman movie	'Wonder Woman' has biggest opening ever for a female director	Business (Economy)
CNN	2018	Fire and Fury book release	Book fury hits Trump where it hurts most -- his image	Politics
CNN	2018	March for our Lives	Student marchers call Washington's inaction on gun violence unacceptable	US
CNN	2018	Roseanne Barr tweet	Roseanne Barr blames Ambien: 'I'm not a racist, just an idiot'	Media (Entertainment)
CNN	2018	Russia probe	A majority has long not believed Trump on Russia	Politics
CNN	2018	Serena Williams and sexism in tennis	Serena Williams is calling out sexism in tennis. Here's why.	US
CNN	2018	Stormy Daniel's allegations	Stormy Daniels says she was threatened to keep quiet about Trump	Politics
CNN	2018	Trump's tariffs	Trump says US will impose steel and aluminum tariffs	Politics
The Atlantic	2016	China as a superpower	How Geography Explains Donald Trump and Hillary Clinton	Global (World)
The Atlantic	2016	Clinton's emails	There's Simply No Comparison Between Clinton's Flaws and Trump's	Politics
The Atlantic	2016	Gay marriage	What Happened When Religious Freedom and Gay Rights Actually Clashed in Court	Politics
The Atlantic	2016	ISIS	How ISIS Weaponized the Media After Orlando	Global (World)
The Atlantic	2016	Melania's plagiarism	Ten Reasons Why Melania Trump's Speech Will Have a Lasting Impact	Politics
The Atlantic	2016	2016 Presidential elections	The Electoral College Was Meant to Stop Men Like Trump from Being President	Politics
The Atlantic	2016	Trump elected	This Is Who We Are The election of Donald Trump to the presidency reveals the true character of America.	Politics

Appendix E

The Atlantic	2016	Trump's sexual assault	How Trump's Abuse of Women Hid in Plain Sight	Election (Politics)
The Atlantic	2016	Trump's taxes	The \$916 Million Loss Hiding in Trump's Tax Returns	Politics
The Atlantic	2016	Turkey coup	The Logic of the Political Purge	Global (World)
The Atlantic	2017	McCain's Healthcare vote	John McCain's 'No' Vote Sinks Republicans' 'Skinny Repeal' Plan	Politics
The Atlantic	2017	#MeToo	The Perils of #MeToo as a Muslim	Global (World)
The Atlantic	2017	Pence visits the Middle East	Religious Bias Is Distorting American Foreign Policy	Global (World)
The Atlantic	2017	Royal Wedding	Barack Obama Reflects on Leaving the Presidency	Politics
The Atlantic	2017	Russia sanctions	Trump Didn't Tell the Truth About Russia	Politics
The Atlantic	2017	Saturday Night Live	Saturday Night Live Faces Off Against the Trump Presidency	Culture (Entertainment)
The Atlantic	2017	Trump's national security	Trump's News-Conference Performance	News (U.S.)
The Atlantic	2017	Trump's Supreme Court nominees	Trump's Formidable Judicial-Appointment Machine Starts to Sputter	Politics
The Atlantic	2017	United drags man off plane	United's Fiasco and the Cult of Low Prices	Business (Economy)
The Atlantic	2017	Women's March	The Significance of Millions in the Streets	Politics
The Atlantic	2017	Wonder Woman movie	Wonder Woman, Heroine of the Post-Truth Age	Culture (Entertainment)
The Atlantic	2018	Fire and Fury book release	Fire and Fury Is a Strange Kind of Coup for Trump Abroad	Global (World)
The Atlantic	2018	Aretha Franklin's death	Aretha Franklin's Revolution	Culture (Entertainment)
The Atlantic	2018	March for our Lives	A Generation Under Siege	Politics
The Atlantic	2018	2018 Midterm elections	How Senate Republicans Will Use Their New Power	Politics
The Atlantic	2018	Roseanne Barr Ambien tweet	Why ABC Finally Had to Cancel Roseanne	Culture (Entertainment)
The Atlantic	2018	Russia probe	Donald Trump Gave Russia Leverage Over His Presidency	Ideas (Politics)
The Atlantic	2018	Stormy Daniel's allegations	Stormy Daniels's Oh-So-Familiar Story	Culture (U.S.)

Appendix E

The Atlantic	2018	Trump & North Korea	South Korea's Ambassador Has a Message for All the North Korea Skeptics	Global (World)
The Atlantic	2018	Trump's tariffs	Trump's 'Smart' Tariffs Don't Make Economic Sense	Business (Economy)
The Atlantic	2018	Winter Olympics	The Out Olympics	Culture (Sports)
The New York Times	2016	Black Lives Matter	Rudolph Giuliani Lashes Out at Black Lives Matter	Politics
The New York Times	2016	China as a superpower	Washington Should Stop Militarizing the Pacific	Opinion
The New York Times	2016	Gay marriage	Gay Rights Advocates Vow to Fight Utah Bill on Adoptions	US
The New York Times	2016	Global Warming hoax	Donald Trump Could Put Climate Change on Course for 'Danger Zone'	Politics
The New York Times	2016	Gun control	What 130 of the Worst Shootings Say About Guns in America	US
The New York Times	2016	Orlando shooting	Before Orlando Shooting, an Anti-Gay Massacre in New Orleans Was Largely Forgotten	US
The New York Times	2016	2016 Presidential elections	Pulling Democrats Back to 'It's the Economy, Stupid'	Politics
The New York Times	2016	Refugee crisis	Europe's Huddled Masses	Opinion
The New York Times	2016	Trump's taxes	The Trump Campaign's Questionable Tax Return Justification	Business (Economy)
The New York Times	2017	Comey's testimony	James Comey Cometh	Opinion
The New York Times	2017	Covfefe tweet	Less Tweeting, Lawyers Beg. 'Covfefe,' the President Says.	Politics
The New York Times	2017	Las Vegas shooting	Nothing Will Change After the Las Vegas Shooting	Opinion
The New York Times	2017	McCain's Healthcare vote	McCain Provides a Dramatic Finale on Health Care: Thumb Down	On Washington (Health)
The New York Times	2017	Melania moves into White House	Melania Trump Leaves New York Behind for the White House	US
The New York Times	2017	#MeToo	The Patriarchs Are Falling. The Patriarchy Is Stronger Than Ever.	Opinion
The New York Times	2017	Mueller firing	Republican Attacks on Mueller and F.B.I. Open New Rift in G.O.P.	Politics
The New York Times	2017	Pence visits the Middle East	Snubs from Mideast Leaders Complicate Pence's Next Move:	World
The New York Times	2017	Pepsi commercial	Pepsi Pulls Ad Accused of Trivializing Black Lives Matter	Business (Economy)
The New York Times	2017	Russia sanctions	Trump Tries to Shift Focus as First Charges Reportedly Loom in Russia Case	Politics
The New York Times	2017	Saturday Night Live	Donald Trump News Conference Gets the 'S.N.L.' Treatment	Arts (Entertainment)

Appendix E

The New York Times	2017	United drags man off plane	United Airlines Is Not Alone	Opinion
The New York Times	2018	Black Panther movie	The Afrofuturism Behind 'Black Panther'	Opinion
The New York Times	2018	Bush's death	It's Now Donald Trump's America. But George Bush's Stamp Endures.	Politics
The New York Times	2018	Fire and Fury book release	North Korea Praises 'Fire and Fury' Book on Trump Administration	World
The New York Times	2018	Hurricane Florence	Tariffs to Raise Cost of Rebuilding After Hurricane Florence	Economy
The New York Times	2018	Kanye West politics	'I've Been Used': Kanye West Says He's Taking a Break from Politics. We'll See.	Politics
The New York Times	2018	2018 Midterm elections	Weak Spots in Democrats' Strong Midterm Results Point to Challenges in 2020	The Upshot (Politics)
The New York Times	2018	Elon Musk social media	Elon Musk Joins #DeleteFacebook with a Barrage of Tweets	Technology
The New York Times	2018	Nike hires Kaepernick	Nike Returns to Familiar Strategy with Kaepernick Ad Campaign	Sports
The New York Times	2018	Roseanne Barr Ambien tweet	Roseanne Barr's Tweets Didn't Come Out of Nowhere	Arts (Entertainment)
The New York Times	2018	Trump as a "very stable genius"	The Worst and the Dumbest	Opinion
The New York Times	2018	Trump's tariffs	Don't Worry About Trump's Tariffs	Opinion
The New York Times	2018	World Cup	FIFA Prepares for a Post-Trump World	Sports
USA Today	2016	Anthony Weiner sexts	FBI email letter raises more questions than it answers	News (U.S.)
USA Today	2016	Refugee crisis	Traffickers use social media to lure migrants - Despite risk of death, refugees are willing to pay steep prices for passage from Turkey to Greece	World
USA Today	2016	Trump elected	Trump's surprise victory leaves a nation unnerved	Opinion
USA Today	2016	Trump's sexual assault	Women say they were groped by Trump	Investigation (U.S.)
USA Today	2016	Zika virus	Zika vaccine waits on Congress - Virus is devastating to the small number of people getting infected during pregnancy	Health
USA Today	2017	Betsy DeVos	What you need to know about Betsy DeVos	News (U.S.)
USA Today	2017	Mueller firing	Smearing Robert Mueller If Trump is innocent, why the effort to discredit the special counsel's findings?: Our view	Editorial (Politics)

Appendix E

USA Today	2017	Russia sanctions	Nearly half of Americans think Trump committed a crime related to Russian election meddling	News poll (U.S.)
USA Today	2017	Saturday Night Live	Alec Baldwin's Donald Trump showers 'SNL' with urination jokes	Entertainment
USA Today	2017	Women's March	At 2.6 million strong, Women's Marches crush expectations	Politics
USA Today	2017	Wonder Woman movie	Spoilers: 5 wonderfully feminist moments in 'Wonder Woman'	Entertainment
USA Today	2018	Mira Ricardel fired	Who is Mira Ricardel? First Lady Melania Trump wants the national security official fired	Politics
USA Today	2018	Nike hires Kaepernick	Just watch it: Nike's Colin Kaepernick TV ad is inspirational, not controversial	Sports
USA Today	2018	Russia probe	Russia inquiry: House Democrats aim to unmask Trump Jr.'s blocked call	Politics
USA Today	2018	Stormy Daniel's allegations	Trump seeks arbitration in Stormy Daniels case	Politics
NPR	2016	Clinton's emails	Did FBI Director James Comey's Email Announcement Break the Law?	Politics
NPR	2016	Gay marriage	In Italy, Protesters Push for Rights for Same-Sex Couples	International
NPR	2016	Global Warming hoax	Has Trump's Stumping Already Frayed Transatlantic Ties?	Environment
NPR	2016	Orlando shooting	What Queer Latinos Are Saying About the Orlando Shooting	Codeswitch (U.S.)
NPR	2016	Scalia death	Scalia And Leonard Nimoy: Justice's Death Spurs Conspiracy Theories	Must reads (U.S.)
NPR	2016	Trump elected	Donald Trump Clinches the Presidency in Major Upset	Politics
NPR	2016	Trump's taxes	How Donald Trump Would Be Able to Not Pay Income Taxes For 18 Years	Politics
NPR	2016	Zika virus	Congress Returns: Much Ado About Little to Do	Politics
NPR	2017	Betsy DeVos	Betsy DeVos Confirmed as Education Secretary	Higher Ed (Politics)
NPR	2017	Mueller firing	3 Ways Trump or His Allies Might Try to Disrupt the Mueller Russia Probe	National Security (U.S.)
NPR	2018	Aretha Franklin's death	Aretha Franklin Was America's Truest Voice	Music (Entertainment)
NPR	2018	Hurricane Florence	Tropical Storm Florence, 'An Uninvited Brute,' Brings Floods, And At Least 5 Deaths	National (Environment)
NPR	2018	2018 Midterm elections	A Boatload of Ballots: Midterm Voter Turnout Hit 50-Year High	Elections (Politics)
NPR	2018	Trump as a "very stable genius"	'A Very Stable Genius': Trump Responds to Renewed Criticism of His Mental State	National (U.S.)
NPR	2018	Trump & North Korea	North Korea Reportedly Expanding Ballistic Missile Production Facility	International (World)

Appendix E

AP	2016	Penn Chapo interview	Drug lord was interested in Mexican actress, not Sean Penn	Entertainment
AP	2017	Women's March	Over 1 million join anti-Trump women's marches worldwide	Politics
AP	2017	Wonder Woman movie	Review: 'Wonder Woman' film and star live up to the name	Review (Entertainment)
AP	2018	Child immigrants	Immigrant kids seen held in fenced cages at border facility	US News
AP	2018	Kanye West politics	Kanye West tweets he's dropping politics, writes big check	Entertainment
AP	2018	Taylor Swift politics	Taylor Swift breaks political silence, backs Tennessee Dems	Entertainment
