In Plain Sight: Media Bias Through the Lens of Factual Reporting (Supplementary Material)

Lisa Fan[1,*] **Marshall White**[1,*] **Eva Sharma**[1] **Ruisi Su**[1]
Prafulla Kumar Choubey[2] **Ruihong Huang**[2] **Lu Wang**[1]
[1] Khoury College of Computer Sciences, Northeastern University, Boston, MA 02115
[2] Department of Computer Science and Engineering, Texas A&M University
{fan.lis, white.mars, sharma.ev, su.ruis}@husky.neu.edu
{prafulla.choubey, huangrh}@tamu.edu, luwang@ccs.neu.edu

A Sample Annotations

On the right, several sample annotations from the BASIL dataset illustrate some aspects of our annotation schema and highlight characteristics of informational bias.

Indirect Bias. Though not as prevalent as bias spans with direct aim, indirect aim is nevertheless important to study because readers may find it more difficult to detect bias consciously when it does not directly implicate the main entity. Indirect bias can be aimed through an intermediary ally or opponent, or may be based on contextual information. In each case, the sentiment towards the intermediary entity alters sentiment toward the main target entity.

Figure 1a shows an example of indirect bias where Donald Trump is negatively targeted via the negative framing of an ally, Donald Trump Jr. Readers are required to know the relationship between the two men in order to notice the bias, and the information itself would be irrelevant to the article were it not for their relationship.

The span from HPO in Figure 1b shows an indirect bias span where contextual information unconnected to the rest of the article reflects negatively on Trump without mentioning him in the text. It requires several leaps in logical thinking: children and families seeking asylum are sympathetic :: turning them away is bad :: Trump wants a border wall :: Trump is framed negatively. This type of informational bias is difficult to detect algorithmically as there is no mention of Trump, the target main entity.

Informational Bias Strategies. Inspecting the informational bias spans in our dataset reveals several trends and strategies that journalists tend to

[*]Equal contribution. Lisa Fan focused on annotation schema design and writing, Marshall White focused on data collection and statistical analysis.

Main Event: Trump reverses decision to allow import of elephant trophies

Main Entity: Donald Trump

NYT: On social media, photos were being shared of Mr. Trump's two elder sons hunting on safari in Zimbabwe, [*including one photo that showed Donald Trump Jr. with a severed elephant tail in one hand and a knife in the other.*]$_{\text{Trump}}$

(a) Indirect negative informational bias against Donald Trump, using the intermediary entity Donald Trump Jr.

Main Event: Trump declares national emergency over border wall

Main Entity: Donald Trump

HPO: [*Since 2014, a high proportion of those crossing have been Central American children and families seeking to make humanitarian claims such as asylum.*]$_{\text{Trump}}$

FOX: President Trump said Friday he is declaring a national emergency on the southern border ... [*despite his criticisms of former President Barack Obama for using executive action.*]$_{\text{Trump}}$

NYT: Mr. Trump's announcement came during a free-wheeling, 50-minute appearance ... [*The president again suggested that he should win the Nobel Peace Prize, and he reviewed which conservative commentators had been supportive of him, while dismissing Ann Coulter, who has not.*]$_{\text{Trump}}$

(b) Example annotations showing negative informational bias from all three media sources for one article triplet.

Main Event: Raul Labrador challenges Kevin McCarthy for House majority leadership

Main Entities: Raul Labrador, Kevin McCarthy

HPO: [*Labrador is an ambitious, sometimes savvy politician.*]$_{\text{Labrador}}$ He is in Idaho this weekend chairing the state GOP convention.

(c) Example annotation of positive informational bias.

Figure 1: Excerpts showing different types of informational bias, annotated in italics. The target of the negative bias is noted at the end of each span. Underlined entities are intermediary targets in indirect bias spans.

use. The examples from FOX and NYT in Figure 1b show the strategy where objective but tangential information frames the target in a negative light given the context of the article. The example from FOX uses nonessential background information to imply Trump is hypocritical, and the NYT example includes a detail peripheral to the main event that portrays Trump as rambling.

Figure 1c is an example of subtle informational bias where the author's opinion masquerades as fact. The writing is in a neutral tone and appears objective, but it is actually the author's perception of the situation and uncovers their bias towards the topic. The span is categorized as informational bias rather than lexical because there is no way to rephrase or remove parts of the sentence without changing the overall meaning. This span is also an example of the rarer positive bias span.

B Data Collection

BASIL contains 100 triplets of articles, each with 3 articles about the same main event from the New York Times (NYT), Fox News (FOX), and the Huffington Post (HPO). According to Budak et al. (2016), FOX is considered strongly right leaning, NYT slightly left leaning, and HPO strongly left leaning. As an initial annotation set, 16 triplets of highly visible, polarizing events were directly selected from the media source websites by our annotators.

The remaining triplets were aligned algorithmically from the Common Crawl corpus.[1] Articles with less than 200 words or more than 1,000 words were filtered out, and only political, non-editorial articles published within 3 days of each other were considered. Article similarity was calculated using the cosine similarity of the TF-IDF vectors of each article's title combined with its first 5 sentences. For each FOX article, the most similar NYT article was found, then the most similar HPO article was found using this pair. An annotator manually selected the final triplets from this list of automatically aligned triplets.

Main event and entities were manually annotated for each article by one annotator. Articles in a triplet share the same main event, which the annotator produced after reading the leads of the three articles. Main entities sometimes differ across the triplet, as stories about the same event can emphasize different characters, but at least one main entity is consistent across each triplet. A single article contains an average of 2.04 main entities and at most five main entities.

[1]http://commoncrawl.org

	Exact Matching			Lenient Matching		
	Prec.	Rec.	F1	Prec.	Rec.	F1
Lexical Bias						
A + B	11.04	14.17	12.41	12.34	15.83	13.87
A + C	8.57	9.76	9.13	11.43	13.01	12.17
B + C	–	–	–	15.38	15.38	15.38
Informational Bias						
A + B	19.90	17.22	18.46	39.80	34.33	36.92
A + C	19.47	22.05	20.68	34.40	38.97	36.54
B + C	15.29	10.83	12.68	32.94	23.33	27.32

Table 1: Inter-annotator span agreement for lexical and informational bias. Dashes indicate that there were no exact matching lexical text spans between annotators B and C.

		Dimensions (Cohen's κ / % Agr.)		
	# Res.	Target	Polarity	Aim
A + B	123	0.93 / 93.7	0.84 / 96.3	0.12 / 93.7
A + C	138	0.88 / 89.5	0.75 / 95.0	0.54 / 89.9
B + C	39	0.96 / 96.9	0.92 / 96.9	– / 96.9

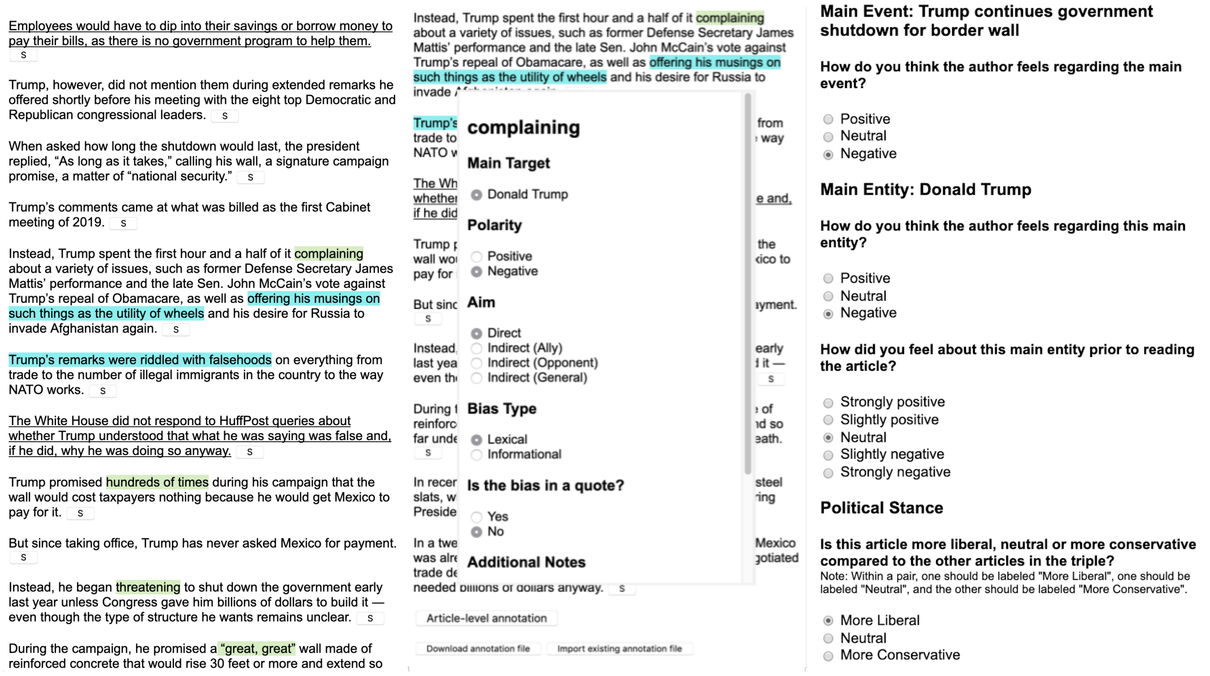
Table 2: Number of articles resolved by each annotator pairing, along with Cohen's κ and percent agreement for IAA on auxiliary dimensions for overlapping spans.

During the annotation process, the order of articles is randomized within each triplet and annotators are not aware of the media source of the article. The entire dataset was annotated by three unique annotators.

C Inter-annotator Agreement

Our study of inter-annotator agreement consists of two parts: the agreement of the text spans selected and the agreement on the dimensions within each annotation span. To find text span agreement, a similar method to Toprak et al. (2010) is used in which precision, recall, and F1 are calculated between two annotators using the agreement metric from Wiebe et al. (2005), treating one annotator's spans as the gold standard and the other annotator's spans as the system. Results are calculated for *exact matching*, where the text spans must overlap exactly to be considered correct, and *lenient matching*, where text spans with any overlaps are considered correct (Somasundaran et al., 2008).

Table 1 shows that span agreement is higher for spans of informational bias than for spans of lexical bias due to the sparsity of lexical bias in our



(a) Tool with loaded annotations. Informational bias spans are shown in blue, lexical bias spans are shown in green.

(b) Tool with sentence-level annotation form.

(c) Document-level annotation form.

Figure 2: Our Javascript annotation tool at various steps.

dataset (see Table 1 in the main paper).

Dimension agreement is reported in Table 2 only for lenient matching spans, as the results are not significantly different from that of exact matching spans. Cohen's κ is used to measure attribute agreement for target, polarity, and aim, and we find high levels of agreement for both polarity and target. Because of the metric's sensitivity to class imbalance, Cohen's κ is impractical for measuring the agreement on aim for one annotator pairing (B + C), which had fewer article triplets to resolve and nearly all overlapping lexical annotations were marked as *direct* (31 / 32 spans). To account for this imbalance, the percent agreement for all attributes is also included in Table 2.

D Javascript Annotation Tool

A Javascript based tool[2] was developed to annotate our dataset. Annotations created in the tool can be downloaded in JSON format and analyzed or imported at a later date. Users can highlight spans of text or select an entire sentence, then answer dimensional questions (see Figure 2b). Users can also answer document-level questions (see Figure 2c). Figure 2a shows the tool after annotations have been made, where blue spans are informational bias and green spans are lexical bias. In order to alleviate eye strain, annotations of the entire sentence are shown underlined rather than highlighted.

[2] `https://github.com/marshallwhiteorg/emnlp19-media-bias`

References

Ceren Budak, Sharad Goel, and Justin M Rao. 2016. Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, 80(S1):250–271.

Swapna Somasundaran, Josef Ruppenhofer, and Janyce Wiebe. 2008. Discourse level opinion relations: An annotation study. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, pages 129–137.

Cigdem Toprak, Niklas Jakob, and Iryna Gurevych. 2010. Sentence and expression level annotation of opinions in user-generated discourse. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 575–584. Association for Computational Linguistics.

Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2-3):165–210.