



ANNOTATION MANUAL EXPERT DATASET

Media Bias Project

This version: November 6, 2020

1 Introduction

The way in which journalists report on newsworthy events, or choose not to, can subtly influence consumers in their perception of political issues. While objectivity is a principle of journalistic standard and ethics shared across most news-reporting organizations, news are not always presented based on objective criteria. Deliberately or not, ideological views and prejudices might interfere in the work of professional journalists. This phenomenon is commonly referred to as *media bias* and finding its different instances is the research objective of this study. Media bias manifests itself in news coverage in different ways. The focus of this study is on the language conveying bias. Instances of text in which events are described by using a particular word choice or framing expose readers to non-neutral news reporting.

In figure 1 one can easily see how bias in media can lead the reader to a conclusion different than what a neutral reporting might have caused. A possible neutral point of view would be to name the disease 'Coronavirus' instead of 'Wuhan Virus'. Both terms describe the same concept but frame it differently. The nuances of word choice and labelling cause different sentiments and potentially lead readers to different conclusions. Here, news consumers are more likely to follow the narrative that an ethnic group of a particular region is responsible for the outbreak of the disease by linking the word 'Wuhan' with 'Virus'. This, in return, increases the risk of stigmatization.

Trump Defends Calling Wuhan Virus 'Chinese Virus'

Figure 1: Example of Media Bias in a Headline of *The Federalist*

In this survey we will show you sentences selected from news articles reporting on a variety of issues published in a broad spectrum of news outlets. We ask you to annotate instances of biased language according to our detailed instructions presented in this document. By enriching the corpus with linguistic information on media bias you will contribute to the training of machine learning algorithms to perform the same task in an automated fashion.

2 What needs to be annotated

Media bias is a linguistically complex phenomenon and moreover sometimes difficult to detect because of its subtleness. Generally, the more difficult a problem is the more complex theory and annotation guidelines need to be. Thus, we will provide you with linguistic cues of when bias occurs and elaborate them with examples. There are two linguistic realizations of media bias. One type of bias can be labelled as **framing bias** and the other type as **epistemological bias**. In the following elaboration on the different manifestations of biased language, examples are elaborated in *italics*. When available, possible neutral formulations are provided as well.

2.1 Framing bias

A framing bias alters the reader's opinion by describing an event only from one particular point of view or *frame*. Linguistically, this is realized by the use of word choices or subjective labelling of concepts.

- **Labelling** is a technique that journalists can use when referring to a semantic concept such as an actor, location or event with subjective intensifiers. The usage of adjectives or adverbs that convey a strong opinion in that context introduce bias.

- (1) *In the months that followed, all independent polling has found the American mainstream not only blames Trump and his party for the shutdown, but also does not want to spend billions of taxpayer dollars on an ineffective and unnecessary border wall.*
- (2) *More than 100 gang members from El Salvador are among the massive wave of migrants that's poured across the border in recent months, [...], beleaguered border officials said.*
- (3) *In July 2019, Rep. Alexandria Ocasio-Cortez, D-N.Y., and other Democrats visited a similar processing center in Clint, Texas, and made explosive allegations about the conditions there.*
- (4) *Rather than reducing federal subsidies to universities and placing the leftist institutions at the mercy of the free market, Obama wants to increase the amount students can borrow with interest from the government.*

- The **word choice** when referring to a semantic concept can also induce bias. One-sided terms or ideologically-driven depictions of concepts may alter the reader's point of view about the issue.

- (5)
 - a. *The explosion of the Hispanic population has long-term job prospect consequences as well: Both legal and illegal aliens will occupy 75 percent of new American jobs in as little as five years.*
 - b. *The explosion of the Hispanic population has long-term job prospect consequences as well: Both legal and illegal immigrants will occupy 75 percent of new American jobs in as little as five years.*
- (6)
 - a. *Abortionists, who prey on desperate people, should no longer receive your hard-earned tax dollars.*
 - b. *Abortion-right activists, who prey on desperate people, should no longer receive your hard-earned tax dollars.*

2.2 Epistemological bias

This class of biased language is often harder to detect than framing bias because of its subtle nature. It encompasses language constructs that aim at manipulating the believability of a statement. The typical mechanisms are either presupposing the truth or the falsehood of a statement. As a result, statements seem to be either more believable or less.

- (7)
 - a. *He realized that the oppression of black people was more of a result of economic exploitation than anything innately racist.*

- b. *His opinion was that the oppression of black people was more of a result of economic exploitation than anything innately racist.*
- (8) a. *The Republican president assumed he was helping the industry at the expense of the environment - a trade-off Trump was happy to make since he rejects climate science anyway.*
- b. *The Republican president aimed at helping the industry at the expense of the environment - a trade-off Trump was happy to make since he rejects climate science anyway.*
- (9) a. *Former ESPN personality Jemele Hill attacked the NFL for its lack of black coaches, and blasted the Rooney Rule as an insufficient solution to the problem.*
- b. *Former ESPN personality Jemele Hill attacked the NFL for its lack of black coaches, and accused the Rooney Rule as an insufficient solution to the problem.*
- (10) a. *So while there may be a humanitarian crisis driving more vulnerable people to seek asylum in the United States, there is no security crisis.*
- b. *So while there is a humanitarian crisis driving more vulnerable people to seek asylum in the United States, there is no security crisis.*
- (11) In general, verbs such as *say* and *state* are neutral, while words such as *claim* or *boast* lead the reader to examine the statement more critically.

3 What should not be annotated

There are some closely related topics to media bias such as **fake news** or **opinionated reporting**. The aim of this study is not to expose the factual correctness of sentences. Instead, we focus on how concepts are portrayed linguistically, be it in a favourable or negative fashion. Even if you encounter statements containing untrue information, they do not necessarily contain biased language as following example demonstrates: *Ingraham began the exchange by noting American graduates' salaries have been suppressed by the flood of foreign graduates*. While the notion that graduates' salaries are suppressed by foreign graduates cannot be considered a fact, it is not an instance of media bias.

Furthermore, do not annotate words as biased merely because they are a concept of a controversial topic. For instance, *abortion* is not a biased word, while the term *abortionist* from example (6) is biased. This one-sided depiction for abortion-right activists leads the reader to believe that those activists are solely motivated by the act of abortion itself and not by the freedom for women to chose. In the following example the concept of *neo-Nazi propaganda* can be considered a controversial topic, still, the sentence cannot be considered biased: *A review of his posts in online message boards revealed he wanted to plant neo-Nazi propaganda inside Nevada middle and high schools*.

4 How to annotate

During the survey, you will be shown both biased and impartial sentences. Each sentence is associated with 3 following tasks:

1. We would like you to highlight the words that convey strong opinion and are most likely to cause bias as explained in section 2. You also can highlight short phrases, e.g., "New York".
2. Label the sentence as biased or non-biased.
3. Label the sentence as factual or opinionated.

While annotating, it is important to follow the following guidelines:

- Keep the bias annotation as short as possible. Please select multiple words together only when they keep the meaning that they have in that particular context, only when they are used together. Otherwise, please select the words separately.

- Leave your personal opinion about the sentence aside and label all instances of biased language irrespective of its political context.