Ad Fontes Media's Multi-Analyst Content Analysis White Paper

Revised September 2021

Author: Vanessa Otero

Download PDF Here: Multi-Analyst Ratings Project White Paper

Overview

This white paper describes, in detail, the process by which Ad Fontes Media produces the various iterations of the Media Bias Chart. Ad Fontes Media's founder, Vanessa Otero, created the first Media Bias Chart in October of 2016. She initially created it as a hobby, for the purpose of creating a visual tool to discuss the news with friends and family, so she created the taxonomy and analyzed the initial set of news sources herself. However, because it grew to be

extremely popular, she sought out to improve the methodology, make the process more data-driven, and mitigate bias (hers and that of any new analysts). To do so, she recruited teams of politically diverse analysts and trained them in the methodology. Over time, this process evolved into Ad Fontes Media's current method of multi-analyst content analysis ratings.

Ad Fontes finished our first extensive multi-analyst content ratings research project in June 2019. From June 2019 to August 2020, a group of nine analysts from that initial project continued to rate several dozen articles per month to add new sources and update previously existing ones. From August-October 2020, we conducted a second large multi-analyst content ratings project to rate over 2000 articles and 100 new news sources with some existing analysts and over 30 new analysts. From October 2020 to present, we have used our team of nearly 40 paid, trained analysts to continuously rate new sources and update ratings on existing sources. Our current data set includes nearly 20,000 multi-analyst ratings of articles, and episodes from over 1600 Web, podcast, and TV shows, and growing.

All the scores we collect in our analysis rely on an assumption that the taxonomy of the Media Bias Chart is a valid way of rating news sources on the dimensions of reliability and political bias. This taxonomy was created by [Vanessa Otero](), the founder of Ad Fontes Media, who has a B.A. in English from UCLA and a J.D. from the University of Denver Sturm College of Law. Most of her educational and career background focused on analytical reading and writing. Aspects of the systems and methods described herein are patent pending.

Since the release of the original Media Bias Chart in 2016, millions of observers have found the system of classification to be useful, regardless of whether they personally agree with where news sources are placed on the Chart. Social scientists, data scientists, statisticians, and news organizations may find areas of our work to be improved, and we welcome criticisms and suggestions, recognizing there is always sharpening that may be done, and that the news landscape itself is always shifting. Some observers take issue with aspects of the taxonomy itself; for example, some object that a left-right spectrum doesn't capture the full extent of political positions. Such points may be (and likely will be) debated forever, but based on the proven utility of this taxonomy for so many,

and with increasing demand for the underlying data, we continue to use it while improving and refining it along the way.

II. Core Taxonomy

Otero also created the original content analysis methodology. However, the methodology has evolved over time and with input from many thoughtful commentators and experts, including Ad Fontes Media Advisor and long-time journalist and journalism professor [Wally Dean](). Prof. Dean co-authored *We Interrupt this Newscast*, a book detailing one of the largest content analysis studies ever done--a study of local news broadcasts from 1998 to 2001.

We currently provide in-depth discussions of our taxonomy and methodology in various public-facing webinars. For example, we teach ongoing news analysis training that is available for members of our site and free teacher training which we conduct from time to time. Several videos from these trainings are linked throughout this paper to provide additional detail into our process.

A. Framework

Our taxonomy continues to be a two-dimensional framework for rating the reliability and bias of content, shows, and sources.

The horizontal axis (political bias, left to right) is divided into seven categories, three of which represent the spectrum on the left, three of which represent the spectrum on the right, and one in the middle. Each category spans 12 units of rating, so the total numerical scale goes from -42 on the left to +42 on the right. These values are somewhat arbitrary, though there are some good reasons for them, including that they 1) allow for at least seven categories of bias, 2) allow for more nuanced distinction between degrees of bias within a category (allowing analysts to categorize something as *just a bit* more biased than something else), and 3) they correspond well to visual displays on a computer screen or a poster.

Bias scores are on a scale of -42 to + 42, with higher negative scores leaning more to the left, higher positive scores leaning more to the right, and scores closer to zero being either centrist, minimally biased, and/or balanced.

(It should be noted that centrism, neutrality, and balance are not the same thing from our point of view. For example, an article may score near 0 because it contains an even-handed debate between multiple similarly informed and authoritative sources.)

The vertical axis (overall reliability, top to bottom), is divided into eight categories, each spanning eight rating units, for a total numerical scale of 0 to 64. Again, these are somewhat arbitrary, but the eight categories provide sufficient levels of classification of the types of news sources we are rating and sufficient distinction within the categories. Reliability scores are on a scale of 0-64, with source reliability being higher as scores go up.

Overall source ratings are composite weighted ratings of the individual article and show scores. There are many specific factors our analysts take into account when considering the reliability and bias of an article, episode, or other content. The main ones for *Reliability* are defined metrics we call "Expression," "Veracity," and "Headline/Graphic," and the main ones for *Bias* are ones we call "Political Position," "Language," and "Comparison."

Observers may argue that numerical scales for bias and reliability should be more or less granular, but our ratings are based on the scale described above. Therefore, an underlying premise of the scoring is that this scale is useful for conveying its meaning to its observers.

B. Definitions

The horizontal (or "bias") categories are defined by the policy positions of current US elected officials. For more on why, [see this methodology video](). This video also discusses how the US left-right spectrum shifts over time (which relates to a concept known as the Overton Window) and how that affects our rating of contemporary media content. There are three important definitions we use for defining areas of the horizontal axis, which are as follows:

1. The line between "Most Extreme Left/Right and Hyper-Partisan Left/Right" is defined by the policy positions of the most extreme *elected* officials significantly relevant to the scope of the issue being considered. Any

position in a news source that is more extreme than what that official advocates falls in the "Most Extreme" category.

2. Most of the categories of "Hyper-partisan Left/Right and Skews Left/Right" are defined by the current positions of Democratic and Republican party politicians, with the midpoint position in each party (the policy positions held by the leaders of each party, respectively) represented by the line between "skews left/right" and "hyper-partisan left/right."

3. The "middle" position on the horizontal axis represents narratives with no discernable political position taken, a relatively complete and balanced survey of key competing positions on a given issue, as well as narratives representing a "centrist" political perspective, which is itself a political bias. An article, episode, or source placing near the midpoint on the horizontal axis may land there for any of these reasons; thus the position does not necessarily represent "neutrality." Nor is the midpoint on the horizontal axis intended to imply that the position is best or most valid.

The vertical (or "reliability") axis represents a continuum between fact and falsehood, as follows:

1. Analysts score content they deem to be primarily fact reporting between 48 and 64, with the highest scores reserved for encouraging the hard (and socially essential) work of original fact reporting that is subsequently corroborated by additional sources. However, fact-based content that is widely corroborated has the highest likelihood of veracity, so fact reporting generally falls between 48 and 56 on the vertical axis.

2. Content that includes analysis scores between 32 and 48, with the higher scores in this range reserved for analysis that is supported by well-argued fact reporting. In terms of "reliability," the taxonomy places opinion (24-32) below analysis. However, as with analysis, opinion that is well-argued and defended based with facts also scores higher within the category.

3. Content scoring below 24 generally has a reliability problem. When it scores between 16 and 24, very likely an important part of the story was omitted. It is likely (and literally) a "partial" story representing - at least in that sense - an "unfair" attempt at persuasion. Content scoring below 16 has been determined by our analysts to be misleading or downright false, at least based on the best evidence presented to date.

It should be noted that our taxonomy and methodology constitute a rubric used to *describe* content on both the horizontal ("bias") and vertical ("reliability") axes. Bias scores are *descriptive* in relation to the current politics of the country as a whole. They are not intended to rate the moral quality of a position; nor are they measured against a timeless or universal norm. Reliability scores are similarly *descriptive*, though veracity is one of the metrics considered on the vertical axis. While veracity is part of what we consider when rating the reliability of content, there is a categorical difference between the "rightness" or "wrongness" of content expressing an opinion and the "truth" or "falsehood" of content stated as fact. Moreover, there are limits to human knowledge, and our methodology considers "likelihood of veracity" to be more accurate than a"true/false" toggle when considering the accuracy of content presented as fact.

The overall source rating is a result of a weighted average, algorithmic translation of article raw scores. Low-quality and highly-biased content weight the overall source down and outward. The exact weighting algorithm is not included here because it is proprietary, but generally, lower reliability and more biased content is weighted more heavily. Aspects of what is disclosed here are patent pending.

III. Analysts

Prior to October 2020, all analysts were volunteers receiving perks and/or small stipends. Since October 2020, Ad Fontes Media has contracted and hired a team of analysts to rate news content on an ongoing basis. At the time of this 2021 revision, we have 33 active professional analysts.

Currently, our analyst application process requires the following:

- Submission of a professional resume, CV, or similar written summary of qualifications.
- Completion of an online application enabling further assessment of qualifications.
- Completion of a political viewpoint assessment used to build an analyst team that can provide equal input from those leaning politically to the left, those leaning to the right, and those with centrist leanings. A self-reported classification of their political leanings. Each analyst submitted a spreadsheet about their political views overall and per listed political topic.

The "political position assessment" can be viewed on our site on the [analyst application](#) page.

Submission of basic demographic information is optional but helpful in maintaining an analyst team that is relatively representative of the country as a whole.

A. *Education and Qualifications*

Our current qualification expectations for new applicants are as follows:

Minimum Qualifications

- Has completed minimum high school + 2 years (60 hours) of undergraduate work
- Lives in the United States, and is politically/civically engaged
- Has a personal computer and reliable internet, and is able to troubleshoot basic technical issues
- Is familiar with a range of news sources
- Is familiar with party platforms and government systems in the US
- Is willing to divulge political leanings internally as required by our approach to analysis
- Demonstrates excellent reading comprehension skills
- Demonstrates excellent analytical skills
- Demonstrates ability to engage in sometimes difficult conversations, including on sensitive issues
- Demonstrates ability to see issues from multiple perspectives while also respectfully expressing a dissenting perspective when applicable

Strong Qualifications

- Demonstrates a passionate interest in news media and contemporary US politics
- Demonstrates a desire to make a positive difference.
- Has achieved an advanced degree, or a highly relevant undergraduate degree, in Media, Journalism, Political Science,

Linguistics, History, Sociology, Philosophy, or other field requiring strong skills in analyzing information content.
- Helps contribute to the diversity of our team
- Helps contribute to the range of special subject expertise within our team
- Helps contribute to the range of skills within our team
- Demonstrates familiarity with identifying bias and reliability in news sources
- Demonstrates interest in Ad Fontes Media and our mission

The extent to which applicants demonstrate the qualifications above is assessed by a politically balanced team of application reviewers using a shared rubric to identify the most qualified applicants.

All of our current analysts hold at least a bachelor's degree, and most have completed at least one graduate degree program. Approximately one third have completed a doctoral degree program, or are current doctoral students.

While education is an important qualification, a number of other factors are considered as well, particularly familiarity with US politics and the ability to engage in rigorous critical reflection on written and spoken content. Analysts come from a wide range of professional backgrounds - including federal service, law, and management - the backgrounds represented most within the team are media and communications, education, and research.

B. Political Leanings

Because analysts use a shared methodology (described below), the descriptive placement of each piece of content on the Media Bias Chart is generally quite close regardless of the analyst's political bias. However, since we moved to a multi-analyst approach in 2019, each piece of content we rate has been rated by an equal number of analysts who identify as left-leaning, center-leaning, and right-leaning politically.

To arrive at the classification of the analysts, we lean heavily on their own sense of political identity, along with a self-assessment currently using the following categories:

1. Abortion-related policy
2. Affirmative Action Action & Reparations
3. Campaign Finance
4. Climate-related policy
5. Criminal Justice Reform
6. Defense/Military Budget
7. Subsidized Food & Housing
8. Gun-related policy
9. Higher Education Policy
10. Immigration
11. International Affairs
12. K-12 Education Policy
13. LGBTQ Related Policies
14. Marijuana Policy
15. Private/Public Healthcare Funding
16. Regulation of corporations
17. Social security
18. Tax related policies

For each of the issues above, we request that each analyst identify their perspective as:

- "Decidedly to the left"
- "Moderately to the left"
- "Centrist or undecided"
- "Moderately to the right"
- "Decidedly to the right"

For each issue in which the analyst identifies their perspective as "decidedly to the left," they score "-2," for each "moderately to the left," they score "-1," for each "centrist or undecided," they score "0," and so on.

Analysts scoring more than 10 points from 0 are initially categorized as left-leaning or right-leaning. Analysts scoring fewer than 4 points from 0 are initially categorized as centrist. Analysts falling between 4 and 10 points from 0

are considered on a case-by-case basis, with the analyst's political identity being considered most heavily.

While individuals' political outlooks are generally quite complex, and are often varied across issues, the analysts are generally able to identify their own perspective on these issues quickly using the framework above. To do so assumes a level of familiarity with US politics, which is assessed during the application process. When combined with the practice of having each piece of content analyzed by an equal number of left-, center-, and right-leaning analysts, we have found that this system of classifying analysts yields a diversity of perspectives represented when analyzing content, though the horizontal axis analysis is politically descriptive and not intended to judge the moral value of political positions.

C. Training

Before joining the analyst team, each analyst trainee reads an article overviewing each step within Ad Fontes Media's eight-step core analysis methodology. The eight steps are made up of Veracity, Expression, Headline/Graphics, and Overall Reliability for the "Reliability" (vertical) metric, along with Political Position, Language, Comparison, and Overall Bias for the "Bias" (horizontal) metric. For each of these eight metrics, analysts also attend a 60-minutes presentation on the same topic.

As they complete the assignments above, trainees also practice rating content using the methodology as a rubric. They also observe live analysis shifts where experienced analysts consider content together. Trainee ratings are observed; however, during this time, trainee scores are not included in the data used in our overall source and content ratings.

Upon successful completion of 20-25 hours of training described above, and once significant outlier scores are rare, trainees enter a probationary period where they score articles along with two experienced analysts. At this point, trainees scores are included in source and content ratings, and any outlier scores are managed as described below.

All analysts attend periodic additional training, which may include occasional fine-tuning to the methodology and awareness of the shifting meaning of categories such as "left" and "right" when applied to specific issues over time.

IV. Process of Analysis

   A. Content Selection

To date, we have fully rated nearly 1000 sources and shows, including Web/print, podcast, and television. Members of our team use reach data, source lists, and user requests in order to select sources to be rated. While all sources gain additional article and episode scores over time, some sources have many more data points than the minimum, no source or show is considered to be "fully rated" until our team has rated a minimum of 15 articles for web/print content or 3 complete episodes of podcast or television content. This being the case, our team has scored over 20,000 articles and episodes to date in order to arrive at approximately 1000 sources fully rated, and analysis is ongoing with several shifts of live analysis running daily.

Upon selecting a source to be rated, sample articles or episodes are selected for analysis, and these samples are used to arrive at an overall source score.

Articles are currently selected manually based on their "prominence," as determined by page placement, size of print headline, or when available, based on reach. In the future, we would like to consider reach more systematically when selecting sample content to be rated. Prominence functions partly as a proxy for reach when necessary currently, though prominence is also an important part of our methodology because many publishers feature highly opinionated or biased articles to drive engagement, even if most content they publish is more fact-based and neutral. Public perceptions of bias of large publishers are often driven by the extensive reach of lower-reliability, highly biased content.

For TV networks, content is similarly selected based on reach and its prominence in terms of when it is scheduled to air. For podcasts and TV shows, sample episodes are selected based on representativeness of the show overall.

For some sources, current ratings are based on a small sample size from each source. We believe these sample articles and shows are representative of their respective sources, but these rankings will certainly get more accurate as we rate more articles over time.

We rate all types of articles, including those labeled analysis or opinion by the news source. Not all news sources label their opinion content as such, so regardless of how it is labeled by the news source, we make our own methodology determinations on whether to classify articles as analysis or opinion on the appropriate places on the chart.

The content rating period for each rated news source is performed over multiple weeks in order to capture sample articles over several news cycles. Sources that have appeared on our Media Bias Chart for longer have articles over much longer periods of time.

Often, our sample sets of articles and shows are pulled from sites on the same day, meaning that they were from the same news cycle. Doing so allows analysts to incorporate evaluations of bias by omission and bias by topic selection.

We update all sources periodically by adding new articles. Because we have so many news sources, and because the most popular sources are important to the public, we generally update the most popular sources the most frequently and less popular sources less frequently. For example, we update a tier of the top 15 sources with about 15 new articles each month, and the next tier of 15 sources with about 7 articles per month. The top 200 get updated with about 5 articles per quarter and the next 200 about 5 articles per 6 months. We strive to balance rating new sources and updating existing ones.

B. Analysis

Each individual article and episode is rated by at least three human analysts with balanced right, left, and center self-reported political viewpoints. That is, at least one person who has rated the article self-identifies as being right-leaning, one as center-, and one as left-leaning.

The main principle of Ad Fontes (which means "to the source" in Latin) is that we analyze content. We look as closely as possible at individual articles, shows, and stories, and analyze what we are looking at: pictures, headlines, and most importantly, sentences and words.

In 2020, we began to rate most content live because the live process requires each analyst to justify their score when needed, aids in exposing analysts to multiple perspectives, and allows analysts to point out features that may have been missed by a single person in the group.

Articles and episodes are rated in three-person live panels conducted in shifts over Zoom, with each pod containing a left-leaning analyst, a center-leaning analyst, and a right-leaning analyst. Analysts first read each article and rate it on their own, and then they immediately compare scores. If there are notable discrepancies in the scores, they discuss and adjust scores if convinced. The three analysts' ratings are then averaged to produce the overall article rating. Sometimes articles are rated by larger panels of analysts for various reasons—for example, if significant outlier scores remain, the article may be sent to an additional pod to receive additional input.

Occasionally, content is rated asynchronously by experienced analysts, as was our initial multi-analyst rating project of 2019. However, when analysis is done asynchronously, the commitment to a politically-balanced multi-analyst approach to each piece of content remains. Additionally, though rare, an outlier score continues to move the content into a review stage for additional input.

The type of rating we ask each analyst to provide is an overall coordinate score on the chart (e.g., "40, -12"). The rating methodology is rigorous and rule-based. There are many specific factors we take into account for both reliability and bias because there are many measurable indicators of each. The main ones for Reliability are defined metrics we call "Expression," "Veracity," and "Headline/Graphic," and the main ones for Bias are ones we call "Political Position," "Language," and "Comparison." There are several other factors we consider for certain articles. Therefore, the ratings are not simply subjective opinion polling, but rather methodical content analysis.

The "Veracity" factor is of particular importance in the reliability score. Our analysts use a veracity-checking methodology which incorporates best practices of fact-checking, such as lateral reading and consulting primary sources, but which is designed to be broad enough to cover claims that are not fact-checkable and quick enough to make an evaluation on every article. For more information on our veracity evaluation methodology, see this video.

Overall source ratings are composite weighted ratings of the individual article and show scores.

We continue to refine our methodology as we discover ways to have analysts classify rating factors more consistently. Our analysts use our software platform called CART–Content Analysis Rating system. This ratings software is currently available for use by educators in classrooms, and by individual adult learners in our news literacy courses. Educators and individuals can learn how to rate news articles like Ad Fontes Media. Our courses include detailed video and written explanations of the factors we use to rate articles.

V. Data Analysis Results

The easiest way to see the resulting ratings for each article and show is, of course, on the Interactive Media Bias Chart. By clicking on a button for a particular source, you can see a scatter plot of each article or episode rated for that source. The overall score of an article or show was the average of at least three individual scores.

To see an individual article or TV show and its score, you can search the table function just below the chart. Searching for a source name will pull up all the individual articles for that source along with their scores, so if you like, you can click on the URL to read the story and compare it to the score.

Close observers of the Interactive Media Bias Chart will notice that, particularly for low-scoring sources, the overall source scores appear to be lower than what would be expected from a straight average. As previously mentioned, this is because in our overall source-ranking methodology, we weight extremely-low-reliability and extremely-high-bias article scores very heavily.

The reason is this: the lowest rows of the chart indicate the presence of content that is very unreliable, including selective or incomplete stories, unfair persuasion, propaganda, misleading information, inaccurate, and even fabricated information (these are listed in order of egregiousness). Therefore, it is unacceptable for reputable news sources to include this type of content, even if it is infrequent or not the majority of the content. A source that has even 5% inaccurate or fabricated information is highly unreliable. A source that "only" publishes misleading or inaccurate content 33% of the time is terrible. In our system, they do not get credit for the 67% of stories that are merely opinion, but factually accurate.

A straight average, in such cases, would result in a higher overall source score—one that is inconsistent with the judgment of most savvy news consumers. Therefore, article scores of less than 24 for reliability are weighted very heavily.

We also rate bias scores more heavily the further the scores are away from zero. This results in sources with left or right leaning opinion content mixed with neutral/balanced content skewing more overall toward the bias of their opinion content. For example, The New York Times and Wall Street Journal skew left and right, respectively, due in large part to their opinion section content.

All other article scores for sources were straight-averaged. For example, if a news source had a mix of "fact reporting," "complex analysis," "analysis," and "opinion" articles, those would be straight averaged. As shown, our taxonomy rewards high percentages of fact reporting and complex analysis in sources and slightly down-ranks them for high percentages of opinion content (via straight averages). It does not punish a source for opinion content, because opinion content does have a useful place in our information ecosystem. However, our system does punish unfair opinion and worse content—that which we view as most polarizing "junk news."

VI. Next Steps

Our analysis of news sources began with funding from an Indiegogo crowdfunding campaign that raised $32K. These funds were used to develop the technological infrastructure for collecting the data, and the labor involved in the

initial project, which means that everyone who helped get us started spent a lot of time and effort for not very much money.

Since the completion of the original project, we have continued to raise additional funds, including over $350K from a WeFunder investment campaign. We have also built a number of products and services, the sales of which are moving us close to being self-sustaining. All the while, we continue to refine our methodology, improve our technical infrastructure, and dramatically expand the size of our data set.

Moving forward, we plan to continue rating more articles and shows in an ongoing fashion. We have built the software and database infrastructure to do so, and we are excited to continue this work.

Rating more articles and shows does require financial resources to train and pay analysts. Although many people have a passion for this, and some could do it for a bit of time for free, it is time consuming, and people deserve to be paid for the hard and diligent work they do to rate sources. Therefore, we will continue to roll out products and services based on our Media Bias Chart data to fund our ongoing efforts. We hope you will find them valuable and purchase them if they meet your needs as a news consumer, an educator, a marketer, a data services provider, a publisher, or an aggregator.

We will refine our training and methodology with what we have learned so far and will continue to iterate and improve. We will also implement technology tools, including AI, in order to increase our rating capability.

The scope of content we plan to rate ultimately includes anything that presents itself as a source of news or news-like information—including sources that exist solely on social media. It's a big vision, but that is what is required to help people navigate the news landscape, clean up our information ecosystem, and maintain the health of our democracy.

More to come in the future.