

Running head: A QUESTION OF BALANCE

A Question of Balance:

Are Google News search results politically biased?

By Eric Ulken

USC Annenberg School for Communication

May 5, 2005

Abstract

This study examines search results from the popular online news portal Google News in an effort to determine whether they are politically biased. By analyzing the content of third-party articles returned in a search on a political candidate (“George W. Bush,” for example), it is possible to assess the level of bias in the search results. Articles returned in searches on the two leading presidential candidates in the weeks before the 2004 election were collected, and a random sample of the highest-ranking results was analyzed for favorability to each candidate. Results from the same searches on Yahoo News were used as a benchmark for comparison. The data show that articles returned in Google News searches are significantly more likely to have a political bias than those returned in searches on Yahoo News, but there is no evidence of an overall conservative bias in search results on Google News, as has been suggested.

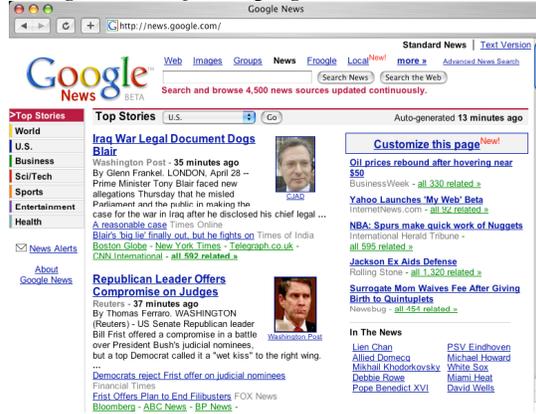
A Question of Balance: Are Google News search results politically biased?

As online news has grown in popularity, a number of sites have sprung up to catalog the wealth of news content available on the Internet. These so-called news portals include Google News, Yahoo News, Topix.net and MSNBC's Newsbot. Through an automated process known as "spidering" or "crawling," these aggregators index the content of selected news sources and allow users to browse and search recent news stories, usually linking to the source of the article for the full text.

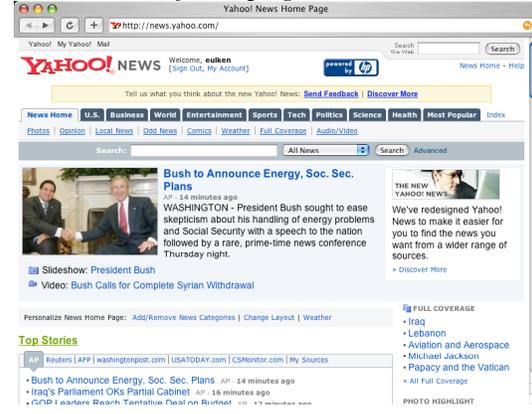
Google News, launched in 2001 but still in the "beta-testing" phase, has become one of the Internet's most popular news portals, drawing about 5.9 million visitors a month (Gaither, 2005). It indexes the top stories on some 4,500 English-language news sites, updating its index roughly every 15 minutes (Google, 2004). Google's innovative method of identifying top stories based on how frequently they appear on sites in its index – and doing so entirely without human intervention – has made the site a target of criticism since its inception. The efficiency with which Google News is able to automatically determine relative importance of stories and present a "front page" with top stories in different subject areas has been seen by some as the first ominous sign that computers will eventually make human editors obsolete. At the same time, users have ridiculed flaws in Google News's algorithms that cause it to occasionally attach a photo

to an unrelated article or elevate a relatively minor story to a prominent spot on its front page.

Google News front page



Yahoo News front page



Google does not share the list of sources it crawls, but searches often reveal results from relatively obscure online-only news sites – including some that are best described as weblogs – leading to questions about Google News’s criteria for inclusion and the notion that there might be some political imbalance in the sites it crawls. Google has taken a lot of criticism recently for the quality of news content in its index. Earlier this year, it removed from its index several sites, including the white supremacist journal National Vanguard, after users complained that hate speech was turning up in searches. Practices at Google News have come under additional scrutiny since March, when Agence France-Presse filed a lawsuit alleging that Google infringed its copyright by displaying AFP material on Google News pages. The Associated Press has also expressed “concern” about Google’s use of its material without payment (Gaither, 2005).

Janice Castro, director of graduate programs at Northwestern University’s Medill School of Journalism and one of the founders of the Online News Association, told CNET News.com the problem with Google News is that it gives users no way of

evaluating the quality of news sources (Olsen, 2005). “The best is mixed up with things that are far from the best,” Castro said.

Web journalism pioneer J.D. Lasica was among the first to suggest a conservative bias in Google News (2004). “What’s going on?” he wrote in an article for *Online Journalism Review* as the 2004 presidential race was heating up. “Have Google’s search results been hijacked by Fox News?” Lasica cited several stories from “second-tier” online-only news and commentary sites in a search on the words “John Kerry.” The headlines returned included the following:

- “John Kerry Said ‘Bring It On,’ Now Whines To Bush To Stop The Ads”
- “The Imploding John Kerry”
- “Swift Boat Veterans for Truth Expose John Kerry’s Lies”
- “John Kerry is Definitely ‘Unfit for Command’”

Political bias in news coverage has been the topic of numerous academic studies, most of which have themselves been subjected to charges of bias. Bias is practically impossible to quantify absolutely, but it can be measured in relative terms. One of the most widely cited and controversial recent studies attempts to do just that. Groseclose and Milyo (2003) assess bias among major news outlets – including *The New York Times*, *USA Today* and Fox News’ “Special Report” – by looking at how often they cite certain politically active “think tanks” and comparing this with how frequently members of Congress cite the same sources in floor speeches. The authors’ assumption is that conservatively biased news organizations, say, will cite a certain think tank with the same frequency as a conservative member of Congress. The study’s conclusion is that the mainstream media have an overwhelming liberal bias and the most unbiased news source

is Fox News' "Special Report." This finding was met with as much outrage as praise when the study first made the rounds of politically oriented weblogs (Dallas, 2004; Tabarrok, 2004) – proof that bias is always relative to the observer.

This study attempts to scientifically test Lasica's casual observation of bias in articles linked from Google News. Unlike the Groseclose/Milyo study, it does not have the benefit of an independent benchmark for comparison. Instead, it compares Google News with a more established competitor, Yahoo News, in an attempt to determine relative bias. The study looks at *balance* within stories as an indicator of bias. A balanced article will presumably have roughly as many favorable references to the search term (i.e., the candidate) as it has unfavorable ones. Given what Lasica and others have reported about Google News search results, the goal of this study is to prove or disprove quantitatively the assertion that Google News displays a conservative bias.

The research question and hypothesis are as follows:

RQ1: Are Google News search results politically biased?

H1: Results of Google News searches on the two major-party presidential candidates will reveal a conservative bias.

The research hypothesis is tested by means of a quantitative text analysis of articles returned in Google News and Yahoo News search results.

Method

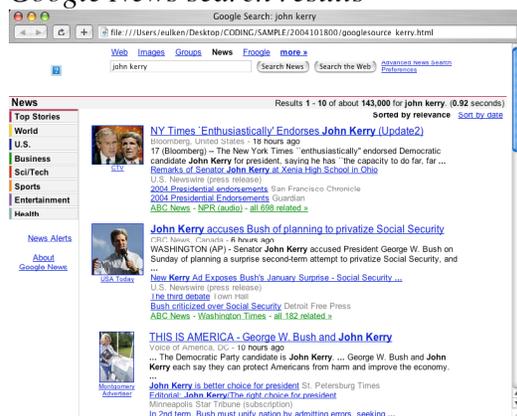
This study analyzes articles returned in searches on the full names of the two major-party presidential candidates ("George W. Bush" and "John Kerry") in the weeks

leading up to the 2004 election in order to determine a bias score for each article and, ultimately, to quantify the overall bias of the search results.

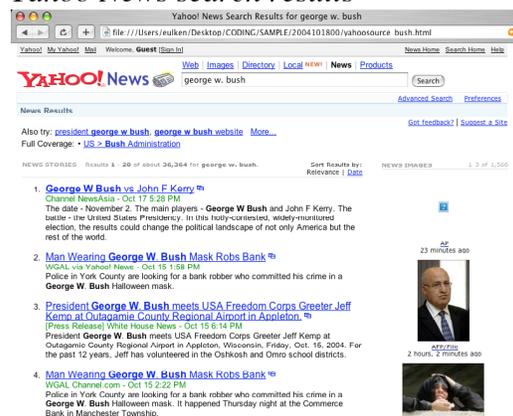
Data acquisition

Since the news search engines update their indexes frequently over the course of a day, the results for a particular search term can change from one minute to the next. A data acquisition scheme was devised that respects the dynamic nature of the search results. A computer program was written to retrieve the first 10 articles returned by Google News and Yahoo News for each search term (“George W. Bush” and “John Kerry”) at four-hour intervals and save them.

Google News search results



Yahoo News search results



The program was run for the period of Oct. 17-30, 2004, the two weeks preceding the Nov. 2 presidential election, resulting in a total of 80 “snapshots.” Each snapshot contained four sets of search results: “George W. Bush” on Google News, “George W. Bush” on Yahoo News, “John Kerry” on Google News and “John Kerry” on Yahoo News. The program also downloaded the full text of the first 10 articles returned in each result list.

Sampling scheme

Taking the top 10 articles in each list would yield 3,200 texts. A more manageable sample of 100 was selected for analysis. In order to generate a representative sample, a two-stage sampling process was devised that divided the data collection period into five sequential periods of equal length and then randomly selected one snapshot from each period. The stratified selection provided for a sample that was spread fairly evenly over the two weeks, so that a single news event would be unlikely to dominate the sample. The random selection stage ensured that the final five snapshots represented a variety of dayparts and days of the week.

The following snapshots were selected:

- Monday, Oct. 18, 2004, midnight
- Thursday, Oct. 21, 2004, 8 a.m.
- Monday, Oct. 25, 2004, midnight
- Thursday, Oct. 28, 2004, 4 p.m.
- Saturday, Oct. 30, 2004, 8 a.m.

For each snapshot, the first five articles from each of the four result lists were selected for analysis, ensuring an equal number of Bush and Kerry results and an equal number of Google News and Yahoo News results. In a couple of cases, the complete text of an article was behind a paid subscription wall. Where possible, a shortened free version was used; otherwise the article was skipped and the next highest-ranked article was used instead.

Units of observation

The articles were subdivided by sentence – with a sentence representing a single unit of observation. The decision was made to use sentences, rather than propositions, as the units of observation for two reasons:

- The texts can be parsed into sentences with minimal work, an important consideration given the volume of the data.
- Propositions – that is, groups of words expressing a distinct idea, whether as a phrase, a sentence or multiple sentences – must be manually identified and parsed by coders. This introduces the problem of unitizing reliability.

Because a number of the texts were extremely long, only the first 25 sentences of each article were coded. It was assumed that an article’s overall bias would be apparent within the first 25 sentences. The results section offers data that support this assumption. Overall, 1,587 sentences were coded.

Coding scheme

Each sentence could be coded in one of five ways:

1. Unfavorable to Kerry
2. Favorable to Kerry
3. Neutral
4. Unfavorable to Bush
5. Favorable to Bush

A coding manual was created to guide coders in the process of evaluating units.

(See Appendix A for the full coding manual.) Because the coding scheme involves

assessing latent meanings of sentences – and the terms “favorable” and “unfavorable” are imprecise – the most important purpose of the coding manual is to define these terms and specify how to apply them. The basic coding rules used are as follows:

- A unit (sentence) can only be coded as favorable or unfavorable if it contains an unambiguous message that, taken independently of other units, is favorable to one candidate or the other. Otherwise it must be coded neutral.
- If a unit contains both favorable and unfavorable references to the same candidate, or it contains only favorable or unfavorable references to both candidates, it should be coded neutral.
- If a unit contains one or more favorable references to one candidate and one or more unfavorable references to the other candidate, only the first reference in the sentence should be considered in coding.
- If there is uncertainty about how different people might interpret a unit, it should be coded neutral.

For purposes of assessing favorability, direct quotations and other attributed statements are treated no differently from statements made by the article’s author, since the choice of one particular quote over another can also represent bias. For example, the sentence, “Democrats accused Bush of misleading the nation about the justification for war in Iraq,” would be coded as unfavorable to the president.

The coding scheme attempts to make the rating process as objective as possible. However, coders’ personal biases could affect how they evaluate elements of the texts. While it is true that such coder subjectivity may skew balance scores of individual

articles in one direction or another, it is still possible to compare average scores for Google News and Yahoo News and thus assess bias of one site relative to the other.

Coding procedure

Two coders were used in this study. A primary coder – the author of the study – analyzed all 100 texts. A second coder – also a graduate student in communication with experience in content analysis – analyzed 28 of the texts, selected at random, for the purpose of assessing the validity of the coding scheme and the reliability of the primary coder. Articles were assigned to coders in random order using a computer-based coding system. The coders had no knowledge of whether a particular article came via Google News or Yahoo News (or which search term returned it). Coders were given only the headline and the name of the source organization for each article. The coders were trained and the coding scheme was initially tested using sample texts from articles not included in the actual sample.

In addition to coding each sentence, coders were asked to assess the overall favorability of the article in the form of two variables, each with five possible values:

- Overall favorability to Bush: highly favorable, favorable, neutral, unfavorable, highly unfavorable
- Overall favorability to Kerry: highly favorable, favorable, neutral, unfavorable, highly unfavorable

These overall favorability scores were intended only for the purpose of validating the unit-by-unit coding scheme and ultimately were discarded.

Intercoder reliability

Intercoder reliability is assessed on two levels: by sentence (the unit of observation) and by article (the unit of analysis).

At the sentence level, Cohen's *kappa*, a measure of agreement between coders on nominative variables, is computed as 0.72 – just above the 70% threshold considered an acceptable level of agreement. When intercoder reliability is tested at the article level, the agreement between the two coders is closer. Because the favorability scores computed for each article are ratio measurements, the two coders' scores are fit against each other, with the *r-square* statistic used to express intercoder reliability. Values of *r-square* for the Kerry favorability and Bush favorability variables are 0.94 and 0.90, respectively – indicating a relatively high level of agreement between the coders.

When computing reliability using such a small number of measurements (only 28 articles were coded by both coders), a single outlying value can greatly affect the outcome. It was necessary to omit one such outlier from consideration in computing agreement on the Kerry favorability score. The offending score was for an article that contained only one sentence, which one coder deemed favorable to Kerry and the other coder recorded as neutral. As a result, the article took a Kerry favorability score of 1 from one coder and 0 from the other. This greatly affected the reliability calculation, causing *r-square* to plummet to 0.57. The article was excluded from the data analysis, which is why the data show one article fewer for Yahoo News than for Google News.

Results

Using the values for each sentence, two scores are calculated for each article,

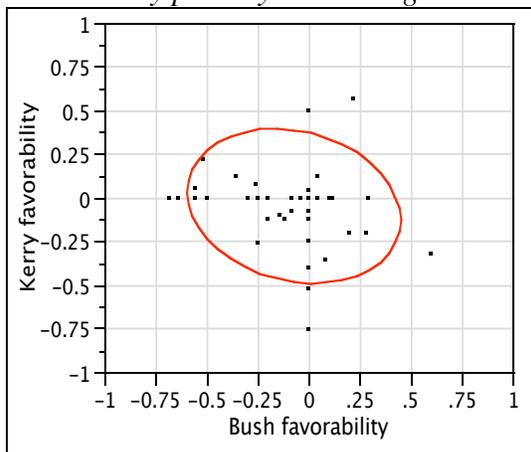
measuring the degree of the article’s overall favorability to each candidate. Bush and Kerry favorability scores for each article are computed using the following formula:

$$\text{favorability score} = \frac{(\text{sum of favorable units}) - (\text{sum of unfavorable units})}{(\text{total units coded})}$$

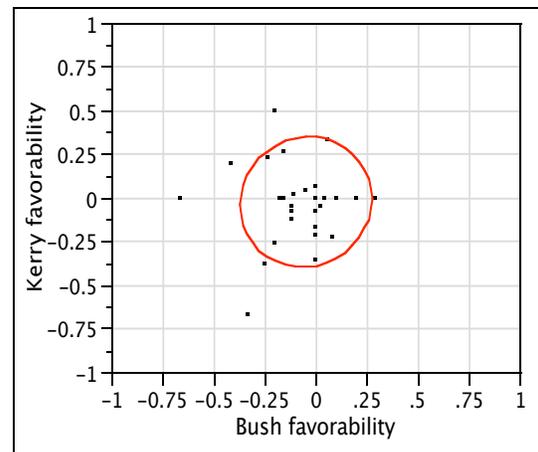
Favorability scores can thus take values of –1 (completely unfavorable) to 1 (completely favorable), with 0 being neutral. For instance, a Kerry favorability score of –0.3 for an article would indicate that, on balance, 30% the content of an article is unfavorable to John Kerry (the actual proportion of unfavorable units might be 35% but offset by 5% of units coded as favorable to Kerry). Because even the most biased articles contain a lot of neutral (or irrelevant) content, the scores tend to be closer to 0 than to either extreme.

Two scatterplots – one for Google News and the other for Yahoo News – provide a basic summary of the data. They show the two candidates’ favorability scores for each article, plotted against each other. This facilitates comparison of the overall favorability of the two portals’ search results.

Favorability plots by search engine



Google News



Yahoo News

Each data point represents an article, and its placement on the chart represents its favorability to the two candidates:

- Upper left quadrant: Article is favorable to Kerry and unfavorable to Bush
- Upper right quadrant: Article is favorable to both
- Lower right quadrant: Article is favorable to Bush and unfavorable to Kerry
- Lower left quadrant: Article is unfavorable to both

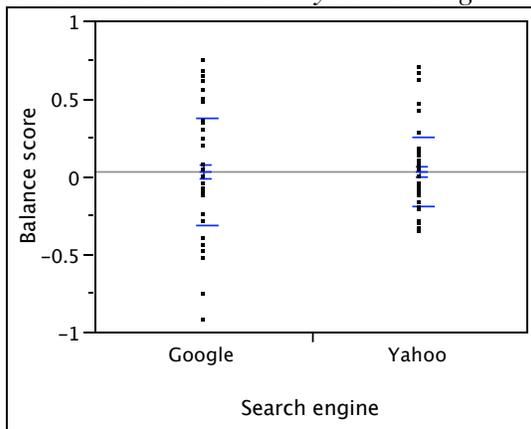
In other words, articles in the upper right and lower left are more balanced than those in the upper left and lower right. Articles closer to the center are more neutral. The circular boundary is a density ellipse drawn around 90% of the data points, which makes it easier to see patterns in the data. One fact that is not apparent in the scatterplots is that a large number of data points are at the coordinates (0, 0). This is because many of the articles – 22% for Google News and 45% for Yahoo News – exhibited no bias at all, either because they discussed both candidates with complete neutrality or because they were not relevant to either candidate. A glance at the two plots reveals what can be seen empirically in the search results from the two sites: Articles returned in the searches using Google News are more likely to be biased in favor of one candidate and against the other, while those that turn up in the Yahoo News searches are generally more balanced.

In order to illustrate article bias in one dimension, a measurement that takes into account favorability ratings for both candidates is needed. Two related scores are devised for this purpose. The first, the article balance score, shows the degree to which articles favor one candidate over the other. It is computed using a simple formula:

$$\textit{balance score} = \textit{kerry favorability} - \textit{bush favorability}$$

An article's balance score takes a value between -1 and 1 , with positive numbers indicating greater favorability to Kerry and negative numbers indicating greater favorability to Bush. Articles with balance scores of 0 are equally favorable (or unfavorable) to both candidates.

Article balance scores by search engine

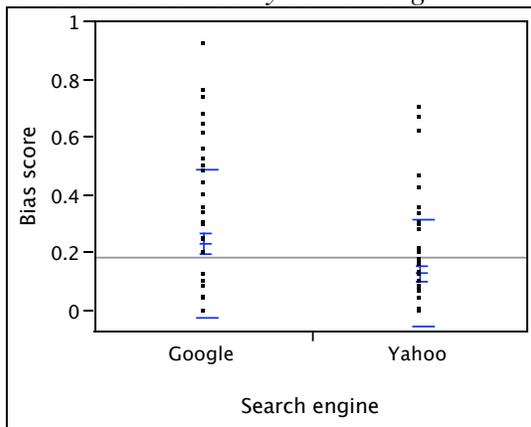


By taking the average balance scores for articles returned by Google News and Yahoo News, each search engine's overall bias can be determined. A balance score that favors Bush is presumed to show a conservative bias, while one that favors Kerry would indicate a liberal bias. The average balance scores for both Google News and Yahoo News are not significantly different from 0 , indicating an absence of overall bias in the search results for both sites. Thus, while the data do show bias in many of the articles returned by Google News, there is no evidence of an overall conservative (or liberal) slant to the site's search results, as has been alleged.

The second measurement, the article bias score, is simply the absolute value of the balance score. It takes a value between 0 and 1 and represents the proportion of an article that is biased, regardless of the direction of bias. For example, an article in which

half the sentences are coded as favorable to Kerry and the other half as unfavorable to Bush would have a bias score of 1, meaning 100% of the article is biased. In fact, bias scores tend to be closer to 0, though one article returned by Google News had a bias score of 0.92. The mean article bias score for each search engine describes the degree to which the average article returned by that search engine is likely to be biased.

Article bias scores by search engine



As seen in the plot above, the articles returned by Google News have a higher mean bias score than those returned by Yahoo News (0.23 compared with 0.13, a statistically significant difference). This means that a search on Google News is likely to turn up articles that are more biased than those returned by its competitor.

Besides being coded for favorability, articles were also classified by whether they came from an independent, online-only source (such as Salon.com) or a website affiliated with a traditional news source. A traditional news source is defined as a wire service, newspaper, magazine, TV station, radio station, broadcast network or cable network. (Content from one of these sources that is syndicated on a news aggregator such as Yahoo News is also classified as traditional.) Of the articles returned by Google News, 40% were from non-traditional news sources, while only 24% of the Yahoo News results

came from non-traditional sources. (See Appendix B for a list of the sources of all the articles coded.) Notably, almost all of the bias in Google News's search results can be attributed to its use of non-traditional sources. In other words, when articles from non-traditional sources are left out of the calculation, the average bias scores for Google News and Yahoo News are virtually identical.

Finally, articles that exceeded the 25-sentence length limit for coding tended to be scored as slightly more biased than shorter articles, on average. A possible explanation for this is that most of the articles from traditional sources – those less likely to exhibit bias – were shorter than 25 sentences. If this arbitrary limit were hindering the coding scheme's ability to ascertain bias in longer texts, one would expect to find a lower average bias score for longer articles.

Discussion

The data show that articles returned in Google News searches are more likely to have a bias toward a particular candidate than those returned in searches on Yahoo News, but there is no evidence of an overall conservative bias in search results on Google News, as has been suggested. Both Google News and Yahoo News searches returned articles that were, on the whole, equally favorable to both George W. Bush and John Kerry. This is what one would expect to see of balanced search results at a time when public opinion is evenly divided between the two candidates.

Accordingly, the research hypothesis, H1, is rejected.

For both candidates, a slight tendency toward negativism (that is, more unfavorable content than favorable) can be seen in articles returned by the two news

portals. This can be explained in two ways:

- By the time the data were collected – in the two weeks before the 2004 general election – the race for the presidency had turned increasingly negative.
- The news media by their nature generally place greater emphasis on negative stories than positive ones.

With the abundance of well-respected, credible sources on the Internet, why does Google News return so many articles from biased sources? An explanation offered by Nathan Stoll, Google’s associate product manager for Google News, has to do with the search terms themselves (2004): A search for “John Kerry” will first return entries in which the entire search term appears in the headline. Even though Google News examines the full text of articles when looking for a search term, it puts extra weight on the headline when it ranks the results. So, stories with headlines such as “John Kerry lies about his record” will receive a higher rank than stories with headlines such as “Kerry campaigns in Ohio” (omitting his first name). Traditional media – which tend to be less biased than many alternative, online-only news sources – generally identify people in headlines by their last names only. As a result, articles from these news organizations may often be outranked in Google News search results by those from sites that do not follow this practice. Given this peculiarity, the use of full names in the searches analyzed here could be seen as a weakness in the study, but in fact it emulates the behavior of an average user. Unaware of the distinction, a user presumably is more likely to search for a full name than just a last name, resulting in a disproportionate number of results from non-traditional news sources. Searches on Yahoo News do not appear to exhibit this tendency as frequently.

It is important to understand that this study is not an indictment of Google News's practice of automatically ranking the top stories on its front page and section fronts. While Google News's ranking methods may be flawed, as some have charged, this study is concerned only with the site's search results. It should also be noted that Google News does not distinguish between factual and opinion pieces in its search results (Stoll, 2004). Thus, an editorial may appear along with straight-news stories, even though the former represents a particular point of view while the latter are supposed to be reasonably balanced recitations of fact. It is not clear that average users can make the distinction, especially given the many online-only sources that often peddle a confusing mixture of fact and opinion. Accordingly, this study makes no attempt to separate news from editorial content.

The main flaws in the study are with the coding scheme. Better coder training, a more detailed coding manual and a more precise definition of "favorability" would almost certainly have improved intercoder reliability, which, while not low enough to call into question the results, is below expectations. Additionally, using sentences as units of observation makes for some ambiguity in the coding process. If one sentence contains multiple distinct propositions, or a single proposition stretches across multiple sentences, some of this granularity is lost in the current coding method.

If users are looking for current factual information about a political candidate, this study concludes that they are more likely to find it by searching Yahoo News. If, on the other hand, users want a wide range of alternative viewpoints, then Google News may be their best bet.

References

- Dallas, Jim. (2004) "From the Department of 'Huh?'" Burnt Orange Report [weblog]. Retrieved April 12, 2005, from <http://www.burntorangereport.com/archives/001234.html>
- Gaither, Chris. (2005, April 11). "Web Giants Go With Different Angles in Competition for News Audience." *Los Angeles Times*, p. C-1.
- Google. (2004). *About Google News*. Retrieved Dec. 9, 2004, from http://news.google.com/intl/en_us/about_google_news.html
- Groseclose, Tim, and Milyo, Jeff. (2003). *A Measure of Media Bias* [working paper]. Retrieved Dec. 9, 2004, from <http://www.stanford.edu/~wacziarg/mediapapers/GrosecloseMilyo.pdf>
- Krippendorff, Klaus. (1980). *Content analysis: An Introduction to its Methodology*. Beverly Hills, CA: Sage.
- Lasica, J.D. (2004, Sept. 24). Balancing act: how news portals serve up political stories. *Online Journalism Review*. Retrieved Dec. 9, 2004, from <http://ojr.org/ojr/technology/1095977436.php>
- Lee, Martin A., and Solomon, Norman. (1990). *Unreliable Sources: A Guide to Detecting Bias in News Media*. New York: Carol Publishing.
- Olsen, Stefanie, and Hansen, Evan. (2005, March 25). "All the news that robots pick." CNET News.com. Retrieved April 28, 2005, from http://news.com.com/2100-1038_3-5635161.html
- Stoll, Nathan. (2004, Nov. 13). [personal communication with Google News associate product manager]
- Tabarrok, Alex. (2004). "Surprise! Fox News Is Fair and Balanced!" *Marginal Revolution* [weblog]. Retrieved April 12, 2005, from http://www.marginalrevolution.com/marginalrevolution/2003/09/surprise_fox_is.html
- Weber, Robert Philip. (1990). *Basic Content Analysis*. Newbury Park, CA: Sage.

Appendix A: Coding instructions

Procedure

When you are assigned an article to code, you will evaluate it in two ways: unit-by-unit and overall. The units of observation are sentences. When you code unit-by-unit, you must consider only the individual unit you are coding. When you code the article overall, you can consider aspects of the article, such as the headline, that can't be taken into account in a unit-by-unit analysis. If there are any technical or procedural irregularities in the coding, please make a note of it in the comments field.

When you have completed the coding process for an article, double-check your response (since you can't go back) and hit the "Submit responses" button. Your responses will be recorded, and you'll be given the opportunity to continue on to another article.

Guidelines

We are looking for favorable and unfavorable references to John Kerry in the results of a search on his name, and the same for George W. Bush. You will be coding the text of articles returned in search queries Yahoo News and Google News.

To keep coders from spending an inordinate amount of time on any one story, stories longer than 25 units (sentences) will be truncated. Most stories are shorter than this anyway.

Here are the basic rules for coding individual units:

- A unit can only be coded as favorable or unfavorable if it contains an unambiguous message that, taken independently of other units, is favorable to one candidate or the other. Otherwise it must be coded neutral.
- If a unit contains both favorable and unfavorable references to the same candidate, or it contains only favorable or unfavorable references to both candidates, it should be coded neutral.
- If a unit contains one or more favorable references to one candidate and one or more unfavorable references to the other candidate, only the first reference in the sentence should be considered in coding.
- If there is uncertainty about how different people might interpret a unit, it should be coded neutral.

On the coding form, mark each sentence as favorable to Bush, favorable to Kerry, unfavorable to Bush, unfavorable to Kerry or neutral. Please observe the following definitions when considering what are favorable and unfavorable references.

The following may be considered favorable or unfavorable references:

- Direct references to the candidate (by name or other obvious identifier -- e.g., "my opponent", "the senator")
- Quotes from candidates (or their surrogates) about themselves or each other
- References to actions or statements by the Bush or Kerry campaigns
- News directly related to candidates' issues or policies where it is clear that the news is damaging or helpful to a particular candidate

The following should be left marked as "neutral":

- Mentions of the actions of parties, aides, colleagues, etc., unless they are directly related to the campaign
- General ideological assertions and political observations that are subjective and can't be considered positive or negative for either candidate (e.g., "big government is bad" or "social security is broken")
- Any citation of poll results (since it is difficult to weight poll results fairly)
- Any mention that cannot be clearly determined to be favorable or unfavorable to a particular candidate

What constitutes favorable and unfavorable?

- If a reference cannot be clearly construed as favorable or unfavorable (e.g., "John Kerry has a rich wife" could be interpreted either way), it should be ignored
- Instances where a favorable adjective is used to describe a neutral or unfavorable action (e.g., "...efficient in his criticism of Bush") do not count as a favorable mention. Same with unfavorable adjectives.
- Historical references can be coded as favorable or unfavorable only if there is a clear relationship to the candidate and it can be clearly discerned as being favorable or unfavorable (e.g., "Truman didn't apologize for war mistakes, so Bush shouldn't have to either" could be coded as favorable to Bush)
- If a candidate's actual or alleged associate or ally is portrayed negatively (or positively), the ally's relationship to the candidate counts as a single unfavorable (favorable) reference. (e.g., "Arafat is a murderer. Arafat is a

thug. Arafat is derailing the peace process. Arafat endorses Kerry." Only the last sentence is coded as unfavorable to Kerry.)

Appendix B: Sources of articles returned

Yahoo News returned more articles from traditional media sources (in boldface) than Google News did. A traditional news source is defined as a wire service, newspaper, magazine, TV station, radio station, broadcast network or cable network, accessed either directly or through a news aggregator.

Google News

VOANews
truthout
Chicago Maroon
Ft. Worth Star-Telegram
Daytona Beach News-Journal
NME.com
UselessKnowledge.com
UnconfirmedSources.com
The Jewish Press
PRNewswire
AxisofLogic
The (Carlisle, Pa.) Sentinel
Sioux Falls Argus Leader quoted on Lucianne.com
Biloxi (Miss.) Sun-Herald
Denver Post quoted on Lucianne.com
Bloomberg
CBC News (Canada)
VOANews
Jerusalem Post
International Herald Tribune
New York Post
UselessKnowledge.com
Xinhua
AP via San Jose Mercury News
AP via Duluth News Tribune
Salon.com
Bloomberg
The Australian
INDOLink
Lawrence (Kan.) Journal-World
CNN
AP via canada.com
Business-Standard.com (India)
MichNews.com
RushLimbaugh.com
TVM (Maldives) via MaldivesInfo
PRNewswire via Yahoo News
MichNews.com
westcoastmusic
ReadaBet.com
ThisDay (Nigeria) via AllAfrica.com
AP via Canada.com
The Washington Dispatch
RushLimbaugh.com
UselessKnowledge.com
s5000.com
Washington Times via The Conservative Voice
TVM (Maldives) via MaldivesInfo
Scranton (Pa.) Times Tribune
AP via WHEC-TV (Rochester, N.Y.)

(60% traditional media sources)

Yahoo News

ChannelNewsAsia.com
TheWGALChannel.com (Harrisburg, Pa.)
Whitehouse.gov
TheWGALChannel.com (Harrisburg, Pa.)
AP
PRNewswire via Yahoo News
PRNewswire
Whitehouse.gov
Whitehouse.gov
Guardian Unlimited (U.K.)
The Southern Illinoisan
AP via Daily Herald (Arlington Heights, Ill.)
The Smoking Gun
AFP via Yahoo News
Whitehouse.gov
Bloomberg
AFP
WCPO.com (Cincinnati)
New Zealand Herald
AP
AP via Yahoo News
WPXI.com (Pittsburgh)
AP via Duluth News Tribune
India Daily
AFP via Yahoo News
Bloomberg
INDOLink
Knight Ridder
Bloomberg
Reuters via Australian Broadcasting Corp.
AP via Canada.com
WhiteHouse.gov
WorldNetDaily.com
Whitehouse.gov
AP via PhillyBurbs.com (N.J.)
PRNewswire via Yahoo News
Kyodo News via Yahoo Asia
AP via WNEP-TV (Scranton, Pa.)
AP via Canada.com
KYW Newsradio 1060 (Philadelphia)
The (Youngstown, Ohio) Vindicator
The Times of India
AP via Canada.com
IndieWire
India Online
AP via WHEC-TV (Rochester, N.Y.)
AFP via Khaleej Times (U.A.E.)
Editor and Publisher via Yahoo News
AP via Yahoo News

(76% traditional media sources)