



PROJECT FOR EXCELLENCE IN JOURNALISM

The Year in News 2010:

Disasters, Economic Anxiety, but Little Interest in War

Methodology

The content analysis research in **The Year in News 2010: Disaster, Economic Anxiety, but Little Interest in War** is the summation of a year's worth of coding conducted by PEJ. The coding is done throughout the year with weekly findings reported in the News Coverage Index, or NCI reports.

All coding was conducted by PEJ's trained staff of in-house researchers.

The 2010 analysis totals 50,508 stories examined from January 1 through December 19, 2010. This consists of 5,414 newspaper articles, 7,523 online stories, 12,524 stories from network television, 16,338 stories on cable news and 8,709 stories from radio programs.

The **2010 Year in News Interactive** uses the data from PEJ's **News Coverage Index**. It was designed to allow users to explore and answer questions about media coverage in 2010.

This Year in News Report also includes a year's worth of analysis of social media from the **New Media Index**, an analysis of millions of blog and twitter posts on the web tracking sites Icerocket and Tweetmeme. It covers the period from January 1 through December 10, 2010.

In addition, the measures of public interest in the news are from the **News Interest Index**, which is derived from weekly national public opinion surveys conducted in 2010 by the Pew Research Center for the People & the Press.

News Coverage Index

The central focus of the study is to analyze a wide swath of American news media to identify what is being covered and not covered—the media's broad agenda.

The Universe: What we are Studying

Because the landscape is becoming more diverse-in platform, content, style and emphasis-and because media consumption habits are also changing, even varying day to day, the Index is designed to be broad. Therefore, our sample, based on the advice of our academic team, is designed to include a broad range of outlets-illustrative but not strictly representative of the media universe.

The sample is also a purposive one, selected to meet this criteria rather than to be strictly random. It is a multistage sampling process that cannot be entirely formulaic or numeric because of differences in measuring systems across media. It involves the balancing of several factors

including the number of media sectors that offer news, the number of news outlets in any given sector, the amount of news programming in each outlet and the audience reach. In addition to front-end selections, we have also weighted the various sectors on the back end to account for differences in audience. The weighting process is discussed further down in this document.

The mainstream or establishment daily news media in the United States can be broken down into five main sectors. These are:

Network TV News
Newspapers
Online News Sites
Cable News
Radio News

Within each media sector, the number of outlets and individual programs vary considerably, as do the number of stories and size of the audience. We began by first identifying the various media sectors, then identifying the news media outlets within each, then the specific news programs and finally the stories within those.

The primary aim of the Index is to look at the main news stories of the week across the industry. With that in mind, for outlets and publications where time does not permit coding the entire news content offered each day (three hours of network morning programming, for instance), we code the lead portion. In other words, we code the first 30 minutes of the cable news programs, the first 30 minutes of the network morning news programs, the front page of newspapers, etc. This may skew the overall universe toward more "serious" stories, but this is also the most likely time period to include coverage of the "main" news events of the day, those that would make up the top stories each week or each month.

Below we describe the selection process and resulting sample for each main sector.

Note: The statistics cited here are the statistics that were accurate at the time of the launch of the Index (January 2007). When available, updated data are included in the footnotes section.

Network News

Sector Reach

Each evening, the three broadcast news shows on ABC, NBC, and CBS, reach approximately 27 million viewers. The morning news shows on those networks are seen by 14.1 million viewers.^[1] In addition, the nightly newscast on PBS reaches roughly 2.4 million viewers daily, according to their internal figures. Because the universe of national broadcast channels is limited to these four channels, it is practical to include all of the networks as part of our sample universe.

Sector Sample

Each of the three major broadcast networks produces two daily national general interest news shows, one in the morning (such as Good Morning America) and one in the evening. Therefore it is practical to include at least part of all these news programs on ABC, CBS, and NBC in our sample. (The magazine genre of programs are not included in the universe both because in most cases they are not daily-except for Nightline-and because they are not devoted predictably to covering the news of the day). At the same time, because the Newshour with Jim Lehrer is considered by many as an alternative nightly news broadcast to the three major networks, and it reaches a substantial audience, we also include that program.

Units of Study

For the commercial evening newscasts, the study codes the entire program. For the morning programs, it codes the news segments that appear during the first 30 minutes of the broadcast, including the national news inserts but not local inserts. By selecting this sample of the morning shows, it is possible that we will be missing some news stories that appear later in the programs. However, through prior PEJ research, we have learned that the morning shows generally move away from the news of the day after the first 30 minutes-save for the top-of-the-hour news insert-and present more human interest and lifestyle stories after that point. The stories that the networks feel are most important will appear during the first 30 minutes and be included in our study.

For PBS NewHour, where the second half of the program differs from the first half, beginning March 31, 2008, we began rotating between the first and second half hour of the show in order to get a closer representation of the program's overall content.

The resulting network sample is:

Commercial Evening News: Entire 30 minutes of 2 out of 3 programs each day (60 minutes)
Commercial Morning News: 1st 30 minutes of 2 out of 3 programs each day (60 minutes)
PBS NewsHour: Rotate to code the 1st 30 minutes one day, the 2nd 30 minutes the next day and then skip

This results in either 2 or 2.5 hours of programming each day.

Cable Television

Sector Reach

According to ratings data, the individual programs of the three main cable television news channels-CNN, MSNBC and Fox News-do not reach as many viewers as those of the broadcast network news shows. During prime time hours, 2.7 million viewers watch cable news, while 1.6 million watch during daytime hours.[\[2\]](#) But ratings data arguably undercount the reach of cable news. Survey data now find that somewhat more people cite cable news as their primary or first source for national and international news as do broadcast network news.

Sector Sample

The most likely option was to study CNN, MSNBC and Fox News. These represent the dominant channel of programming from each news-producing cable company. (This means selecting MSNBC as opposed to CNBC, and CNN as opposed to CNN Headline News, and MSNBC over Headline News, which now sometimes beats MSNBC in ratings).

Units of Study

Since these channels provide programming around the clock, with individual programs sometimes reaching fairly small audiences, it is not practical for us to code all of the available shows. On the one hand, there is a great challenge in selecting several times out of the day to serve as a sample of cable news overall.

On the other hand, earlier studies have shown that for much of the day, most people find one cable news program on a channel to be indistinguishable from another. If one were to ask a daytime viewer of cable news which program he or she preferred, the 10 a.m. or the noon, you might get a confused look in response. For blocks of hours at a time, the channels will have programs with generic titles such as CNN Newsroom, Your World Today or Fox News Live. Our studies have shown that there are four distinct programming sectors to cable: early morning, daytime, early evening and prime time.

Working with academic advisors we weighed various options. A selection based on the most-watched programs would result in the O'Reilly Factor (1.8 million viewers a night) for Fox and Larry King Live (500,000 viewers a night) for CNN.^[3] However, some of these shows are not news programs per se, but rather their content derives from the host's opinions and guests on any given day. Separating news and talk also proved problematic because it is often difficult to distinguish between the two categories, and several programs offer both news and talk in the same hour.

The best option, we concluded, was to draw from two time periods:

- 1) The daytime period, to demonstrate what on-going or live events are being covered. The study includes two 30 minute segments of daytime programming each day, rotating among the three networks.
- 2) Early Evening and Prime time (6 PM - 11PM) together as a unit, rather than separating out talk and news or early prime and late prime. Within this five hour period, we included all programming that focuses on general news events of the day. Basically, this removes three programs: Fox's Greta Van Susteren, which is more narrowly focused on crime, CNN's Larry King which as often as not is focused on entertainment or personal stories rather than news events and MSNBC's documentaries program.

Prior to January 1, 2009, 3 out of 4 evening cable programs were coded each evening for both Fox News and CNN. For MSNBC, 2 out of 4 evening programs were coded each evening. These

shows were rotated each weekday. In the past, MSNBC's ratings were significantly lower than the ratings for Fox News and CNN, and that was the justification for including one fewer of their shows each evening.

At the start of 2009, PEJ made a change to the evening cable sample. While MSNBC's ratings still trail both Fox News and CNN, the general trend over the last few years has been that the differences have decreased. Therefore, it becomes harder to justify having a different amount of programming for the 3 stations. Consequently, beginning on January 1, 2009, all 3 stations have 2 of their 4 evening cable shows coded on a nightly basis.

To include the most cable offerings possible each week, the study codes the first 30 minutes of selected programs and rotates them daily. Morning shows were not included because those shows are run at the same time for every part of the country - meaning that a broadcast that starts at 7 a.m. on the east coast will begin at 4 a.m. on the west coast. Those programs appear far too early for much of the country to actually view. This is in contrast to the broadcast morning programs, which are shown on tape delay in different parts of the country, in the manner of other broadcast programs.

This process resulted in the following cable sample:

Daytime

Rotate, coding two out of three 30-minute daytime slots each day (60 minutes a day)

Prime Time

Two 30-minute segments for Fox News (60 minutes)

Two 30-minute segments for CNN (60 minutes)

Two 30-minute segments for MSNBC (60 minutes)

The Index rotates among all programming from 6 to 11 p.m. that was focused on general news events of the day excluding CNN's Larry King Live and Fox's Greta Van Susteren.

Below is the current list of evening cable programs included in our sample as of January 2010.

	CNN	Fox News	MSNBC
6 p.m.	Situation Room	Special Report w/Bret Baier	The Ed Show
7 p.m.	CNN Tonight/ John King, USA	Fox Report w/ Shephard Smith	Hardball
8 p.m.	CNN Prime Time/ Parker Spitzer	The O'Reilly Factor	Countdown w/Keith Olbermann
9	-----	Hannity	The Rachel

p.m.			Maddow Show
10 p.m.	Anderson Cooper 360	-----	-----

**CNN made certain programming changes during 2010. Our sample included replacement shows for any given time slot when appropriate.*

This results in 4 hours of cable programming each day (including daytime).

Newspapers

Sector Reach

Roughly 54 million people buy a newspaper each weekday.^[4] This number does not include the "pass along" rate of newspapers, which some estimate, depending on the paper, to be approximately three times the circulation rate. In addition, specific newspapers, such as the New York Times and Washington Post, have an even greater influence on the national and international news agenda because they serve as sources of news that many other outlets look to in making their own programming and editorial decisions. So while the overall audience for newspapers has declined over recent years, newspapers still play a large and consequential role in setting the overall news agenda that cannot be strictly quantified or justified by circulation data alone. There is a growing body of data to suggest that the total audience of newspapers, combining their reach in print and online, may actually be growing slightly.

Sector Sample

To create a representation of what national stories are being covered by the 1,450 newspapers around the country, we divided the country's daily papers into three tiers based on circulation: over 650,000; 100,000 to 650,000; and under 100,000. Within each tier, we selected papers along the following criteria:

First, papers need to be available electronically on the day of publication. Three websites, www.nexis.com, www.newsstand.com, and www.pressdisplay.com, offer same day full-text delivery service. Based on their general same-day availability (excluding non-daily papers, non-U.S. papers, non-English language papers, college papers, and special niche papers) a list of U.S. general interest daily newspapers was constructed. The original list included seven papers in Tier 1, 44 papers in Tier 2, and 22 papers in Tier 3.

Tier 1: In Tier 1, we wanted to include a representation from the large nationally reputed or distributed papers, so each day we code two out of four of the largest papers, the New York Times, Los Angeles Times, USA Today, and Wall Street Journal.

Tiers 2 and 3: Four newspapers were selected from Tier 2 and Tier 3 respectively. To ensure geographical diversity, each of the four newspapers within Tier 2 and Tier 3 was randomly selected from a different geographic region according to the parameters established by the U.S.

Census Bureau, i.e., Northeast Region, Midwest Region, South Region and West Region. An effort was also made to ensure diversity by ownership.

We rotate two of the four newspapers in Tier 2 and one or two of the three newspapers in Tier 3 each day.

The following is the current list of newspapers included in the Index.

1st Tier

The New York Times
LA Times
USA Today
Wall Street Journal

2nd Tier

Washington Post
Tampa Tribune
Seattle Times
Columbus Dispatch

3rd Tier

The Day
Rome News Tribune
Ventura News

Units of Study

For each of the papers selected, we code only articles that begin on page A1 (including jumps). The argument for this is that the papers have made the decision to feature those stories on that day's edition. That means we do not code the articles on the inside of the A section, or on any other section. The first argument for ignoring these stories is that they would be unnecessary for our Index, which measures only the biggest stories each week. If a story appears on the inside of the paper, but does not make A1 at any point, it would almost certainly not be a big enough story to make the top list of stories we track each week. The weakness of this approach, arguably, is that it undercounts the full agenda of national and international news in that it neglects those stories that were not on Page 1 on certain days but were on others. While this is perhaps less pertinent in the weekly Index, at the end of the year, when trying to assess the full range of what the media covered, those stories that appeared on the inside of the paper but didn't vanish may be undercounted.

Part of the reasoning for excluding those national and international stories that begin inside the front section of the paper is practical. Coding the interior of the papers to round out the sample for year-end purposes would be an enormous amount of work for relatively minimal gain.

The other argument for foregoing national and international stories that fail to make Page 1 is more conceptual. We are measuring what newspapers emphasize, their top agenda. Given the cost versus the benefit, capturing the front page of more newspapers seemed the better alternative. In the same regard, we do not code every story that might appear on a Web site, an even more daunting task, but instead code just the top stories.

The other challenge with newspapers that we did not face with some other media is that we only include stories that are national or international. National is defined as a story being covered by newspapers from different locations, as opposed to a local story that is only covered in one paper. The only local stories included in the study are those that are pertinent to a larger national issue - how the war in Iraq is affecting the hometown, for instance, or new job cuts at the local industries because of the sliding economy.

This results in a newspaper sample of approximately 20 stories a day.

Online News Sites

Sector Reach

About 30 million internet users go online for news each day.^[5] About 6.8 million people read a blog each day, some of the most popular of which are news oriented.^[6] Both online news sites and web blogs are becoming more important in the overall news agenda. Any sample of the modern news culture must include a representation of some of the more popular online news sources.

Sector Sample

The online news universe is even more expansive than radio and has seemingly countless sites that could serve as news sources. To get an understanding of online news sources we chose to include several of the most popular news sites in our universe as a sample of the overall online news agenda. We also wanted balance in the type of online news sites, between those that produced their own content and those who aggregated news from throughout the web.

When the Index was originally launched in 2007, the sample included 5 prominent Web sites that were tracked each weekday. These sites were Yahoo News, MSNBC.com, CNN.com, AOL News, and Google News. However, considering the increased usage of internet for news shown recent surveys conducted by the Pew Research Center for the People & the Press, PEJ decided to expand our Internet content.

The increase in the number of sites included in the NCI took effect on January 1, 2009.

To choose the sites to be included in our expanded online sample, we referred to the lists of the top news sites based on the averages of six months of data (May 2008-October 2008) from two rating services: Nielsen and Hitwise. (Data providing the most-visited news sites ranked by specific Web addresses were not available from Comscore at the time of our sampling.)

First, we found the top general interest news sites ranked by their average unique audience data for six months based on Nielsen Netview monthly data on a subdomain level. Second, we found the top general interest news sites ranked by their average market share for six months based on monthly rankings for top news and media Web sites provided by Hitwise. We then averaged the ranks of the top sites on these two lists to determine the top 12 general interest news websites.

In 2009, we updated our sample by averaging 7 months of data (May - November 2009) provided by Nielsen Media Research.

The sites included in our current sample are as follows:

Yahoo News
MSNBC.com
CNN.com
NYTimes.com
Google News
AOL News
Foxnews.com
USAToday.com
Washingtonpost.com
ABCNews.com
HuffingtonPost.com
Wall Street Journal Online

Units of Study

For the online news sites, the study captures each site once a day. We rotate the time of day that we capture the Web sites between 9-10 am Eastern time and 4-5 pm Eastern time. For each site capture, we code the top five stories, since those have been determined to be the most prominent at that point in time by the particular news service. As is true with our decision about page A1 in newspapers, if a story is not big enough for the online sites to highlight it in their top five stories, it is likely not a story that will register on our tally of the top stories each week.

This results in a sample of 30 stories a day.

Radio

Sector Reach

Radio is a diverse medium that reaches the majority of Americans - 94 percent of Americans 12 years and older listen to traditional radio each week. [7] Approximately 16% of radio listeners tune into news, talk and information radio in an average week, which ranks it as the most popular of all measured radio formats. [8] Many more Americans get news from headlines while listening to other formats as well.

The challenge with coding national radio programs is that much of radio news content is localized, and the number of shows that reach a national audience is only a fraction of the overall programming. On the other hand, our content analysis of radio confirms that news on commercial radio in most cities has been reduced to headlines from local wires and syndicated network feeds (plus talk, much of which is nationally syndicated itself). The exception is in a few major cities where a few all-news commercial radio stations still survive, such as Washington, D.C., where WTOP is a significant all-news operation.

Sector Sample

The Index includes three different areas of radio news programming.

1. Public radio: The Index includes 30 minutes of a public radio's broadcast of National Public Radio's (NPR) morning program, Morning Edition or its broadcast of NPR's 4 - 6 pm program, All Things Considered.

NPR produces two hours of Morning Edition, and two hours of All Things Considered each day, which also include multiple news roundups produced by a different unit of NPR. Member stations may pick any segments within those two hours and mix and match as fits their programming interests. Thus, what airs on a member station is considered a "co-production" of NPR and that member station rather than programming coming directly from NPR.

In order to account for this unique relationship, in addition to rotating coding for each show, PEJ also rotates between coding the first 30 minutes of the first hour and the first 30 minutes of the second hour of the member station that we record the show from, WFYI. This gives us a closer representation of the overall content of Morning Edition and All Things Considered for a week.

NPR coding for a sample week may look like this:

Radio Show	Monday	Tuesday	Wednesday	Thursday	Friday
Morning Edition (5 -5:30 am)	X				X
Morning Edition (6 - 6:30 am)		X			
All Things Considered (4 - 4:30 pm)			X		
All Things Considered (5 - 5:30 pm)				X	

In this case, we would begin the next week by coding the 6 - 6:30 broadcast of Morning Edition.

2. Talk Radio: The Index includes some of the most popular national talk shows that are public affairs or news oriented. Since the larger portion of the talk radio audience, and talk radio hosts, are of a conservative political persuasion, we included more conservative hosts than liberal hosts. We code the first 30 minutes of each selected show.

The two most popular conservative radio talk shows are Rush Limbaugh and Sean Hannity. We code each of these shows every other day. Since the politically liberal audience for talk radio is much smaller, we only code one liberal talk show every other day: Ed Schultz, who is the top liberal radio host based on national audience numbers. The Arbitron ratings, according to Talker's Magazine online, for spring 2006 are as follows:

Minimum Weekly Cume (in millions, rounded to the nearest .25, based on Spring '06 Arbitron reports) [\[9\]](#)

Rush Limbaugh (13.5)
Sean Hannity (12.5)
Michael Savage (8.25)
Ed Schultz (2.25)
Randi Rhodes (1.25)
Alan Colmes (1.25)

3. Headline Feeds: Hourly news feeds from national radio organizations like CBS and CNN appear on local stations across the country. These feeds usually last approximately 5 minutes at the top of each hour, and are national in the sense that people all over the country get the same information. They frequently supplement local talk and news shows.

To get a representation of these feeds, we code two national feeds, one from ABC radio, and the other from CBS radio. Each network airs two feeds a day (9 am and 5 pm Eastern time). We code one the 9 am broadcast from one network and the 5 pm feed from the other network every weekday.

The stations used to capture each program are selected based on the availability of a solid feed through the station's web site. We have also compared their shows to that of other stations to ensure that the same edition is aired on that station as on other stations carrying the same program.

This results in the following sample:

News: 30 minutes of NPR each day, rotating between Morning Edition and All Things Considered, as broadcast on a selected member station.

Talk: The first 30 minutes of either one or two talk programs each day. Every weekday, we code one conservative talker: either Rush Limbaugh or Sean Hannity. We code one liberal talk show every other day (Ed Schultz).

Headlines: Two headline segments each day (one from ABC Radio and one from CBS Radio), about 10 minutes total.

This results in a sample of roughly 1 or 2 hours of programming a day.

Universe of Outlets

Each day, then, the NCI includes approximately 8 hrs of broadcast (television and radio), 5 or 6 newspapers (approximately 20 stories), and 6 news web sites (30 stories).

The current universe is as follows:

Newspapers (Eleven in all, Sun-Fri)

Code 2 out of these 4 every day

The New York Times
Los Angeles Times
USA Today
Wall Street Journal

Code 2 out of these 4 every day

The Washington Post
The Columbus Dispatch
Tampa Tribune
Seattle Times

Code 1 or 2 out of these 3 every day

The Day (CT)
Rome News Tribune (GA)
Ventura News (CA)

Web sites (Code 6 of 12 each day, Mon-Fri)

Yahoo News
MSNBC.com
CNN.com
NYTimes.com
Google News
AOL News
FoxNews.com
USAToday.com
WashingtonPost.com

ABCNews.com
HuffingtonPost.com
Wall Street Journal Online

Network TV (Seven in all, Mon-Fri)

Morning shows - code 2 out of 3 every day

ABC - Good Morning America
CBS - Early Show
NBC - Today

Evening news

Code 2 out of 3

ABC - World News Tonight
CBS - CBS Evening News
NBC - NBC Nightly News

Code two consecutive days, then skip one

PBS - Newshour with Jim Lehrer

Cable TV (Fifteen in all, Mon-Fri)

Daytime (2-2:30 pm) - code 2 out of 3 every day

CNN
Fox News
MSNBC

Nighttime CNN - code 2 out of the 4 every day

Situation Room (6 pm)
CNN Tonight/ John King, USA
CNN Prime Time/ Parker Spitzer
Anderson Cooper 360

Nighttime Fox News - code 2 out of the 4 every day

Special Report w/ Bret Baier
Fox Report w/ Shepard Smith
O'Reilly Factor
Hannity

Nighttime MSNBC - code 2 out of the 4 every day

The Ed Show
Hardball (7 pm)

Countdown w/ Keith Olbermann
The Rachael Maddow Show

Radio (Eight in all, Mon-Fri)

News and Headlines - every day
ABC Radio headlines at 9am or 5pm
CBS Radio headlines at 9am or 5pm

NPR - code 1 of the two every day

Morning Edition
All Things Considered

Talk Radio
Rotate between:

Rush Limbaugh
Sean Hannity

Code every other day
Ed Schultz

That brings us to 27 - 30 outlets each weekday. (Between 5 and 6 newspapers only for Sunday)

Universe Procurement and Story Inclusion

Newspapers

For each of the seven newspapers included in our sample, we code all stories where the beginning of the text of the story appears on the front page of that day's hard copy edition. If a story only has a picture, caption, or teaser to text inside the edition, we do not include that story in our sample. We code all stories that appear on the front page with a national or international focus. Because we are looking at the coverage of national and international news, if a story is about an event that is solely local to the paper's point of origin, we exclude such a story from our sample. The only exception to this rule is when a story with a local focus is tied to a story that we have determined to be a "Big Story" - defined as one that has been covered in multiple national-news outlets for more than one news cycle. For example, a story about a local soldier who has come back from the Iraq War has a local angle but is related to a national issue and is important in the context of our study.

We code the entirety of the text of all the articles we include. If an article includes a jump to an inside page in the hard copy edition, we code all the text including that which makes up the jump.

When possible, we have subscribed to the hard copies of the selected newspapers and have them delivered to our Washington, D.C. office. This is possible for national papers that have same-day delivery methods (the New York Times, the Washington Post, the Wall Street Journal and USA Today). For these papers, we use the hard copy edition to determine the placement on the front page of the edition, and to get all the text we will code. We use the LexisNexis computer database to determine the word count for each of the stories.

For all of the other papers that we are not able to get hard copies of within the same day of publication, we take advantage of internet resources that have digital copies of the hard copy editions. Pressdisplay.com and Newsstand.com have subscription services offering same-day access digital versions of the hard copy. From these digital versions, we obtain the text of the relevant articles and also determine the word counts. To get the word counts, we copy the text of the articles (not including captions, titles, bylines, or pull-out quotes) into the Microsoft Word software program and run the "word count" function to get the final number. When necessary, we go to the paper's web site in order to find the text of articles that are not available on either of the two web services. Through examination of each individual article, we are able to determine when the text of the article on the web site is the same as it would be in the hard copy of the paper.

Network and Cable Television

For all television programs, we code the first 30 minutes of the broadcast (with the exception of PBS Newshour), regardless of how long the program lasts. As with newspapers, we code all stories that are news reports that relate to a national or international issue. Therefore, we do not code stories that are part of a local insert into a national show. For example, each half-hour, NBC's Today Show cuts to a local affiliate which will report local stories and local weather. We do not include those local insert stories.

We also exclude from our sample commercials, promos, and teasers of upcoming stories. We are only interested in the actual reporting that takes place during the broadcasts.

Any story that fits the above criteria and begins within the first 30 minutes is included in the study, even if the story finishes outside of the 30 minute time period. A three-minute story that begins 28 minutes into a program would be coded in its entirety, even though the final minute ran after our 30-minute cutoff mark. The exception to this rule is when a television station is showing a speech or press conference that runs longer than the 30-minute period (often much longer). In those cases, we cut off the coding at the 30-minute mark in order to prevent that event from unduly impacting our overall data.

The method of collection of all television programs is the same. PEJ subscribes to DirectTV satellite television service. We use a software program called Snapstream to capture the shows we code. The Snapstream service digitally records each broadcast on an in-house server and then we archive the programs onto DVDs. There is redundancy in our recording method so that each show is also recorded on TiVo recording boxes that are directly linked to DirectTV. We do this in order to avoid problems in our capture that might result from technical error.

Occasionally outlets deviate from the regularly scheduled news programs. When a show is pre-empted for a special *live* event, such as a presidential campaign debate or the State of the Union address, we do not include that period as part of our sample.

Radio

The rules for capturing and selecting stories to code for radio are very similar to television. We code the first 30 minutes of each show regardless of how long the show lasts. We also exclude local inserts from local affiliates, and continue coding any story that runs past the 30-minute mark.

For each of the radio shows selected, we have found national feeds of the show that are available on the web. As with television, we have two computers capturing each show so as to avoid errors if one feed is not working. The actual recording is done using a software program called Replay A/V which captures the digital feeds and creates digital copies of the programs onto our computers. We then archive those programs onto DVDs.

Online

For each of the web sites we are including in our sample, we capture and code the top 5 stories that appear on the site at the time of capture. Our capture times rotate on a regular basis. They occur either between 9 and 10 am Eastern time or between 4 and 5 pm Eastern time each weekday. The captures physically occur with a coder going to each site using an internet browser and saving the home page and appropriate article pages to our computers, exactly as they appear in our browsers at the time of the capture. We rely on people rather than a software package to capture sites because some software packages have proved invasive to Web sites.

With the current rotation of Web sites along with the rotation of the times of day that we capture the sites, we wanted to make sure that we did not always capture the same sites at the same time (CNN.com always at 9 am, for example). We also wanted to assure that for the Web sites where we coded another outlet from the same news organization, such as USA Today's newspaper and the usatoday.com Web site, we did not code both of those outlets on the same days. In order to avoid these two concerns, we created a method of rotation where the capturing times for the Web site rotate every two days.

This means that the pattern the capture times follow is 9 am, 9 am, 4 pm, 4 pm, 9 am etc.

Here is an example of how the online rotation works:

	<u>Week 1</u>					<u>Week 2</u>					<u>Week 3</u>					<u>Week 4</u>				
	4 pm	9 am	9	4	4	9	9	4	4	9	9	4	4	9	9	4	4	9	9	4
CNN.com		x		x		x		x		x		x		x		x		x		x
Yahoo News		x		x		x		x		x		x		x		x		x		x
MSNBC.com		x		x		x		x		x		x		x		x		x		x
Wall Street Journal Online		x		x		x		x		x		x		x		x		x		x
Foxnews.com		x		x		x		x		x		x		x		x		x		x
USAToday.com		x		x		x		x		x		x		x		x		x		x
NYTimes.com	x		x		x		x		x		x		x		x		x		x	
AOL News	x		x		x		x		x		x		x		x		x		x	
WashingtonPost.com	x		x		x		x		x		x		x		x		x		x	
ABCNews.com	x		x		x		x		x		x		x		x		x		x	
HuffingtonPost.com	x		x		x		x		x		x		x		x		x		x	
Google News	x		x		x		x		x		x		x		x		x		x	

As with newspapers, some stories are longer than one web page. In those cases, we include the entire text of the article for as many web pages as the article lasts.

Because each web site is formatted differently, we came up with a standard set of rules to determine which stories are the most prominent on a given home page. We spent a significant amount of time examining various popular news sites and discovered patterns which led us to the best possible rules. First, we ignore all advertisements and extra features on the sites that are not reported news stories. We are only interested in the main channels of the web sites where the lead stories of the day are displayed. Second, we determine the top "lead" story. That is the story with the largest font size for its title on the home page. The second most prominent story is the story that has a picture associated with it, if that story is different than the story with the next largest title. By considering many sites, we realized that a number of sites associate pictures with stories that they find particularly interesting, but are clearly not intended to be the most important story of the day. However, we do want those stories to be in our sample because the reader's eye will be drawn to them.

Having figured out the first and second most prominent stories, we then rely on two factors to determine the next three most prominent stories. We first consider the size of the headline text and then the height on the home page. Therefore, for determining the third most prominent story, we look for the story with the largest title font after the top two most prominent stories. If there are several stories with identical font sizes, we determine that the story that is higher up on the page is more prominent. In cases where two articles have the same font size and the same height on the screen, we choose the article to the left to be the more prominent.

For the first two years of the NCI, we did not include online news stories that were audio or video features. Starting in 2009, PEJ changed its method of measuring online stories to allow for the inclusion of audio and video stories. See the section below entitled "Inclusion of Online Audio and Video in Index Calculations" for details on how the changes to the Index statistics have been incorporated.

Coding Procedures and Intercoder Reliability

A coding protocol was designed for this project based on PEJ's previous related studies. Seventeen variables are coded, including coder ID, date coded, story ID number (these three are generated from the coding software automatically), story date, source, broadcast start time, broadcast story start timecode, story word count, placement/prominence, story format, story describer, big story, sub-storyline, geographic focus, broad story topic, lead newsmaker, and broadcast story ending timecode.

The source variable includes all the media outlets we code. The variable for broadcast start time applies to radio and TV broadcast news and gives the starting time of the program in which the story appears. Broadcast story start timecode is the time at which a story begins after the start of the show, while broadcast story ending timecode is the time at which a story ends. The variable for story word count designates the word count of each individual print/online news story. The placement/prominence variable designates where stories are located within a publication, on a website, or within a broadcast. The location reflects the prominence given the stories by the journalists creating and editing the content. Story format measures the type and origin of the text-based and broadcast stories, which designates, at a basic level, whether the news story is a product of original reporting, or drawn from another news source. Story describer is a short description of the content of each story. Big stories are particular topics that occurred often in news media during the time period under study. Sub-storyline applies to stories that fit into some of the long-running big stories, reflecting specific aspects, features or narrower elements of some big stories. The variable for geographic focus refers to the geographic area to which the topic is relevant in relation to the location of the news source. The variable for the broad story topic identifies which of the type of broad topic categories is addressed by a story. The variable for lead newsmaker names the person or group who is the central focus of the story.

The daily coding operation is directed by a coding manager, a training coordinator, a methodologist, and a senior researcher.

In 2010, the coding team responsible for performing the content analysis was made up of twenty individuals. Several of the coders have been trained extensively since the summer of 2006 and most of the coders have more than a year's worth of coding experience.

Numerous tests of intercoder reliability have been conducted since the inception of the NCI in order to ensure accuracy among all coders.

2010 Intercoder Tests

PEJ conducted one major intercoder test over the course of 2010. This test focused on the complex variables that require extensive training, expertise and effective communication among the coders. These are the Main Variables.

Main Variables

Over the course of Spring and Summer of 2010, we tested intercoder agreement for Main Variables. One hundred and five stories (representing 10% of a week's sample) were randomly selected from across all 5 media sectors. Coders were asked to recode 12 stories from the newspaper sector; 15 from online; 20 from network; 34 from cable and 24 from the radio sector.

In this round of testing multiple coders were asked to recode the same stories. A total of 17 coders participated in this test. The number of stories recoded by a coder depended on the amount of coding executed by the coder in the previous six weeks.

The percent of agreement was as follows:

Format: 83%

Big Story: 85%

Substory: 82%

Geographic Focus: 91%

Broad Topic: 78%

Lead Newsmaker: 84%

Lead Newsmaker 2: 95%

Select Housekeeping Variables

In addition to these Main Variables, we also tested certain Housekeeping variables in the same round of testing. Housekeeping variables are those that are necessary for each story, but require little or no inference from each other.

Placement: 97%

Broadcast Only Variable

Broadcast Start Time: 97%

2009 - Early 2010 Intercoder Tests

In 2009, PEJ conducted two phases of major intercoder testing to ensure continuing accuracy among all coders.

The first phase tested for variables that require little to no subjectivity from the coder. We refer to these codes as Housekeeping Variables. The second phase of testing was conducted in the fall of 2009. In this phase we tested for variables that are more complex and require more training and expertise. We call these the Main Variables.

Housekeeping Variables

In summer 2009, we tested intercoder agreement for Housekeeping variables. These are variables that are necessary for each story but involve little inference from each coder.

We used a random sample of 131 stories, representing all five media sectors that we code. This sample represented 10% of the number of the stories we code in an average week.

A total of 15 coders participated in the study. Each coder was asked to recode each of the 131 stories.

A total of 27 print (12 newspaper, 15 online) and 104 broadcast (44 network, 36 cable and 24 radio) stories were sampled.

The percent of agreement was as follows:

Story Date: 99%

Source: 97%

Placement: 94%

Print Only Variable:

Story Word Count (+/- 20 words): 84%

Broadcast Only Variables:

Broadcast Start Time: 98%

Story Start Time (+/- 6 seconds): 97%

Story End Time (+/- 6 seconds): 91%

Main Variables

The second group of variables we tested was referred to as the main variables, and they involve more training and interpretation. Having already demonstrated that we had a high level of agreement for all of our housekeeping variables, we then had the coders participate in separate tests for these main variables.

In the fall of 2009, we conducted intercoder testing for main variables. One hundred and thirty stories coded were randomly selected from all five media sectors (20 newspaper articles, 10 online stories, 36 network stories, 41 cable stories and 23 radio stories). These stories were coded over the course of 10 weeks.

A total of 16 coders participated in this test.

For main variables, we achieved the following level of agreement:

Format: 86%

Big Story: 85%

Substoryline: 83%

Geographic Focus: 89%

Lead Newsmaker: 86%

Lead Newsmaker 2: 90%

For our most complicated variable, Broad Story Topic, we conducted multiple tests in mid to late 2009 and early 2010. The average agreement for Broad Story Topic was 81%.

Testing Details

All the percentages of agreement for the above variables were calculated using a software program available online called PRAM.[\[10\]](#)

Since the inception of the News Coverage Index, as new coders were hired and included in the coding team, they were given extensive training from both the training coordinator, content supervisor, and other experienced coders. New coders were not allowed to participate in the weekly coding for the project until they had demonstrated a level of agreement with experienced coders for all variables at an 80% level or higher.

Each coder works between 20 and 37.5 hours a week in our Washington D.C. office and was trained to work on all the print and broadcast mediums included in the sample. The schedule for each coder varies, but since all of the material included in the Index is archived, the actual coding can be performed at any point during the week.

To achieve diversity in the coding and ensure statistical reliability, generally no one coder codes more than 50% of a particular media sector within one week. Each coder codes at least three mediums each week. In the case of difficult coding decisions about a particular story, the final decision is made by either the coding administrator or a senior member of the PEJ staff.

The physical coding data is entered into a proprietary software program that has been written for this project by [Phase II Technology](#). The software allows coders to enter the data for each variable, and also allows coders to review their work and correct mistakes when needed. The same software package compiles all of the coding data each week and allows us to perform the necessary statistical tests.

Total Media Combined: Creation and Weighting

The basis of measurement used to determine top stories in broadcast and cable is time, and in text-based media, it is words. Thus for cable news, for example, we refer to the percent of total seconds that a certain story received. In other words, of all the seconds analyzed in cable news this week, ground events in Iraq accounted for xx% (or xx seconds out of a total of xxx). The industry term for this is "newshole"-the space given to news content.

The main Index considers broadcast and print together, identifying the top stories across all media. To do this, words and seconds are merged together to become total newshole. After considering the various options for merging the two, the most straightforward and sensible method was to first generate the percent of newshole for each specific medium. This way all media are represented in the same measurement-percent.

Next, we needed to create a method for merging the various percentages. There were several options. Should we run a simple average of all five? Should we average all print and all broadcast and then average those two? Or, should we apply some kind of weight based on apparent audience?

Because each medium measures its audience differently (ratings per month in television, weekly circulation in newspapers, unique visitors in online), any system based on audience figures raises serious issues of discontinuity. Nonetheless, several of our advisors thought some kind of weight should be applied. Various options were considered, including a combination of different metrics, such as audience data alongside supplemental survey data. One consistent measure is that of public opinion surveys. The same question is posed about multiple media. Two such questions are asked regularly by the Pew Research Center for the People & the Press. One asks about "regular usage" and the other asks where people go for "national and international news."

Before arriving at a method for the launch of the Index in January 2007, we tested multiple models:

Model 1: compile percentages for big stories for each of the five media sectors (newspapers, online sites, network TV, cable TV and radio), and then average those five lists into one final list.

Model 2: Divide the media sectors into two groups, text-based media (newspapers, online sites) and broadcast (network TV, cable TV, and radio). Average the lists of percentages between the two groups to get one final list.

Model 3: compile percentages for big stories for each of the five media sectors, and then add the weighted five lists together into one final list. The weights given to each media sector were calculated by averaging the three most recent survey data in terms of where people get news about national and international issues, collected by the Pew Research Center for the People & the Press (June 2005, November 2005, and August 2006). First, we take the average response for each media sector across the three time periods. Next, we rebalance the average percents to match the five media sectors in the Index—newspapers, online, network TV, cable TV, and radio—to equal 100%. Thus, the weight for newspapers would be 0.28, online would be 0.16, network TV would be 0.18, cable TV would be 0.26, and radio would be 0.12.

Model 4: compile percentages for big stories for each of the five media sectors, and then add the weighted five lists together into one final list. The weights assigned to each media sector were generated based on the regular media usage survey data, collected by the Pew Research Center for the People & the Press in their Biennial Media Consumption Survey 2006. Thus, the weight for newspapers would be 0.307, online would be 0.218, network TV would be 0.165, cable TV would be 0.201 and radio would be 0.109.

By testing two trial weeks' data, we found that the lists of top five stories were exactly the same (top stories' names and their ranks) using all four of these models, although some percentages varied. In the end, the academic and survey analysts on our team felt the best option was model 3. It has the virtue of tracking the media use for national and international news which is what the Index studies. Also, the Pew Research Center for the People & the Press asks this question about once every six months so we can reflect changes in media use. We adopted this model and plan to update the weights when appropriate.

Note: The weights used for data in Model 3 have been updated three times since the inception of the News Coverage Index.

2010 weights were calculated by averaging survey data from July and December 2009. These data were collected by the Pew Research Center for the People & the Press. The weights currently in use for the Index in 2010 are:

2010 Weights

Newspapers: 0.22

Online: 0.26

Network: 0.15

Cable: 0.24

Radio: 0.13

Inclusion of Online Audio and Video in Index Calculations

The decision to include audio and video stories for Online beginning in 2009 has meant that PEJ needed to create a method to incorporate different ways of measuring length (time in seconds versus amount of words) within the same media sector.

Prior to this change, the calculations for the percent of newshole for the online sector have been the percent of words (in text). By now including multi-media elements in our Web sample, this created a challenge for coming up with a percent of newshole calculation for that can incorporate both text and the length of multi-media stories in seconds. PEJ undertook several tests to come up with a simple, yet accurate method of creating an equivalent measure.

The process PEJ uses for valuing multi-media stories is to take the length of the multi-media story (in seconds) and multiply by 4 to get an approximate equivalent value to a text story of that number of words. For example, an online video that is 30 seconds in length would be given an equivalent value of 120 in words ($30 * 4$). An online video that is 60 seconds in length would be given an equivalent value of 240 words ($60 * 4$).

PEJ arrived at this method by first timing how long it takes for people to read news stories out loud. After having five people timed reading different types of news stories, we discovered that people read approximately 3 words per second. However, simply multiplying the length of a story by 3 would not accurately reflect the value of a multi-media story.

We then compared the distribution of length of online text stories to the distribution of length of online multi-media stories. To make this comparison, we took the distribution of 3500 online text stories over the last 6 months in the NCI and compared that to the distribution of length for 280 video stories compiled from 7 separate Web news sites. The median length of a text story was approximately 600 words while the median length of a video story was approximately 150 seconds.

Drawing from these comparisons, we determined that multiplying the length by 4 gave a reasonable approximate value to use in comparison to the length of a text story in words. No single multiplier would match exactly since the distribution of the length of Web videos is not linear, and because there is no simple way to quantify the value of visuals within multi-media stories along with the text. However, this simple method of multiplication gives us a straightforward way to make approximate equivalents between two measures (seconds and words) that are not otherwise easily compared.

Updated January 10 , 2010

FOOTNOTES:

[1] Data from 2009 indicate that the three evening networks reach about 22 million viewers and the three morning newscasts average about 12.7 million people daily. Nielsen Media Research, used under license.

[2] Data from 2009 indicate that 3.9 million viewers watch cable news during prime time hours and over 2 million watch during daytime hours. PEJ Analysis of Nielsen Media Research, used under license.

[3] For August 2009, the O'Reilly Factor averaged 3.4 million viewers a night while Larry King Live averaged 1.1 million viewers. Nielsen Media Research on Media Bistro.com.

[4] For 2008, circulation numbers indicate that 48 million people buy a newspaper each weekday. 2008 Editor and Publisher Yearbook Data.

[5] According to the December 2007 survey, 27% of adults go online for news each day. Pew Internet, December 2007 survey. A more recent survey shows that 37% of Americans regularly go online for news. Pew Research Center for the People and the Press, July 2008 survey.

[6] Pew Internet, 2005 survey.

[7] By the spring of 2007, 93% of the population 12 and older listened to radio on a weekly basis. [Arbitron ratings, Spring 2007](#). Radio reaches 235 million Americans over the course of a week. Arbitron ratings, March 2008.

[8] Arbitron, "Radio Today: How Americans Listen to Radio, 2007 Edition," April 13, 2007. March 2008 data shows that News/ Talk is the top or second top category of listening in every region of the country except one, ranking it as the most popular of all measured radio formats. Arbitron ratings, March 2008.

[9] Current ratings data available at Talkers Magazine [online](#).

[10] See "The Content Analysis Guidebook," by Kimberly A. Neuendorf, Sage Publications, 2002.

NEW MEDIA INDEX

The New Media Index is a weekly report that captures the leading commentary of blogs and social media sites focused on news and compares those subjects to coverage in the mainstream press.

The New Media Index is a companion to PEJ's weekly **News Coverage Index**. Blogs and other new media are an important part of creating today's news information narrative and in shaping the way Americans interact with the news. The expansion of online blogs and other social media sites has allowed news-consumers and others outside the mainstream press to have more of a role in agenda setting, dissemination and interpretation. PEJ aims to find out what subjects in the national news the online sites focus on, and how that compares with the narrative in the traditional press.

Universe

Two prominent Web tracking sites, **Icerocket** and **Tweetmeme**, monitor millions of blogs, Tweets, and other pieces of social media, using the links to articles embedded on these sites as a proxy for determining what these subjects are.

Each of these two sites offers lists of the most linked-to news stories, based on the number of blogs, tweets, or other sites that link to them. These stories are almost always from mainstream news sources such as the New York Times Web site or CNN.com.

Each weekday between 9 and 10 am ET, a PEJ staff member manually captures these lists.

From those lists, the top five linked-to articles are captured for further analysis by PEJ staff. That results in ten articles a day and 50 articles a week.

Coding Procedures

Once the lists of articles are compiled, PEJ staff conducts a content analysis of the subject matter of these mainstream news articles in a similar manner to the **News Coverage Index**.

Almost all of the codes and rules are the same as with the NCI. For details about the variables and PEJ's intercoder testing procedures for these codes, refer to the detailed methodology about the **News Coverage Index**.

Among the variables is what we call Big Story, a variable also used in the NCI, which is a continuously evolving list of the major news storylines of the day. The Big Story variable is the primary one used to see which topics are getting the most attention. A story must be 50% about a particular topic in order to be given a specific Big Story code.

The only additional variables used in the NMI are identifying the original outlet of the news story and tracking the number of links aimed at each story included in the sample. Technorati and Icerocket provide the number of links within their lists.

Calculations

The priorities of the bloggers are measured in terms of percentage of links. Each time a news blog or social media Web page adds a link to its site directing its readers to a news story, it suggests that the author places at least some importance on the content of that article. The user may or may not agree with the contents of the article, but they feel it is important enough to draw the reader's attention to it.

The calculations for the NCI have a different base. That Index uses the percent of time (in seconds) or space (in words) to arrive at the percent of newshole devoted to each topic.

The reason that the New Media Index uses a different measure, links rather than newshole, is because the nature of online media is different from other traditional forms of media. First, there is no limit to the amount of space that can be devoted to a specific story. In a newspaper, there is a limited amount of space on a front page, for example, and a television newscast is limited by its allotted amount of time. Web sites have no such limits.

Second, PEJ determined that in this procedure, the number of blogs that link to a news article are a far greater measure of the significance of that article online than the length of the story. A particular article might be quite long in terms of number of words, but if only a few blogs link to it, that article would have only a small influence in the new media environment. A short story that gets linked to many times has a far greater influence.

The percent of links for each Big Story is determined by taking the total number of links in the sample (from all 50 stories in the week) and then dividing that number by the number of links devoted to each specific Big Story. The percentages are then ranked in order to discover the five storylines that were most present in online commentary.

Differences from the NCI

In addition to the base calculation, there are three differences between the NMI and the NCI to note:

While the capture times for the Web sites included in the News Coverage Index rotate each day, a decision was made to keep the times the same for the New Media Index. The reasoning is that since these lists compile the number of links to stories over a 48-hour window, rotating the time of capture would result in different increments of times between each capture. Through testing, PEJ has discovered that the stories on the lists change significantly more over a 24-hour period than they do over a 12 or 16-hour period.

While the News Coverage Index is comprised of primarily U.S.-based media outlets, the aggregators of blogs and other social media include both U.S. and non-U.S. blogs. In addition, stories that are linked to can be from non-U.S. sources. However, according to PEJ's research over the last two months, the only non-U.S. news stories included in the top lists for Technorati and Icerocket have been the BBC (whose Web site is part of the News Coverage Index) and the Guardian.

PEJ's weekly **News Coverage Index** includes Sunday newspapers while the New Media Index is Monday through Friday.

YouTube Videos

The New Media Index also includes a section of the most popular news video on YouTube each week.

Each Friday at noon ET, a PEJ staff member captures the **list of most viewed news** and politics videos on YouTube over the previous week. These videos are categorized as such on the YouTube site and are often a mix of mainstream news reports, raw footage relating to breaking events, or other types of public affairs clips. PEJ determines the top five most viewed videos as they are listed on YouTube's page at the time of capture.

Note: After consulting various reference guides and outside consultants on usage, the Project has chosen to refer to its several weekly content analysis reports as “indexes”—the version largely accepted in journalism—instead of “indices”—a term used more frequently in scientific or academic writing.

News Interest Index

The *News Interest Index* is a weekly survey conducted by the Pew Research Center for the People & the Press aimed at gauging the public's interest in and reaction to major news events. This project has been undertaken in conjunction with the Project for Excellence in Journalism's *News Coverage Index*, an ongoing content analysis of the news. The News Coverage Index catalogues the news from top news organizations across five major sectors of the media: newspapers, network television, cable television, radio and the internet. Each week (from Monday through Sunday) PEJ compiles this data to identify the top stories for the week. (For more information about the Project for Excellence in Journalism's News Coverage Index, go to www.journalism.org.) The News Interest Index survey collects data from Thursday through Sunday to gauge public interest in the most covered stories of the week.

Results for the weekly surveys are based on telephone interviews conducted under the direction of Princeton Survey Research Associates International among a national sample of approximately 1,000 adults living in the continental United States, 18 years of age or older. For results based on the total sample, one can say with 95% confidence that the error attributable to sampling is plus or minus 4 percentage points.

The combined landline and cell phone sample are weighted using an iterative technique that matches gender, age, education, race, Hispanic origin, region, and population density to parameters from the March 2009 Census Bureau's Current Population Survey. The sample is also weighted to match current patterns of telephone status based on extrapolations from the 2009 National Health Interview Survey. The weighting procedure also accounts for the fact that respondents with both landline and cell phones have a greater probability of being included in the combined sample and adjusts for household size within the landline sample. Sampling errors and statistical tests of significance take into account the effect of weighting.

In addition to sampling error, one should bear in mind that question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of opinion polls.