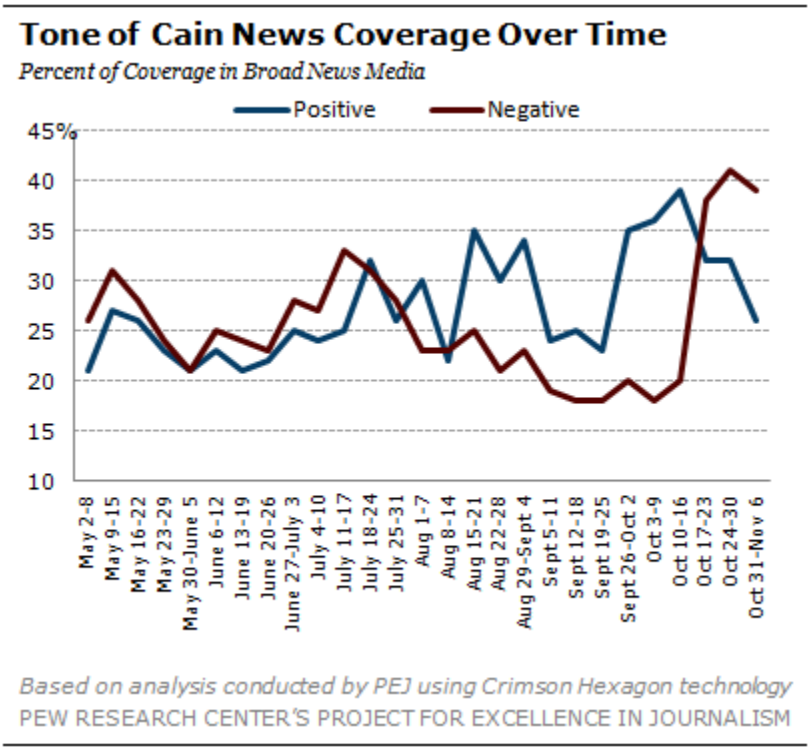
Cain's Bad Stretch—A Campaign Coverage Update:

How Elite Media and Press Overall Compare

While his support continued to hold in the polls, businessman and GOP presidential candidate Herman Cain was the focus of a much tougher narrative in the news media last week, according to an analysis by the Pew Research Center's Project for Excellence in Journalism. The week, October 31 through November 6, was also the third consecutive one in which negative assertions of Cain in the press outnumbered positive, a turn in his narrative that predated allegations of sexual harassment.

Last week, with the news media focused on reports that the trade association Cain once ran compensated women who had alleged he had sexually harassed them, 39% percent of the statements about Cain across a broad spectrum of news media were negative, while 26% were positive (and 36% were neutral).



That 13 point differential between negative and positive coverage is the worst week Cain has undergone in press coverage so far, looking across six months of coverage in some 11,500 news outlets, a sample that represents the bulk of what Americans see in the media.

But the tone of the narrative had already begun to turn negative for Cain the week of October 17-23, as he began to see more media scrutiny of his 9-9-9 tax plan and other aspects of his positions while he gained in the polls.

The new PEJ study, which updates an earlier October 17 report on the tone of campaign news coverage, also compares two samples of news media with each other—that broad spectrum of news media and an "elite" media sample of 47 outlets that are among the largest and most popular.

That comparison shows the subsample of "elite" media news outlets were even tougher on Cain last week. In that smaller group of outlets, negative statements outnumbered positive toward Cain by 31 points (19% positive, 50% negative and 32% neutral).

The comparison between the broad spectrum of news outlets and the smaller group of elite outlets also reveals some interesting distinctions between the samples more generally. The elite media tend to move faster when there is a shift in the media narrative. The broader spectrum of media sometimes shifts a week or two later, and when it does so the differentials can be even more pronounced than in the elite media.

Over time, however, the tone of coverage in the broad spectrum of news outlets and the tone in the elite subsample tend to merge and look very similar.

For instance, overall across the broad range of news outlets, 25% of the statements about Mitt Romney were positive, while 28% were negative and 47% neutral from May 2 to November 6. In the smaller media sample, the numbers were 27%, 29% and 45%.

Coverage of Obama was similarly close. In the broad spectrum of outlets, 9% of statements were positive, 35% negative and 56% neutral. The elite sample was 9% positive, 36% negative and 55% neutral. Across all candidates, the trend lines between the two samples were essentially the same and the variance in positive assertions averaged three points and negative assertions one point.

These are some of the findings of a new report that combines traditional content analysis methods with computer algorithmic coding using software developed by the company Crimson Hexagon.

This research on the tone in news coverage is not a study of media fairness or bias. Rather, it offers a comprehensive, quantitative analysis of whether the messages Americans receive about a candidate in the news media are positive, negative or neutral. The work examines and quantifies all the assertions about a candidate in news coverage, whether they come from journalists, supporters, opponents, citizens, newsmakers, pundits, polling data and other sources, to understand the overall narrative about that candidate. When a candidate is widely criticized by rivals, for instance, Americans are hearing negative statements about that candidate. When a candidate begins to surge in the polls, and his or her candidacy begins to look more viable, Americans are receiving positive statements about that candidate.

About the Study

The findings are based on research that combines the conventional content analysis research methods conducted by human researchers with algorithmic technology developed by the company Crimson Hexagon. In this combined approach, researchers analyze media content for tone, using PEJ's traditional rules and strict intercoder testing methods to assure reliability and accuracy. Those researchers then train the algorithm until it can replicate the results the researchers arrived at themselves. The power of the computers to code massive quantities of content in ways that replicate the human coding makes it possible for the study to examine what comes close to a census of all the news media offered to Americans via RSS feeds, providing a much deeper and more powerful sense of the media in the U.S. than traditional "sampling" can give. Samples of media offer a useful proxy, but only that. The comparison of the elite sample and the broad spectrum of media provide a sense of how those two cuts of news media compare.

To be assured that the algorithm is accurate and current, researchers "retrain" the algorithm each week with new content, and test that the algorithm continues to produce accurate results.[1]

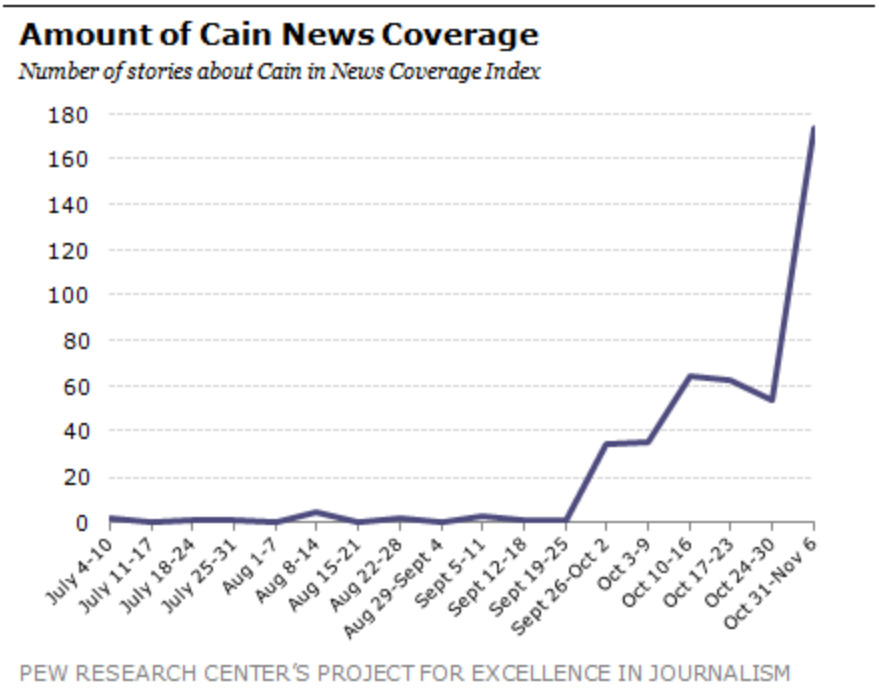
To arrive at the "elite" media sample, PEJ created a list of outlets that mirrored those in the Project's weekly News Coverage Index. That sample involves media from five different sectors of national media—print, cable, broadcast, radio or audio, and online. The list of outlets is derived to include a broad range of outlets representative of the traditional or elite media universe.[2]

A number of people at the Project for Excellence in Journalism worked on this report. Associate Director Mark Jurkowitz and Director Tom Rosenstiel wrote the report. The creation of the monitors using the Crimson Hexagon software was supervised by Tricia Sartor, the manager of the weekly news index, and senior researcher Paul Hitlin. Researchers Kevin Caldwell and Nancy Vogt and content and training coordinator Mahvish Shahid Khan created and ran monitors using the computer technology. Tricia Sartor and researcher Steve Adams produced the charts. Dana Page handled the web and communications for the report.

Cain's Difficult Week

Cain's difficulties in the press had been building for nearly a month, but last week was different for two reasons. The percentage of positive assertions about his candidacy fell. Perhaps just as important, the volume of coverage about him was enormous.

Last week (October 31 through November 6) represented by a substantial margin the high water mark of 2012



[1] Extensive testing by Crimson Hexagon and PEJ has demonstrated that the tool is 97% reliable, that is, in 97% of cases analyzed, the technology's coding has been shown to match human coding. In addition, PEJ conducted examinations of human intercoder reliability to show that the training process for complex concepts is replicable. Those tests came up with results that were within 85% of each other.

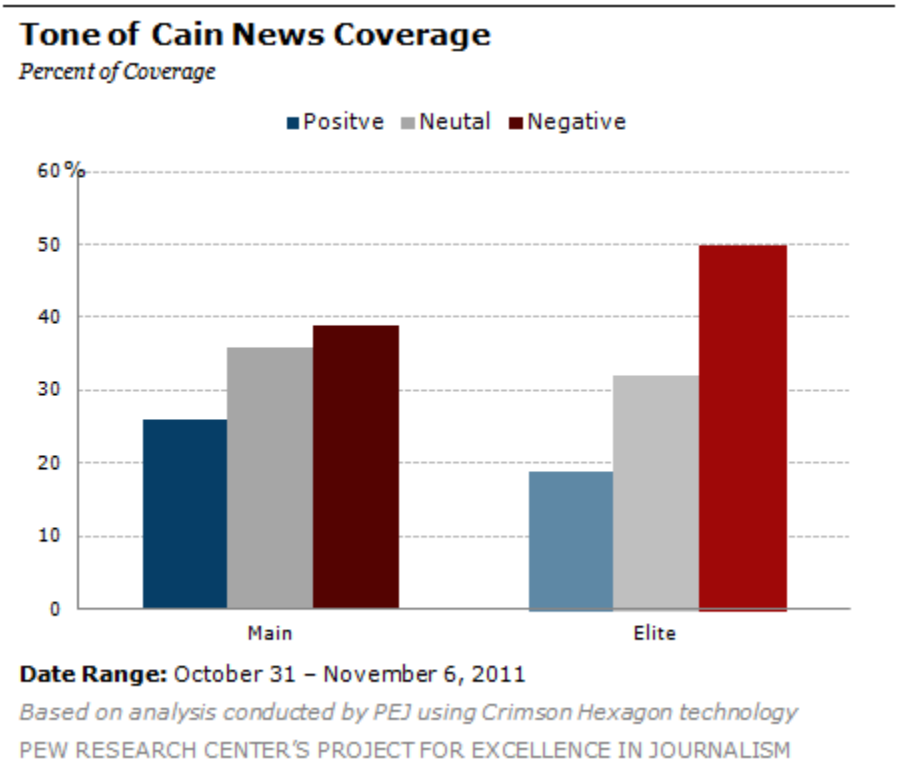
[2] Using Crimson Hexagon's technology, which retrieves media content via RSS feeds, researchers found that five of those outlets could not be included. The radio programs of Rush Limbaugh, Ed Schultz and Sean Hannity did not have RSS feeds, though Hannity's cable web content is coded through the FoxNews feed. Crimson Hexagon's technology could not retrieve data from the Wall Street Journal's RSS feeds because of a paywall, and the content from Google News aggregation of content produced by others was already coded elsewhere in Crimson Hexagon's sample.

presidential coverage thus far. The campaign filled 29% of the newshole studied, according to PEJ's News Coverage Index which monitors the news agenda in the U.S. press each week by examining in real time what topics get what level of coverage.[3] That was roughly a third more attention than the previous high point in coverage, which was 19%, from October 10-16.

Just as important, Cain was overwhelmingly the central figure in that coverage last week. He was a "significant" newsmaker in 77% of the week's campaign stories, and a "dominant" newsmaker in almost three-quarters (72%) of all campaign stories. (To register as a significant newsmaker, a figure must be mentioned in at least 25% of a story; that threshold rises to 50% for a dominant newsmaker).

That represents the single biggest week of coverage for any candidate in the Republican field so far in the 2012 race. Put another way, the media's gaze was focused more intently on Cain last week than it had been so far in any week on any candidate. (By comparison Romney, the second-most covered GOP candidate last week, was a significant newsmaker in 15% of the campaign stories and a dominant newsmaker in 9%.)

As one measure of how widespread news of Cain's troubles became, a new poll by the Pew Research Center for the People & the Press taken late last week found that 75% of Americans said they had heard a lot or a little about the sexual harassment allegations.



The story, initially broken by the political site Politico on October 30, expanded as Cain's campaign had difficulty responding, at one point accusing the campaign of Texas Governor Rick Perry of leaking the news. The story also grew with reports about compensation paid to one of the accusers and then with the number of women allegedly involved growing. (This week, on November 7, another woman, Sharon Bialek, came forward to publicly detail what she said were inappropriate attempts by Cain to secure sexual favors in exchange for his help in finding her a job.)

[3] The list of media outlets included in PEJ's News Coverage Index closely resembles the list of outlets represented in the "elite" media sample.

As the scandal grew, much of the negative coverage of Cain involved speculation about its potential to seriously damage, if not end, what had been a surprisingly successful campaign. Many political analysts and pundits were looking past the rush of developments to handicap the unlikely frontrunner's chances of survival.

By week's end, the result was not so much a jump in negative coverage for Cain across the broad spectrum of media outlets as it was a drop in his positive coverage. For the week, 26% of statements about Cain were positive (down from 32% the week before) and 39% were negative (similar to 41% the week before, which had been his most negative week to date).

The assessment was even more negative in the smaller sample of elite media, where the tone of Cain's coverage last week was 19% positive, 50% negative and 32% neutral.

In these elite outlets, the political calculus of the potential fallout of the accusations—and the competency of Cain in addressing them—were particularly evident in the coverage.

An October 31 story on the CBS News site was headlined, "Can Herman Cain's campaign survive?" warning that "further developments in this story could potentially sink his campaign."

On Nov. 2, the New York Times' Nate Silver blogged that, "I've become convinced that the sexual harassment allegations against Mr. Cain are a real problem" in regard to the viability of his campaign.

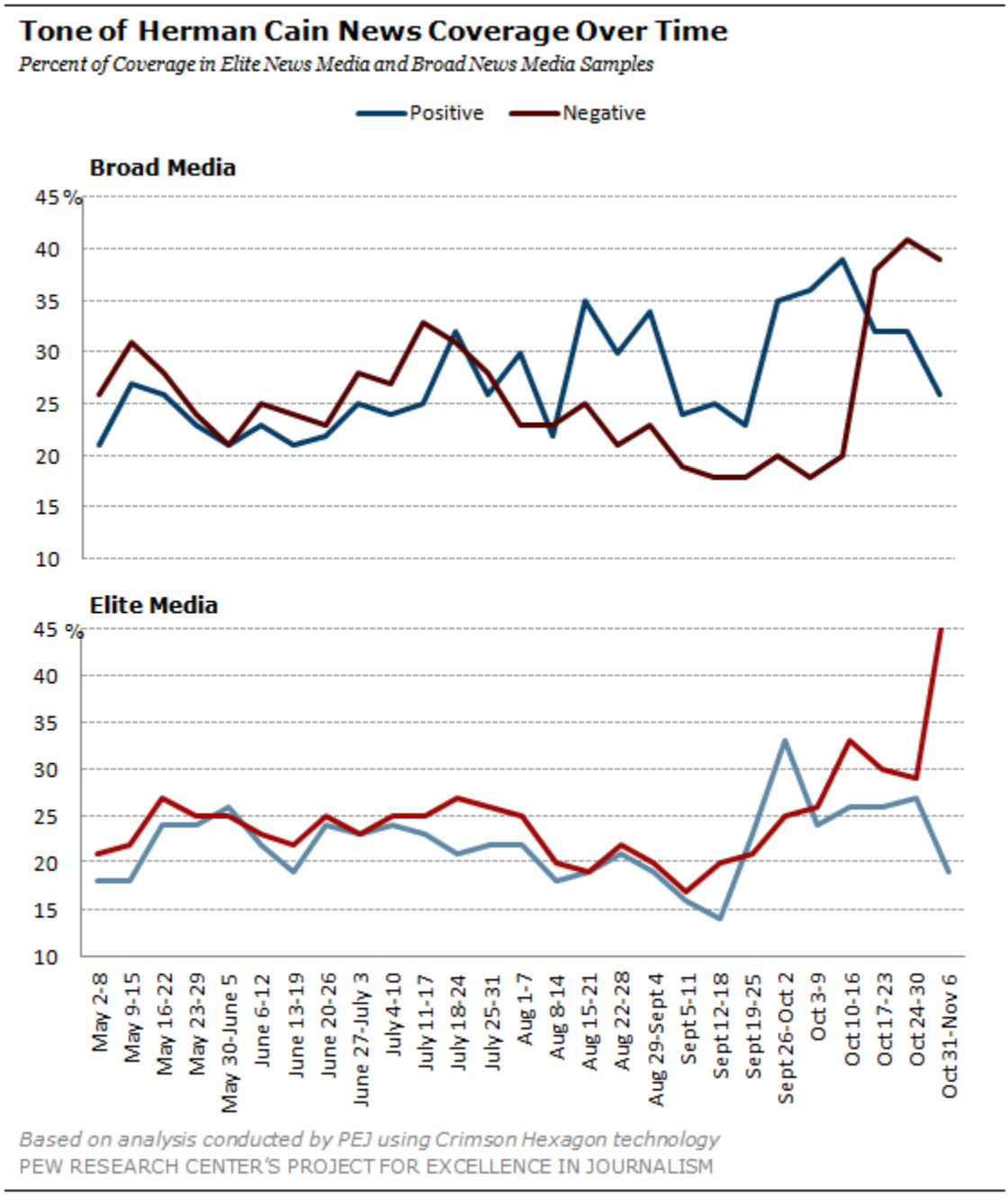
A day later, writing in the Washington Post, conservative commentator Jennifer Rubin went further, asserting that "Republican operatives, pollsters and consultants seem in agreement that Herman Cain is, as a prominent communications guru put it, 'toast.'"

Yet throughout the week there was also was a counter narrative that led to some positive coverage for Cain. Despite the brewing scandal, he was holding firm at the top of GOP presidential polls.

On November 2, the Wall Street Journal reported that "after two days of fending off allegations of sexual harassment dating back a dozen years, GOP presidential hopeful Herman Cain remains popular in the early primary state of South Carolina, a new poll finds. A Rasmussen survey conducted Tuesday found Mr. Cain in the lead with 33% support from likely Republican primary voters."

A November 4 piece in Commentary magazine noted that Cain's apparent staying power in the polls "has led many observers to conclude that Cain is not merely a strong candidate but is actually bulletproof to charges that would destroy other men's hopes."

Whether those poll numbers begin to drop with time remains to be seen. But the rapid downward trajectory in the tone of Cain's news coverage is already apparent. As recently as October 10-16, his positive coverage exceeded negative by 19 percentage points in the broad spectrum of media outlets. Last week across that broad swath of outlets, the negative dominated by 13 percentage points—a dramatic 32-point swing in only three weeks.



Perry, Romney and the Rest of the GOP Field Last Week

If it weren't for the fact that he got significantly less coverage, Rick Perry would have had an even worse week in the broad cross section of news coverage than Cain. From October 31-November 6, only 20% of the assertions about Perry were positive compared with 41% negative and 39% neutral.

Perry however, was far less visible. He was present in a fraction of the coverage that Cain was. Perry for the week was a significant newsmaker in 11% of the campaign stories and a dominant newsmaker in 4%. His problems—including a circulating video of an unusually animated and unbound Perry speaking in New Hampshire—were obscured by the volume of tough Cain coverage.

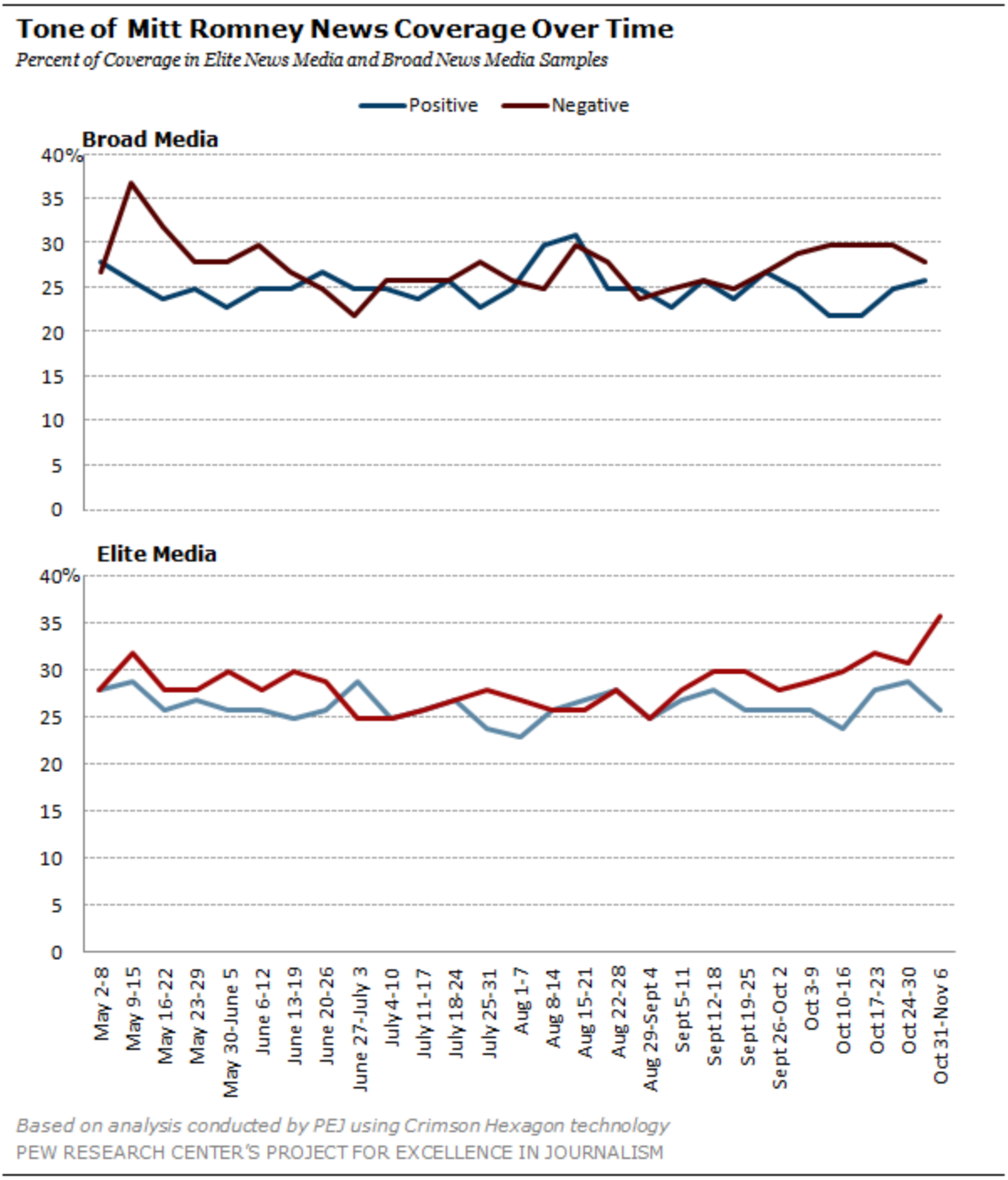
Perry has also seen a recent and dramatic turnaround in the tone of his media narrative. For 22 consecutive weeks, from May 2-October 2, Perry's positive news coverage exceeded negative coverage, based largely on his perceived strengths as a candidate and his instant rocketing to the top of the polls after his August 13 entry into the race.

But for the last five consecutive weeks starting in early October, the percentage of negative statements in the press have outnumbered positive ones about Perry, and the differential between negative and positive assertions has grown each week. Last week, it reached a new high with negative outstripping positive ones by 21 percentage points.

Romney, who received somewhat more coverage than Perry but far less than Cain last week, saw his overall pattern of a largely mixed media narrative continue.

Last week, 26% of the assertions about Romney were positive compared with 28% negative and 46% neutral. That actually marked a slight improvement over the previous four weeks when his negative assertions exceeded his positive ones by between four and eight percentage points.

As was the case with Cain, the tone of Perry's and Romney's coverage was also more negative in the subsample of "elite" media outlets last week. There, 19% of the assertions about the Perry were positive compared with 49% negative and 32% neutral—a 30 point negative to positive difference. For Romney, 26% of the assertions were positive, 36% negative and 38% neutral—a 10 point negative differential.



No other GOP candidates generated significant media coverage last week. Only three of them enjoyed more positive than negative coverage in the broader news sample. Whether coincidentally or not, they are the candidates lagging in the presidential preference polls—Jon Huntsman (25% positive, 17% negative, 57% neutral), Rick Santorum (24% positive, 19% negative, 56% neutral) and Ron Paul, who had the best week at 32% positive, 14% negative and 55% neutral.

Yet, when the sample is narrowed to elite media, only Paul emerged with coverage that was more positive (17%) than negative (13%) last week.

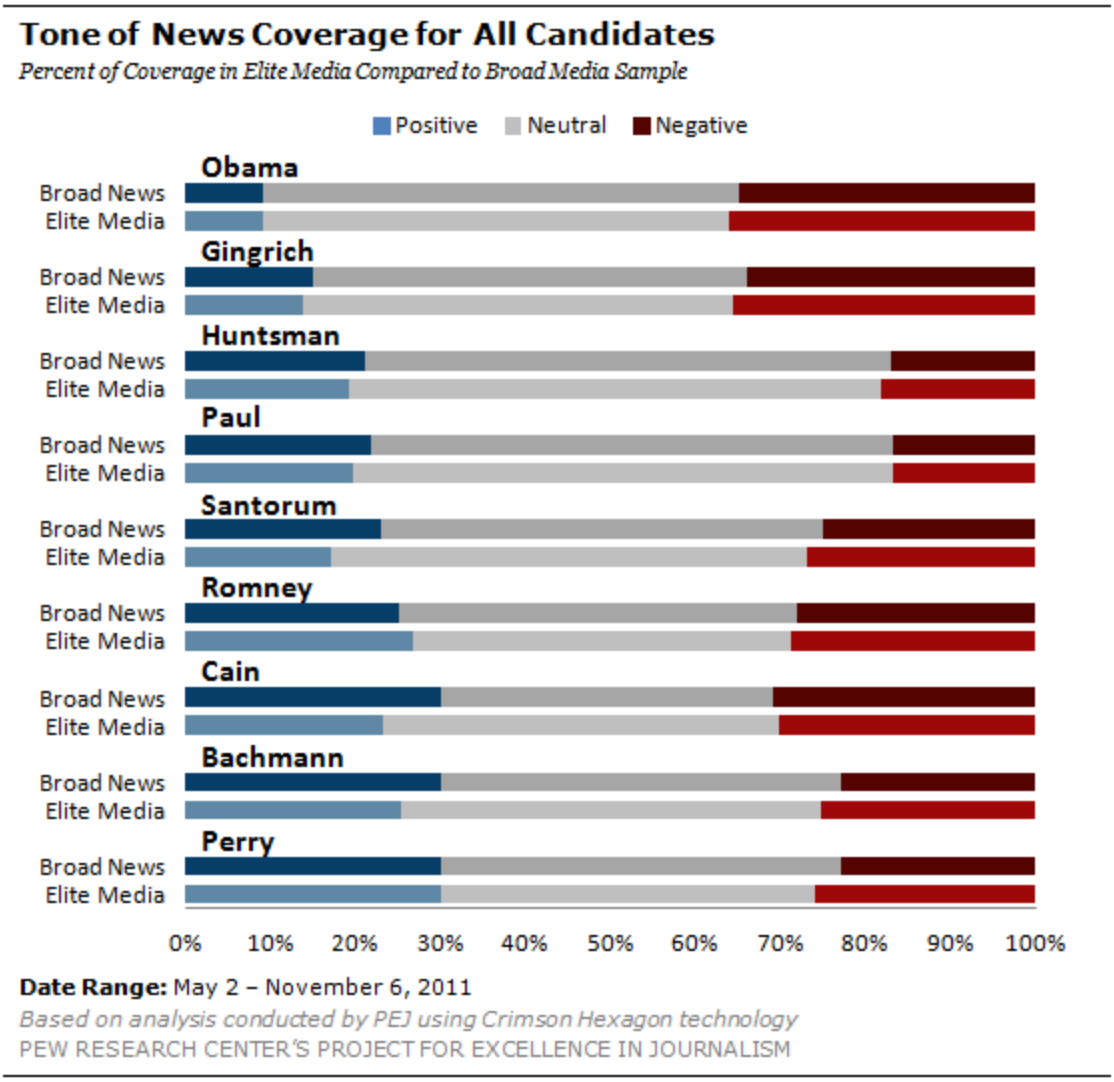
The Mainstream Media, Elite vs. Broad Sample: Differences and Similarities

The Project released its initial study of how the media have covered the campaign on October 17. That broad sample offers a remarkably deep and robust look at what the news media provide the American public. With roughly 11,500 outlets included, it is effectively a census of all the news media provided to Americans via RSS feeds. Only a handful of outlets are not included, notably those that have a strict paywall, such as the Wall Street Journal, but those with more porous paywalls, such as the New York Times, are included.

Some have wondered how that broad spectrum might compare to the smaller sample of outlets typical of more traditional media research that does not involve algorithmic coding. Because of the time involved in human coding, traditional research has used samples of media as a proxy for the media overall, a limitation that computers are able to avoid. PEJ, for instance, examines the output from 52 different outlets for its weekly News Coverage Index to assess media agenda—which topics the media are covering and which they are not. That sample (which includes print, cable, radio or audio, online and broadcast) is selected to provide a cross section of national media based in large part on audience numbers.

For this report, PEJ took the same outlets and coded the content for tone using the same mix of human and algorithmic coding that it used for the broader sample. (Three of the radio talk shows did not have RSS feeds, one newspaper had a paywall and one website does not produce original reporting).

The results show little variation between the two samples over the six month period from May 2-November 6.



For instance, for President Obama, 9% of the statements in the broad media sample, more than 1.7 million from May 2 to November 6, were positive. In the smaller "elite" sample, again 9% of the statements were positive. In the broad sample, 35% of the statements were negative. In the smaller elite sample the number was 36%, a nearly identical ratio of about four-to-one negative.

For Rick Perry, similarly, 30% of statements in both the broad sample and elite sample were positive, while 23% were negative in the broad sample and 26% in the smaller sample.

There are various possible explanations for why the widest spectrum of news outlets and a smaller sample of popular national outlets are so similar in tone. One is that the media conversation moves so quickly today, and is so easily consumed, that the assessments of candidates by elite national media outlets are mirrored by smaller ones relatively quickly. Another explanation is that the media ecosystem has become highly distributive, meaning that the content many local outlets are presenting is actually material that was produced by national outlets—and this distribution system is made even more efficient by digital technology. A local television or radio website that wants to post and distribute a national political story can easily

use wires, or rewrite them, without any of the limitations of space and time those media faced in their legacy platforms. An analysis of the content in the broader sample finds, indeed, that the elite media that cover the campaign regularly provided the highest percentage of the content that was coded in the broad spectrum of media. Some of these outlets include Yahoo.com, Reuters.com, Washingtonpost.com and NYTimes.com.

Another feature of the comparison between the broad media spectrum and the smaller elite sample is how similar they are in coverage of President Obama. Even the differences by week over the course of 27 weeks are negligible. When it came to the percentage of positive statements about him, the two samples were within two percentage points in all the weeks studied.

Why would the two samples of media be so similar about Obama? There are some possible structural reasons that may help explain this. One is that, as president, Obama is simply attached to problems in the country that he cannot escape—an economy that has been weakening, falling poll numbers, the fight over the debt crisis that critics felt was a sign of an increasingly dysfunctional Washington. In addition, Obama as a candidate has a host of Republican candidates making news every day while attacking him. House and Senate Republicans are doing the same. And as the president struggled, and his numbers slipped, members of his own party were critical as well. All of that is reflected in the news coverage.

There were some differences between the two samples, however. First, when the media narrative about a candidate shifts, the elite media tend to reflect the change more quickly. The broad spectrum of news outlets, in turn, tend to follow by about two weeks. And when the tone in the broader spectrum of news outlets shifts, the difference between positive and negative assertions about a candidate becomes even more pronounced there, a kind of amplifying echo effect across the media spectrum, in which matters become less nuanced.

In August, for instance, the tone in elite media became mixed and less positive about Michelle Bachmann two weeks before the tone in the broad media universe took a similar turn. The elite media, it seems, began to discount how much gain Bachmann would enjoy from winning the Iowa straw poll in mid August, as Perry entered the race. In the broad media sample, she enjoyed her two best weeks of coverage in late August, but saw the tone turn negative in September.

Similarly, the tone of coverage in elite media shifted to more negative than positive about Rick Perry the week of September 12-18, after some controversial debate performances. The tone turned in the broad media sample two weeks later, the week of October 3-9, following a particularly difficult debate performance in Florida and as he began to slip behind Romney and Cain in polls.

When the broad media turn, the tone also can become even more pronounced in one direction or the other than in the smaller elite media sample. Once the media overall became more negative about Perry, the differential between negative and positive assertions about his candidacy were at least five percentage points more than in the elite media every week thereafter, except for one. The same was true of Bachmann once negative assertions about her began to significantly outnumber positive ones in early October.

The pattern had already begun to emerge with Cain, even before the allegations of sexual harassment. Coverage of Cain in the broad sample became more positive in mid-August and continued that way through the week of October 17-23. Yet in elite media, the coverage became more mixed the first week of October.

If the pattern seen before between elite and broad media recurs now, the tone of coverage about Cain would become even more negative in the broader spectrum of media.

Crimson Hexagon Methodology

The study, The Media and Herman Cain, uses content analysis data from two sources. Data regarding the *quantity* of coverage is mostly derived from the Project for Excellence in Journalism's in-house coding operation. (Click here for details on how that project, also known as PEJ's News Coverage Index, is conducted.)

To arrive at the results regarding the *tone* of coverage, PEJ employed a combination of traditional media research methods, based on long-standing rules regarding content analysis, along with computer coding software developed by Crimson Hexagon. That software is able to analyze the textual content from billions of messages on blogs, Twitter, Facebook and web-based articles from news sites. Crimson Hexagon (CH) classifies online content by identifying statistical patterns in words.

Use of Crimson Hexagon's Technology

The technology is rooted in an algorithm created by Gary King, a professor at Harvard University's Institute for Quantitative Social Science. (Click here to view the study explaining the algorithm.)

The purpose of computer coding in general, and Crimson Hexagon specifically, is to "take as data a potentially large set of text documents, of which a small subset is hand coded into an investigator-chosen set of mutually exclusive and exhaustive categories. As output, the methods give approximately unbiased and statistically consistent estimates of the proportion of all documents in each category."

Universe

Crimson Hexagon software examines online content provided by RSS feeds of thousands of news outlets from the U.S. and around the world. This provides researchers with analysis of a much wider pool of content than conventional human coding can provide. Specifically, the monitors PEJ created for this study are based on more than 11,500 news web sites. CH maintains a database of all stories available so texts can be investigated retroactively.

While the software collects and analyzes online content, the database includes many news sites produced by television and radio outlets. Most stations do not offer exact transcripts of their broadcasted content on their sites and RSS feeds, however, those sites often include text stories that are very similar to report that were aired. For example, even though the television programs from Fox News are not in the sample directly, content from Fox News is present through the stories published on FoxNews.com.

The universe includes content from all the major television networks along with thousands of local television and radio stations. Two notable television sources, CBS and PBS' NewsHour, do offer transcripts of their television news programs, and those texts are including in the sample.

Elite vs. Broad Sample

For this report, PEJ examined two different samples using Crimson Hexagon's database of news outlets. The universe described as the "broad" sample includes all of the more than 11,500 news sites available. Not all of these outlets contain campaign stories on a regular basis, but any time they do, those stories are included in the sample. For instance, local television newscasts may not offer much coverage of the presidential campaign. However, the sample will include any relevant reports that do appear.

The universe entitled the "elite" sample is made up of a smaller collection of news sites that provide a focused cross section of national media based in large part on audience numbers. This elite sample is based on the 52 different outlets included in PEJ's weekly News Coverage Index (NCI) which includes print, cable, radio, online and broadcast.

Of the 52 outlets found in the NCI, 47 are included in the elite sample. For technical reasons, five sources cannot be represented. Three radio talk shows (Rush Limbaugh, Sean Hannity and Ed Schultz) do not have accompanying RSS feeds. The Wall Street Journal is not included in the algorithmic tone coding due to its paywall. Google News, which does not produce original material but rather pulls stories from other sources, is not included because the same material is coded in the other outlets where it appears.

The 47 outlets' content is distributed through 21 unique URLs. This number is smaller because television websites often serve as umbrellas for web feeds from multiple programs. For example, the URL abcnews.go.com includes feeds from both Good Morning America and ABC's World News with Diane Sawyer. Foxnews.com provides material from Fox News programs including Special Report with Bret Baier, Fox Report with Shepard Smith, the O'Reilly Factor and Hannity.

The list of URLs that are included in the elite sample are as follows:

1. www.msnbc.msn.com
2. www.today.msnbc.msn.com
3. www.ed.msnbc.msn.com
4. www.cnn.com
5. www.foxnews.com
6. www.abcnews.go.com
7. www.cbsnews.com
8. www.NPR.org
9. www.pbs.org/newshour
10. www.nytimes.com
11. www.washingtonpost.com
12. www.usatoday.com
13. www.ajc.com
14. www.latimes.com
15. www.toledoblade.com
16. www.azcentral.com

17. www.thehour.com
18. www.spokesman.com
19. www.joplinglobe.com
20. www.news.yahoo.com
21. www.huffingtonpost.com

Monitor Creation and Training

Each individual study or query related to a set of variables is referred to as a "monitor."

The process of creating a new monitor consists of four steps. (See below for an example of these steps in action.)

First, PEJ researchers decide what timeframe and universe of content to examine - general news stories, blogs, messages on the major social media sites Twitter and Facebook or some combination. For this study, the focus was solely on English-language news outlets.

Second, the researchers enter key terms using Boolean search logic so the software can identify the universe of posts to analyze.

Next, researchers define categories appropriate to the parameters of the study. If a monitor is measuring the tone of coverage for a specific politician, for example, there would be four categories: positive, neutral, negative, and irrelevant for posts that are off-topic in some way.

If a monitor is measuring media framing or storyline, the categories would be more extensive. For example, a monitor studying the framing of coverage about the death of Osama bin Laden might include nine categories: details of the raid, global reaction, political impact, impact on terrorism, role of Pakistan, straight account of events, impact on U.S. policy, the life of bin Laden, and a category off-topic posts.

Fourth, researchers "train" the CH platform to analyze content according to specific parameters they want to study. The PEJ researchers in this role have gone through in-depth training at two different levels. They are professional content analysts fully versed in PEJ's existing content analysis operation and methodology. They then undergo specific training on the CH platform including multiple rounds of reliability testing.

The monitor training itself is done with a random selection of posts collected by the technology. One at a time, the software displays posts and a human coder determines which category each example best fits into. In categorizing the content, PEJ staff follows coding rules created over the many years that PEJ has been content analyzing news media. If an example does not fit easily into a category, that specific post is skipped. The goal of this training is to feed the software with clear examples for every category.

For each new monitor, human coders categorize at least 250 distinct posts. Typically, each individual category includes 20 or more posts before the training is complete. To validate the

training, PEJ has conducted numerous intercoder reliability tests (see below) and the training of every monitor is examined by a second coder in order to discover errors.

Once the training is complete, the software analyzes the entirety of the identified online content. This classification is done by applying statistical word patterns derived from posts categorized by human coders during the training process.

How the Algorithm Works

To understand how the software recognizes and uses patterns of words to interpret texts, consider a simplified example. Imagine the study examining coverage regarding the death of Osama bin Laden that utilizes the nine categories listed above. As a result of the example stories categorized by a human coder during the training, the CH monitor might recognize that portions of a story with the words "Obama," "poll" and "increase" near each other are likely about the political ramifications. However, a section that includes the words "Obama," "compound" and "Navy" is likely to be about the details of the raid itself.

Unlike most human coding, CH monitors do not measure each story as a unit, but examine the entire discussion in the aggregate. To do that, the algorithm breaks up all relevant texts into subsections. Rather than the dividing each story, paragraph, sentence or word, CH treats the "assertion" as the unit of measurement. Thus, posts are divided up by the computer algorithm. If 40% of a story fits into one category, and 60% fits into another, the software will divide the text accordingly. Consequently, the results are not expressed in percent of newshole or percent of stories. Instead, the results are the percent of assertions out of the entire body of stories identified by the original Boolean search terms. We refer to the entire collection of assertions as the "conversation."

Testing and Validity

Extensive testing by Crimson Hexagon has demonstrated that the tool is 97% reliable, that is, in 97% of cases analyzed, the technology's coding has been shown to match human coding. PEJ spent more than 12 months testing CH and its own tests comparing coding by humans and the software came up with similar results.

In addition to validity tests of the platform itself, PEJ conducted separate examinations of human intercoder reliability to show that the training process for complex concepts is replicable. The first test had five researchers each code the same 30 stories which resulted in an agreement of 85%.

A second test had each of the five researchers build their own separate monitors to see how the results compared. This test involved not only testing coder agreement, but also how the algorithm handles various examinations of the same content when different human trainers are working on the same subject. The five separate monitors came up with results that were within 85% of each other.

Unlike polling data, the results from the CH tool do not have a sampling margin of error since there is no sampling involved. For the algorithmic tool, reliability tested at 97% meets the highest standards of academic rigor.

Ongoing Monitors

In some instances, PEJ uses CH to study a given period of time, and then expand the monitor for additional time going forward. In order to accomplish this, researchers first create a monitor for the original timeframe according to the method described above.

Because the tenor and content of online conversation can change over time, additional training is necessary if the timeframe gets extended. Since the specific conversation about candidates evolves all the time, the CH monitor must be trained to understand how newer posts fit into the larger categories.

In those instances, researchers conduct additional training for the monitor with a focus on posts that occurred during the new time period. For every new week that is examined, at least 25 more posts are added to the monitor's training. At that point, the monitor is run to come up with new results for the expanded time period which are added to results that were already derived in the original timeframe.

An Example

Since the use of computer-aided coding is a relatively new phenomenon, it will be helpful to demonstrate how the above procedure works by following a specific example.

PEJ created a monitor to measure the tone of media coverage on news sites for Republican candidate Mitt Romney. First, we created a monitor with the following guidelines:

1. Source: "News" sources only
2. Original date range: May 2 to September 11, 2011
3. English-language content only
4. Keyword: Romney

We then created the four categories that are used for measuring tone:

1. Positive
2. Neutral
3. Negative
4. Off-topic/Irrelevant

Next, we trained the monitor by classifying documents. CH randomly selected entire posts from the time period specified, and displayed them one by one. A PEJ researcher decided if each post is a clear example of one of the four categories, and if so, assigned that post into the appropriate category. If an example post is not clear in its meaning, or could fit into more than one category, such as a story with a mix of positive and negative assertions, the coder skipped the post. Since

the goal is to find the clearest cases possible, coders will often skip many posts until they find good examples.

A story that is entirely about a poll showing Mitt Romney ahead of the Republican field – and that his lead is growing, would be a good example to put in the "positive" category. A different story that is entirely about Romney's record in Massachusetts and how many conservative voters are opposed to him would be put in the "negative" category. A post that is strictly factual, such as a story about a speech Romney gave on the economy that does not include evaluative assessments, would be put in the "neutral" category. And a post that includes the word "Romney" but is not about the candidate at all, such as a story about a different person with the same last name, would go in the "off-topic" category.

The coder trained 260 documents in all - ten more than the necessary minimum of 250. Each of the four categories had more than 20 posts in them.

At that point, the initial training was finished. For the sake of validity, PEJ has another coder check over all of our training and look for stories that they would have categorized differently. Those stories are removed from the training sample because the disagreement between coders shows that they are not clear, precise examples. In the case of the Romney monitor, there were four documents that were removed for this reason.

Finally, we "ran" the monitor. This means that the algorithm examined the word patterns derived from the monitor training, and applied those patterns every post that was captured using the initial guidelines. Since the software studies the conversation in an aggregate as opposed to individual posts or stories, the algorithm divided up the overall conversation into percentages that fit into the four categories.

For the initial monitor, the algorithm examined over 94,00 assertions from thousands of news stories and determined that 34% of the conversation was positive, 33% neutral, and 33% negative. The assertions or statements that are off-topic were excluded from the results.

In order to extend the Romney monitor beyond September 11, coders added at least 25 new pieces of content to the training for each new week examined. This assures that any linguistic changes in the overall coverage or conversation regarding Romney in the new week are accounted for. We then run the monitor again, which now includes the original training of 260 posts plus 25 new ones, for the new week while leaving the earlier results in place.