

Content Analysis Methodology

For more information:

Limor Peer, l-peer@northwestern.edu, 847-467-6578
Mary Nesbitt, m-nesbitt@northwestern.edu, 847-467-4285

The content analysis includes all types of information that appear in daily and Sunday newspapers. Stories, photographs, graphics, display and classified advertising and newspaper promotion all fall within the scope of this study. The analysis consists of counting and classifying types of content.

While we gathered information about the entire paper such as number of pages and number of sections, the bulk of the analysis focused on each content item in depth. We analyzed each story in terms of characteristics such as topic, sources, treatment, geographic focus, photos and writing style. We analyzed newspaper promotions for things such as type, color, topic and size and ads for things such as category, technique, contact information and photos.

Selecting representative newspaper issues

Once a sample of 52 newspapers was established http://www.readership.org/new_readers/sample_selection.pdf, we selected eight publishing days from October 2003 to form a constructed week of seven non-consecutive days, plus one additional Sunday. The sampling plan was designed to minimize distortion in the newspaper's content caused by a single major news event, holidays or other extraordinary circumstances. Each of the 52 newspapers provided a copy of the complete final home edition for the dates of October 1, 4, 5, 7, 10, 12, 13, and 16. The total number of newspaper issues requested for the study was 416, and the final number of newspaper issues analyzed was 414 (one paper publishes only six days and one paper did not send us a paper for one of the dates). We excluded all pre-printed advertisements, syndicated supplements such as Parade and USA Weekend magazines, special sections and stand-alone television guides.

Preparing newspapers for analysis: Identifying and categorizing content

Before beginning to work on each newspaper, we checked to ensure we had a complete edition, removed pre-printed advertisements and verified that sections were in the correct order. Coders then counted and recorded the number of sections, the number of pages (both standard-size or broadsheet pages and non-standard size or tabloid pages) and whether the paper had a classified advertising section.

Next, we examined each newspaper systematically and identified all stories, newspaper promotions and ads. We defined each of these content types and categorized them based on fixed criteria http://www.readership.org/new_readers/analysis.pdf. Once identified, each content type was marked with a color-coded marker and assigned a number. A total

of 66,841 stories were identified in the sample. Based on each newspaper's circulation size, we randomly sampled between 50% and 70% of the stories for analysis. Similarly, a total of 32,987 newspaper promotions were identified, from which we randomly sampled between 50% and 70% from each paper for analysis. A total of 43,035 ads were identified and, again based on circulation size, we randomly sampled between 15% and 25% of the ads for analysis.

The sampling procedure yielded a total of 33,634 stories, 21,279 newspaper promotions and 11,233 ads that were analyzed for this study.

This report includes the main findings for the entire content analysis. We present results for the sample as a whole and for newspapers by circulation size. We also present the results for weekdays (Monday through Saturday) and for Sunday separately. When calculations were made for the entire 7 days of the week, we averaged the data from the two Sundays in each paper to construct an "average Sunday."

Coding procedures and reliability testing

We hired a total of 30 coders for this research project. Coders were divided into teams, each working on a separate part of the analysis. The first team was responsible for preparing the newspapers for analysis and for coding the stories, the second team coded newspaper promotions and the third team coded ads.

All 30 coders were extensively trained by research manager Limor Peer, before beginning coding. For all three teams training consisted of an overview of the project, an explanation of the method and general procedures and an introduction to the specific coding instructions and rules they were to use http://www.readership.org/new_readers/content.htm. Next, coders practiced several sets of selected examples of stories, promotions or ads (depending on what content type they were assigned to analyze). After each set of examples was analyzed, the team convened and discussed the correct coding. This was an ongoing process, which allowed coders to raise questions, clear ambiguities and clarify coding rules when needed. When coders demonstrated a satisfactory level of understanding of the coding rules and procedures and a high degree of accuracy in coding they began coding the newspapers. During the initial phases of the coding, the teams convened for regular meetings to discuss difficulties with coding decisions and any additional clarifications that may have been necessary. Throughout the coding period, coders' work was periodically spot-checked.

While a great deal of effort went into the training phase for the purpose of creating consistency across coders, we also pulled a random sample of 850 stories, 708 promotions, and 487 ads, and had randomly-assigned coders code those a second time to allow us to evaluate inter-coder reliability. We then compared the two sets of coding forms on key variables. The results show a satisfactory level of agreement between the coders (Cohen's Kappa coefficient averaged 0.73 for stories and ads, and 0.83 for promotions).