Communication Methods and Measures, 5(1), 76–91, 2011

ISSN: 1931-2458 print / 1931-2466 online
DOI: 10.1080/19312458.2010.547823

How Much Is Enough? New Recommendations for Using Constructed Week Sampling in Newspaper Content Analysis of Health Stories

Douglas A. Luke and Charlene A. Caburnay
Washington University, St. Louis

Elisia L. Cohen
University of Kentucky

Researchers frequently use constructed week samples to approximate content for larger populations of textual data in content analysis projects. To date, this sampling method has not been validated in longitudinal contexts necessary for the conduct of large-scale health communication research. This study uses Monte Carlo bootstrap sampling to determine the number of constructed weeks necessary to accurately estimate one- and five-year population values for different types of variables in a quantitative content analysis. Five years (1999–2004) of four different daily newspapers were coded for four variables that varied on type (count vs. rating), amount of missing data, and distribution (normal vs. nonnormal). Results suggest that sampling a minimum of six constructed weeks was most efficient for both time frames. Missing data lowers sampling precision, although a correction can be calculated if

This project was funded through the U.S. Centers for Disease Control and Prevention contract U48/CCU710806. The authors thank Matthew Kreuter, Robert Logan, Brooke Asbury, Caren Bacon, Darigg Brown, Vicki Collie-Akers, Kristy Davidson, Mark Graves, Heather Jacobsen, Keri Jupka, Susan Lukwago, Vinay Reddy, Jyothi Varanasi, and Anitha Vempaty for their assistance in planning and carrying out the everyday activities of the study. Some statistical analyses and other support were provided by the Saint Louis University Center of Excellence in Cancer Communication Research (1P50CA095815).

Correspondence should be addressed to Douglas A. Luke, Ph.D., George Warren Brown School of Social Work, Washington University in St. Louis, Campus Box 1009, 700 Rosedale Ave., St. Louis, MO 63112-1408. E-mail: dluke@gwbmail.wustl.edu

the amount of missing data can be estimated. Using an efficient method of sampling newspapers such as constructed week sampling can help communication researchers to more easily study health coverage in the media.

INTRODUCTION

A fundamental challenge for communication science is how researchers choose to identify and acquire appropriate texts when conducting quantitative content analysis studies. As this study will demonstrate, there are limited options available to researchers who seek to approximate content for large populations of textual data. *Constructed week sampling* is a sampling technique that is becoming a popular tool for communication scientists but has not been validated in longitudinal contexts necessary for the conduct of large-scale media and health communication research. As such, this study is a first step in a broader effort to validate the use of constructed week sampling in a longitudinal media research context. Also, the study identifies the methodological strengths of utilizing a Monte Carlo bootstrap sampling technique to validate constructed week sampling and describes the limitations to the study's generalizability by clarifying areas for future research.

BACKGROUND

Content analysis is one of the fastest-growing methods used in quantitative mass and health communication research (Riffe & Freitag, 1997). For example, recent communication studies have used content analysis to demonstrate how particular health-mediated messages target individuals or groups (Campo & Mastin, 2007; Mastin, Andsager, Choi, & Lee, 2007; Stryker, Emmons, & Viswanath, 2007) and to evaluate the ways in which theoretically grounded appeals are deployed in health communication contexts (Cohen, Shumate, & Gold, 2007; Stephenson, 2002; Stephenson & Quick, 2005).

The recent growth in longitudinal content analysis may be related, in part, to the availability of digital media archives and large-scale community-based health communication research projects (Stryker et al., 2007). Frequently, content-analysis studies are the first step in a two-step evaluation process linking media content to viewer and reader perceptions (e.g., Caburnay et al., 2008; Cohen, Caburnay, Luke, Rodgers, & Cameron, 2008; Niederdeppe, Davis, Farrelly, & Yarsevich, 2007; Stryker, 2003; Stryker, Solky, & Emmons, 2005). However, such research also demands that researchers consider the best available means of efficiently and reliably gathering population data.

When the scope of content analysis is expanded to larger populations, researchers determine whether it is necessary to analyze the entire population or if sampling could be used (Stryker et al., 2007; Long, Slater, Boiarsky, Stapel, & Keefe, 2005). However, communication researchers have criticized content analytic scholarship for often ignoring issues of sampling error (Krippendorff, 1980; Long et al., 2005; Neuendorf, 2002; Stryker et al., 2007).

Electronic databases for textual analysis (e.g., LexisNexis) coupled with researchers' ability to capture and store video and digital text allows for examining larger populations of texts with relative ease. At the same time, digital databases require researchers to consider new sampling techniques and introduce new issues of sampling error. Although Stryker and colleagues (2006) have developed recommendations to measure the quality of a database-driven search, these techniques require that the *reach* of the database search extends equally to every available newspaper. However, community-based researchers, particularly those who rely on small- to medium-scale media, have historically relied on media that lack such extensive digital archives (e.g., Ethnic NewsWatch) (Stryker et al., 2007). Accordingly, while electronic databases have aided researchers' sampling strategies, communication researchers continue to use a variety of media sampling methods while drawing inferences to broader populations.

For example, mass communication researchers interested in longitudinal analysis of community-based newspapers or difficult-to-reach media often rely on purposive and stratified samples (Boyle, McCluskey, McLeod, & Stein, 2005, p. 642; Clark & Illman, 2006). Certainly the cost of copying each issue from microfilm for coding is an understandable concern. However, the scientific support for utilizing purposive, stratified samples to generalize to the broader population under longitudinal consideration generally remains unsupported, with communication researchers preferring more robust sampling procedures (Long et al., 2005; Stryker et al., 2006).

For this reason, Fink and Gantz (1996) recommended sampling procedures in content analysis that allow researchers to generalize from the sample to the population with random, or probability, sampling as the ideal. *Constructed week sampling* is a type of stratified random sampling (SRS) technique popular in media studies in which the final sample represents all seven days of the week (Jones & Carter, 1959; Stempel III, 1952) to account for cyclic variation of news content (Riffe, Aust, & Lacy, 1993). For example, constructing one week from a population of one month would involve selecting one Sunday from all four Sundays that month, one Monday from all four Mondays, etc., until each day of the week is represented in the final sample. This technique could be modified for the size of both the final sample and the population.

The overall goal of constructed week sampling is to create maximum sampling efficiency while controlling for cyclical biases (e.g., weekly news patterns). Too few sampling units may lead to unreliable estimates and invalid results, whereas too many may be a waste of coding resources (Riffe et al., 1993). A number of studies have attempted to identify the number of weeks needed for efficient constructed week sampling for different sized populations and media (Lacy, Riffe, & Randle, 1998; Lacy, Riffe, Stoddard, Martin, & Chang, 2001; Riffe et al., 1993; Riffe, Lacy, & Fico, 2005). However, these studies varied in their recommendations of the most efficient number of weeks to construct for a given population size and examined only a limited number of variables. In a daily newspaper-based study, Riffe et al. (1993) recommend selecting one constructed week for a population of 182 days to identify the average number of local news stories per day. In addition, Lacy et al. (2001) propose selecting nine constructed weeks for a population of five years when examining daily newspaper content over multiple years to identify variables such as number of photos, informational graphics, stories, and stories from staff. Given these time frames and results, it is difficult to generalize to other populations unless the study is designed to match these specific recommendations. There has not yet been a consistent framework under which we can apply constructed week recommendations for daily newspapers over a one- to five-year period for different types of variables. There is a need to clarify and extend guidelines and specific recommendations for constructed week newspaper sampling to inform such research examining content across a heterogeneous group of media and health variables with differing degrees of variability.

This study will demonstrate a general approach for determining sampling efficiency in daily newspapers by using Monte Carlo bootstrap sampling, a relatively rare methodology in communication research (Stephenson & Holbert, 2003). Here, a Monte Carlo bootstrap sampling approach can be applied to help understand how health stories are treated in four daily newspapers during a five-year timeframe. Most importantly, until now research validating constructed week sampling has used basic news-type count variables (e.g., total number of local news stories, photographs, graphics, and stories in total inches about the United States) (Lacy et al., 2001; Riffe et al., 1993). However, media analysts, particularly those who analyze health coverage, are often interested in the varied characteristics or qualities of news coverage, not just the frequency of coverage (e.g., Atkin, Smith, McFeters, & Ferguson, 2008; Caburnay et al., 2008; Campo & Mastin, 2007; Cohen et al., 2008; Mastin et al., 2007; Powers & Andsanger, 1999; Stryker et al., 2006). Thus, this Monte Carlo study gives guidance for quantitative content analysis methodology by examining a broader array of content analysis constructs, including types of variables different from previous longitudinal studies of daily newspaper content (Lacy et al., 2001).

This study will build and extend upon previous studies by testing sampling efficiency with different types of health-related variables (i.e., type, missing data, and distribution). In order to expand the current body of research on constructed week sampling of daily newspapers, this study addresses whether a general approach for constructed week sampling of daily newspapers can be developed for use with one- and five-year time frames, and whether different types of health-related variables (count versus rating, higher versus lower levels of rating occurrence, skewed versus normal distribution) contribute to differing sampling efficiencies.

METHOD

Original Study Description and Procedures

As part of a larger study of health coverage in small-market media, four daily print newspapers from mid-size Missouri communities (Cape Girardeau, Poplar Bluff, Joplin, and Hannibal) were tracked every day for five years: February 1, 1999 through January 31, 2004. For more information about the original study, see Caburnay et al. (2003). These four newspapers' circulations range in size from approximately 8,000 to 29,000 ($M = 16{,}000$) and are located in communities ranging in population from 16,000 to 46,000 (Caburnay et al., 2003). These newspapers are typical of smaller market daily newspapers in the Midwest, and particularly in Missouri, where the average daily newspaper circulation is 20,000. Geographically, these newspapers also span four areas of the state: southwest, northeast, lower east central, and southeast (i.e., the "Bootheel").

Inclusion criteria

For a newspaper story to be included in the study, it had to address diet, physical activity, or tobacco and describe the relationship between the health topic and chronic disease within the headline or article. All advertising, comic strips located in the cartoon section, tables and listings of stock and commodities prices, law enforcement surveillance (e.g., police blotters, highway patrol reports), school lunch menus, and recipes were excluded. Over the five years, a total of 8,342 health stories were identified in 1,826 individual newspaper issues for the four newspapers. The number of identified health stories ranged from 1,249 for Hannibal to 2,737 for Cape Girardeau. On average, these four newspapers had approximately 1.14 health stories per daily issue.

Measures

Coding criteria

For the purposes of this study, the newspaper issue is the level of analysis. Data from content analyses of individual story characteristics were aggregated by issue across multiple health stories when there was more than one health story identified in a particular issue. Stories meeting the inclusion criteria were coded for total number of included *stories* per issue, total number of stories that included a *local angle*, average story *prominence*, and average amount of *science* information. Each of these is described below.

Inter-rater reliability was assessed for all coders for determining which newspaper stories met inclusion criteria and in applying all coding criteria using a 10% subsample of newspapers and stories. Krippendorff's alpha (2004) for the coding of these variables ranged from .70 to .79. Trained project staff members identified health stories within each newspaper and coded them on more than 40 journalistic and public health variables; however, this current study will focus on four of these variables.

Total number of included health stories per issue is a count variable that sums the total number of stories meeting inclusion criteria for a particular issue.

A *local angle* was present if information from or about a local individual or organization was included in the story.

Prominence was measured using a 7-point rating scale (1–7; 1 = lowest prominence, 7 = highest prominence) that reflected a story's total area, inclusion of a visual, and size of visual. *Total area* was measured by summing the square inches from across all parts of the story, including text, headlines, bylines, and visuals. *Inclusion of a visual* meant the story had a photograph or other visual element. *Size of visual* was measured by measuring the area of the visual (if any) and determining whether the visual took up 20% or more of the page or less than 20% of the page. This scale had good internal consistency ($\alpha = .80$).

Science was measured using a 4-point rating scale (0–3; 0 = lowest science, 3 = highest science) that reflected a story's mention of data, mention of investigators, and mention of scientific source. *Data mention* required reporting of specific quantitative and/or qualitative data from a research study. *Mention of investigators* required that the story contain the name or names of members of the research team. *Mention of scientific source* required that the story contain a scientific source, such as a journal name or conference. This scale had high internal consistency ($\alpha = .90$).

Table 1 reports summary statistics for these four variables for the five years of the newspaper content analysis for each newspaper. These four variables were chosen for this Monte Carlo study to represent the types of variables commonly used in health communication content analysis studies. In addition, we selected these variables because they reflect three different comparisons by type of variable: count versus rating, skewed versus normal distribution, and higher versus lower rating occurrence.

Count versus rating

Two of these variables are counts (*stories, local angle*), and two are ratings of story characteristics (*prominence, science*). Counts and ratings have differing

TABLE 1
Characteristics of Newspaper Stories Used in Simulation, Four Newspapers

Variable	*Type*	*Mean (μ)*	*SD (σ)*	*Skew*	*% with No Ratings*
Cape Girardeau					
Stories	Count	1.50	1.37	1.16	0%
Local Angle	Count	0.34	0.61	1.88	0%
Prominence	Rating	3.56	1.50	0.14	25.9%
Science	Rating	0.20	0.51	3.31	25.9%
Poplar Bluff					
Stories	Count	0.93	1.14	1.40	0%
Local Angle	Count	0.20	0.51	3.30	0%
Prominence	Rating	3.10	1.60	0.41	46.9%
Science	Rating	0.37	0.80	2.28	46.9%
Joplin					
Stories	Count	1.45	1.31	1.32	0%
Local Angle	Count	0.23	0.54	4.26	0%
Prominence	Rating	3.55	1.33	0.13	25.7%
Science	Rating	0.22	0.55	3.07	25.7%
Hannibal					
Stories	Count	0.68	1.03	2.12	0%
Local Angle	Count	0.16	0.48	4.56	0%
Prominence	Rating	3.73	1.53	0.08	58.2%
Science	Rating	0.27	0.66	2.71	58.2%

Note. $N = 1{,}826$ newspaper issues (5 years).

underlying statistical properties which may influence their sampling properties when using constructed week sampling.

Higher versus lower rating occurrence

More specifically, depending on the design of the content analysis study, ratings of story characteristics may have a greater amount of missing or not applicable data than simple counts. Here, for example, we have count information for every issue. When there is no health story in any particular issue, the count variable has the value 0. However, a rating of a story characteristic such as *prominence* will have no data for issues without any health stories, thus ratings do not apply. Table 1 shows that the two rating variables used in this study have from 25.7–58.2% missing data than the count variables, depending on the newspaper.

Skewed versus normal distribution

Content analysis variables may also vary on their distributional characteristics—some may be normally distributed while others may exhibit varying degrees of skewness, kurtosis, and so forth. Previous literature on constructed week sampling provides no guidance on how variable distributional characteristics may influence the efficacy of constructed week sampling. For example, are more constructed weeks required when the underlying variables show high degrees of non-normality? For this study we have included variables which are relatively normal (*prominence, stories*) and other variables which are much less normally distributed (*local angle, science*).

Statistical Analysis

We developed a Monte Carlo simulation with four variables (described above) to determine two types of estimates: (1) the number of constructed weeks necessary to accurately estimate one-year population values and (2) the number of constructed weeks necessary to accurately estimate five-year population values. The *R* statistics package, version 2.8.1, was used for all analyses (R Development Core Team, 2008).

The bootstrap sampling procedure was as follows. The population (census) of data included five years or 1,826 days of content analysis data from four daily newspapers. For each individual newspaper, and for all newspapers combined, a random one-year window from the population of five years was selected, and k number of constructed week samples (with replacement) were constructed from this random year (note: $k = 1$–10, 12, 15, 20, 30, 40, and 50 constructed weeks). Raw and standardized deviations of each constructed week sample mean were calculated from the one-year population mean. This entire process was repeated for each number of constructed weeks (k) and repeated for 1,000 bootstrap replications (Mooney & Duval, 1993). All raw bootstrap data and summaries were stored for further processing and validation.

For each of the 1,000 replications within a given number of constructed weeks, a standardized deviation score was calculated:

$$S\Delta = \frac{\bar{x}_{bootstrap} - \mu}{\sigma}$$

where $\bar{x}_{bootstrap}$ is the sample mean of the variables (i.e., stories, local angle, etc.) across the number of constructed weeks (1, 2, 3, etc.); μ is the mean of the same variable across the entire one-year or five-year population of newspaper stories; and σ is the population standard deviation. The deviation score ($S\Delta$) tells us how far away the sample bootstrap mean is from the one- or five-year population mean. Our expectation is that as the number of constructed weeks increases, the size of the bootstrap deviation scores will become smaller. The standardized bootstrap 95% confidence intervals were calculated for each simulation and are standardized by dividing by the standard deviation in order to allow comparisons across analyses and variables. These confidence intervals are the primary products of this study. Each confidence interval represents a predicted precision range for a given number of sampled constructed weeks. By examining the confidence intervals from the Monte Carlo study across a number of constructed weeks and different types of variables, we can make recommendations about the minimal number of constructed weeks necessary to obtain reasonably accurate estimates for future quantitative content analysis studies.

RESULTS

Figure 1 presents the primary results of the bootstrap simulations addressing Research Question 1 for the Joplin newspaper, which plots the confidence interval ranges against the number of constructed weeks for both the one- and five-year population windows. As expected, as the number of constructed weeks increases, the confidence interval decreases. The dashed vertical line in the figure shows the six constructed week cutoff. The same general pattern holds whether one is estimating a one-year population value or a five-year population value. Moving from one to six constructed weeks results in a rapid decrease in the bootstrap confidence interval ranges. After this sharp decline the improvement levels off with additional constructed weeks.

The bootstrap simulations were run for all four newspapers, and the results were virtually identical for each newspaper. Figure 2 shows the standardized bootstrap CI ranges for the health story count variable for each newspaper. So, although the variable characteristics differed across the four papers (see Table 1), the bootstrap simulations show that there is a common mechanism underlying the reliability of constructed week sampling numbers. The detailed simulation results and specific data underlying the figures for each newspaper are available from the authors.

Count versus rating

To address Research Question 2, an examination of Figure 1 shows that although all four variables show the same general pattern of prediction accuracy by number of constructed weeks, two of the variables (*Prominence, Science*) show larger confidence interval ranges than the other two variables (*Stories, Local Angle*) across the entire range of constructed weeks.

Higher versus lower rating occurrence

Prominence and *science* are the two ratings variables that have a large amount of missing or not applicable data. The observations are not applicable when the underlying health story count variable is 0. This occurred 25.7% of the time for

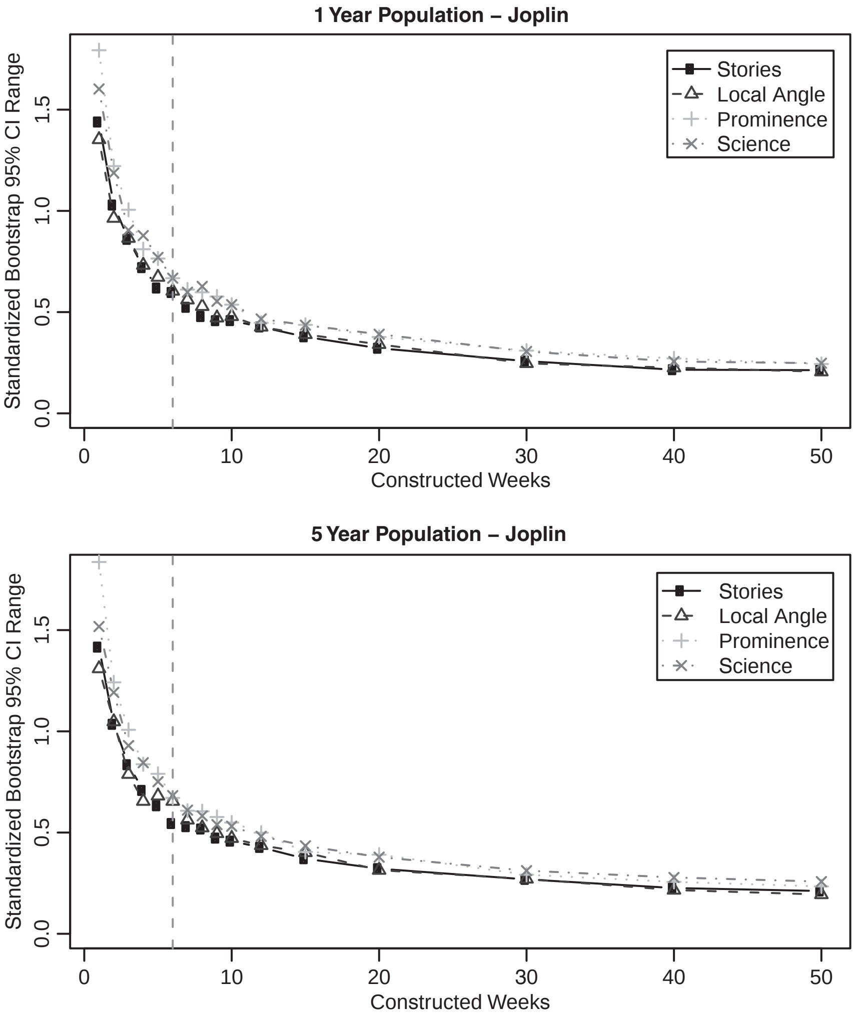

FIGURE 1 Standardized bootstrap 95% confidence interval ranges by number of constructed weeks for one- and five-year time frame for Joplin newspaper only, showing reduced improvements after six constructed weeks.

these data from Joplin. The smaller number of applicable observations leads to reduced accuracy in estimating population values. This makes sense when one considers that fewer observations lead to "shorter" constructed weeks. That is, on average, a constructed week of content analysis data based on a variable with 25% fewer observations will be 25% shorter, or 5.25 days instead of 7.

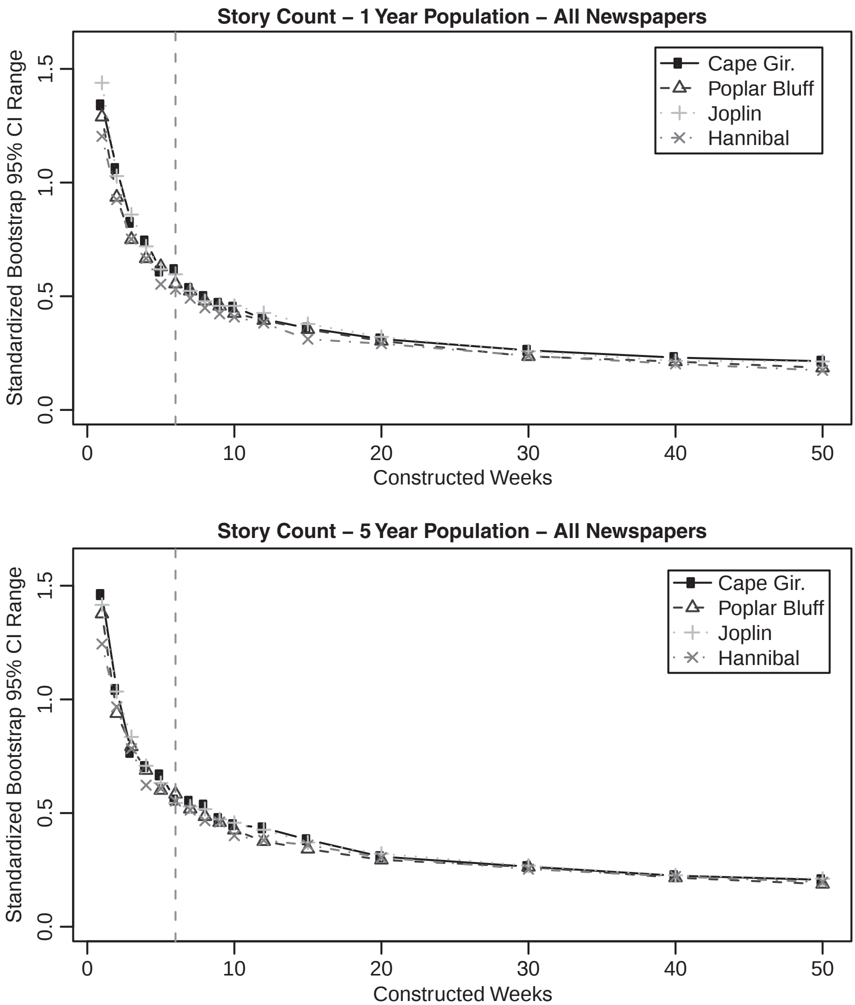

FIGURE 2 Standardized bootstrap 95% confidence interval ranges by number of constructed weeks for one- and five-year time frame for all four newspapers, showing reduced improvements after six constructed weeks.

Skewed versus normal distribution

Because *prominence* and *science* showed larger confidence interval ranges than the others variables, the results suggest that whether or not a variable is a count or rating variable is more predictive of sampling accuracy when compared to whether or not the variable has a skewed or normal distribution.

DISCUSSION

The main objective of constructed week sampling in quantitative newspaper content analysis is to maximize sampling efficiency or to obtain the most accurate estimate of the population by sampling the least number of newspaper issues. Constructed week sampling is more advantageous to simple random sampling because it accounts for variation of news content over a seven-day news week. In the early days of content analysis research, every newspaper issue in the population of interest was reviewed. However, with limited coding resources and technological advances such as LexisNexis, it is not feasible to continue reviewing every issue in the population and furthermore, sampling theory suggests that it is unnecessary.

Past work has clearly demonstrated that constructed week sampling is a cost-efficient and unbiased approach to sample daily newspaper issues for analysis (Lacy et al., 2001; Riffe et al., 1993). The current study builds on these past two studies by examining the utility and efficiency of constructed week sampling for an even wider variety of variable topics and traits.

This study took advantage of a large-scale, five-year content analysis dataset to examine the sampling properties of various numbers of constructed week samples for variables with different distributional properties. The central limit theorem tells us that as the number of constructed week samples increases, we should expect that the accuracy of these estimates will also increase. However, previous work in this area has not taken full advantage of this statistical theory to provide general recommendations across types of media or variables about the required number of constructed week samples.

Overall, this study found that sampling a minimum of six constructed weeks was most efficient based on confidence interval ranges. Although estimate precisions continue to improve past six constructed weeks, this improvement is relatively small. The cost-benefit of collecting more than six constructed weeks we think would be low for most studies. For example, examining the bootstrap confidence intervals in Figure 1 shows that moving from one to six constructed weeks results in an improvement of approximately 59% from the initial one-week confidence intervals. To see the same degree of improvement after six constructed weeks would require an additional 32.5 constructed weeks. This clearly shows the diminishing returns after the first six weeks of constructed week sampling.

The cost of data collection increases linearly with each constructed week, but the return in precision is reduced exponentially. As the central limit theorem indicates, the precision of sample estimates is based on the square root of the sample size (Pagano & Gauvreau, 2000). With each additional constructed week we are adding seven additional data points, one for each day. So this explains why after six constructed weeks (with 42 days of data), additional weeks add only minimal improvements.

The second important finding from this study was that these precision estimates hold whether or not you are generalizing to a one- or five-year time frame. This makes sense once one realizes that it is the number of sampling units (i.e., constructed weeks) that leads to precision, not the size of the population. For quantitative content analysis studies, this suggests that one does not need to increase the number of constructed week samples simply based on the time frame that is to be estimated.

Third, not unexpectedly this study showed that the precision of constructed week sampling estimates is lowered when there are missing data within the sampling timeframe. However, if an investigator can estimate the amount of expected missing, not applicable (NA), or irrelevant data in a study, a correction for the number of constructed weeks can be calculated. This correction is simply the number of originally desired constructed weeks divided by 1 minus the percent missing or NA:

$$CW_{corrected} = \frac{CW_{original}}{(1 - \%NA)}$$

For example, if the variable has 25% fewer observations due to missingness or NA observations, instead of sampling six constructed weeks ($CW_{original}$), we would instead sample eight constructed weeks ($CW_{corrected}$) to account for the missing data. In this case, the longitudinal study provided data for missing value rates to be assessed for each variable after the fact. Researchers may also rely on prior literature or utilize pilot studies to assess this ratio in order to construct a sampling plan.

Fourth, the Monte Carlo simulation suggests that as long as enough constructed weeks are used to create precise estimates, the skewness of the variable will not influence the precision of the estimates. This phenomenon also reflects the underlying statistical theory (i.e., central limit theorem).

Finally, although the detailed results presented are based on a single newspaper's data, the bootstrap simulations run on the other three newspapers match those obtained from Joplin. The number of constructed weeks needed to sample, for either a one- or five-year time frame, was identical across all four newspapers. The only difference between newspapers was the amount of missing values. Thus the correction to account for the missing data was different for each newspaper. Future research can model this approach by examining an even larger set of newspapers (e.g., ranging in circulation size, target audience, regionalism).

One important limitation of this study is that it examines one broad content category (health) in four daily newspapers. We do not believe that the findings and recommendations are specific only to health-related content, except that missing value rates may differ across content areas and require a different correction for the number of constructed weeks. Given that the presence of any health story or other content area is not standard in every newspaper issue, one-fourth of all issues sampled in this particular newspaper did not contain at least one health story. As such, generalizing these results to all other content categories or other types of newspapers may be inappropriate. In addition, while we were able to replicate the bootstrap simulation on four different newspapers with similar results, we acknowledge that these were all smaller daily newspapers in the same state. We might see different results if we examine larger newspapers with larger newsholes for health stories.

This constructed week sampling approach has important implications specifically for health communication. With an efficient method of sampling newspapers, health communication researchers can more easily study health coverage in the media. Because the resource constraints are minimized with this sampling method, it will be more cost efficient to be able to determine how many constructed weeks are needed based on the time frame and types of variables of study. As a result, having more research on health coverage in newspapers may result in greater and more informative coverage of health topics overall, and eventually change the types of health information the public receives from the media.

This study provided a general approach for determining the number of constructed weeks needed for sampling daily newspapers and examined how different types of variables affected sampling efficiencies. While this study examined health stories from four daily newspapers, future studies could replicate this study's process using larger newspapers or other types of news media (e.g., television, radio, Internet) using different lengths of time (e.g., constructed months or years). In addition, this logic could be extended to address issues of seasonality using larger amounts of time. This Monte Carlo study provides a method by which these other directions could be easily tested.

REFERENCES

Atkin, C., Smith, S., McFeters, C., & Ferguson, V. (2008). A comprehensive analysis of breast cancer news coverage in leading media outlets focusing on environmental risks and prevention. *Journal of Health Communication*, *13*, 3–19.

Boyle, M., McCluskey, M., McLeod, D., & Stein, S. (2005). Newspapers and protest: An examination of protest coverage from 1960 to 1999. *Journalism & Mass Communication Quarterly*, *82*, 638–653.

Caburnay, C. A., Kreuter, M. W., Cameron, G. C., Luke, D. A., Cohen, E. L., McDaniels, L. M., Wohlberg, M. T., & Atkins, P. J. (2008). Black newspapers as a tool for cancer education in African American communities. *Ethnicity and Disease*, *18*, 488–495.

Caburnay, C. A., Kreuter, M. W., Luke, D. A., Logan, R. A., Jacobsen, H. A., Reddy, V. C., Vempaty, A. R., & Zayed, H. R. (2003). The news on health behaviors: Coverage of diet, activity, and tobacco in local newspapers. *Health Education and Behavior*, *30*, 709–722.

Campo, S., & Mastin, T. (2007). Placing the burden on the individual: Overweight and obesity in African American and mainstream women's magazines. *Health Communication*, *22*, 229–240.

Clark, F., & Illman, D. L. (2006). A longitudinal study of the *New York Times* Science Times section. *Science Communication*, *27*, 496–513.

Cohen, E., Caburnay, C. A., Luke, D. A., Rodgers, S., & Cameron, G. T. (2008). Cancer coverage in general audience and Black newspapers. *Health Communication*, *23*, 427–435.

Cohen, E., Shumate, M., & Gold, A. (2007). Anti-smoking media campaign messages: Theory and practice. *Health Communication*, *22*, 91–102.

Fink, E. J., & Gantz, W. (1996). A content analysis of three mass communication research traditions: Social science, interpretive studies, and critical analysis. *Journalism & Mass Communication Quarterly*, *73*, 114–134.

Jones, R. L., & Carter, R. E., Jr. (1959). Some procedures for estimating 'news hole' in content analysis. *Public Opinion Quarterly*, *23*, 399–403.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Newbury Park, CA: Sage.

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, *30*, 411–433.

Lacy, S., Riffe, D., & Randle, Q. (1998). Sample size in multiyear content analyses of monthly consumer magazines. *Journalism & Mass Communication Quarterly*, *75*, 408–417.

Lacy, S., Riffe, D., Stoddard, S., Martin, H., & Chang, K. (2001). Sample size for newspaper content analysis in multiyear studies. *Journalism & Mass Communication Quarterly*, *78*, 836–845.

Long, M., Slater, M., Boiarsky, G., Stapel, L., & Keefe, T. (2005). Obtaining nationally representative samples of local news media outlets. *Mass Communication & Society*, *8*, 299–322.

Mastin, T., Andsager, J., Choi, J., & Lee, K. (2007). Health disparities and direct-to-consumer prescription drug advertising: A content analysis of targeted magazine genres, 1992–2002. *Health Communication*, *22*, 49–58.

Mooney, C. Z., & Duval, R. D. (1993). *Bootstrapping: A nonparametric approach to statistical inference.* Sage University Papers Series. Quantitative Applications in the Social Sciences. Newbury Park, CA: Sage.

Neuendorf, K. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Niederdeppe, J., Davis, K., Farrelly, M., & Yarsevich, J. (2007). Stylistic features, need for sensation, and confirmed recall of national smoking prevention advertisements. *Journal of Communication*, *57*, 272–292.

Pagano, M., & Gauvreau, K. (2000). *Principles of biostatistics*. Belmont, CA: Duxbury Press.

Powers, A., & Andsager, J. (1999). How newspapers framed breast implants in the 1990s. *Journalism & Mass Communication Quarterly*, *76*, 551–564.

R Development Core Team. (2008). R: A language and environment for statistical computing (Version 2.8.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.

Riffe, D., Aust, C. F., & Lacy, S. R. (1993). The effectiveness of random, consecutive day and constructed week sampling in newspaper content analysis. *Journalism Quarterly*, *70*, 133–139.

Riffe, D., & Freitag, A. (1997). A content analysis of content analyses: Twenty-five years of Journalism Quarterly. *Journalism & Mass Communication Quarterly*, *74*, 873–882.

Riffe, D., Lacy, S., & Fico, F. (2005). *Analyzing media messages: Using quantitative content analysis in research* (2nd ed.). Mahwah, NJ: Erlbaum.

Stempel, G. H. III. (1952). Sample size for classifying subject matter in dailies. *Journalism Quarterly*, *29*, 333–334.

Stephenson, M., & Holbert, R. (2003). A Monte Carlo simulation of observable versus latent variable structural equation modeling techniques. *Communication Research*, *30*, 332–354.

Stephenson, M. T. (2002). Anti-drug public service announcements targeting parents: An analysis and evaluation. *Southern Communication Journal*, *67*, 335–350.

Stephenson, M. T., & Quick, B. L. (2005). Parent ads in the National Youth Anti-Drug Media Campaign. *Journal of Health Communication*, *10*, 701–710.
Stryker, J. E., Solky, B. A., & Emmons, K. M. (2005). A content analysis of news coverage of skin cancer prevention and detection. *Archives of Dermatology*, *141*, 491–496.
Stryker, J. (2003). Media and marijuana: A longitudinal analysis of news media effects on adolescents' marijuana use and related outcomes, 1977–1999. *Journal of Health Communication*, *8*, 305–328.
Stryker, J., Emmons, K., & Viswanath, K. (2007). Uncovering differences across the cancer control continuum: A comparison of ethnic and mainstream cancer newspaper stories. *Preventive Medicine*, *44*, 20–25.
Stryker, J., Wray, R., Hornik, R., & Yanovitzky, I. (2006). Validation of database search terms for content analysis: The case of cancer news coverage. *Journalism & Mass Communication Quarterly*, *83*, 413–430.