USING NEWS ABSTRACTS
TO REPRESENT NEWS AGENDAS

By Jill A. Edy, Scott L. Althaus, and Patricia F. Phalen

Many scholars rely upon the Vanderbilt Television News Index and Abstracts *to represent the topics covered by network broadcast news. Earlier research has shown that the* Abstracts *do not adequately capture the evaluative tone of news, but the degree of topical correspondence between the abstracts and the full transcripts of newscasts has never been formally tested. This paper uses content analysis of transcripts of ABC's coverage of the 1991 Gulf War and the corresponding* Vanderbilt Abstracts *entries to assess the relationship between the topical content of newscasts and that of their abstracts. It demonstrates that under the right conditions, the topical content of news can be effectively represented in abstracts, but emerging topics and those not discussed by the White House are likely to be underrepresented in abstracts.*

Introduction

The news indexes and abstracts that are widely available to researchers can greatly simplify content analysis. They are an invaluable resource for describing the news agenda and its transformation over time—provided they contain reliable representations of the news. In fact, many scholars use these sources as proxies for news coverage, particularly when faced with especially large data sets. To evaluate the validity of this practice, the authors have been engaged in a program of research designed to answer two questions: Can abstracts and indexes be used to represent the full text of news, and, if so, under what circumstances?

In earlier work,[1] we demonstrated that the *Vanderbilt Television News Index and Abstracts* do not adequately capture the evaluative tone of news. Research on this aspect of the news is common in scholarship on political communication and journalism, where issues related to "objectivity," balance, and fairness in the news constitute a long-standing concern. In the present study, we turn our attention to the utility value of the *Abstracts* for research in the agenda-setting tradition, which has traditionally focused not on how the news might influence the direction of public opinion through some sort of reporting "bias," but rather on the "status conferral"[2] function of media in which topics in the news

become salient for the public. While more recent debate and scholarship have produced a certain amount of overlap between these two fields of study, wherein framing and "second-level" agenda setting are said to influence both the salience of issues and the direction of public opinion, the theorized mechanisms of influence remain distinct. Thus, the ability of the *Abstracts* to capture the topical content of news must be tested as well.

While agenda-setting research is centrally concerned with topics in the news, other scholarly analyses of news content also use coding categories to document topics over some specific period of time. "Topic" may be the primary focus of a study, as in agenda-setting research, or it could be one of several measured variables. In either case, findings from this study should be of interest to all researchers who conduct content analysis of news.

Literature Review

The long tradition of using news indexes as proxies to measure news content dates back at least to 1973 when Funkhouser used the *Reader's Guide to Periodical Literature* to document issue agendas in mass media.[3] More recently, scholars have turned to the *Vanderbilt Abstracts* to represent the agendas of broadcast news. For example, Zhu and associates used the *Abstracts* along with front pages from the *New York Times* to establish the relative importance of issues in their study of agenda setting and social interaction.[4] And Norris used the *Abstracts* to track changes in the foreign news agenda from the Cold War to the post-Cold War eras.[5]

In addition to addressing this classic form of measurement, our study speaks directly to a more recent development in agenda-setting research: the study of aspect[6] or second-level[7] agenda setting. This line of research examines not the overall agenda of news but rather the relative prominence of the various aspects of particular issues. Here, too, abstracts and indexes have played an important role in theory development. Norris's study can be described from this perspective: she examined the relative prominence of countries and issues within the overall context of foreign news. Johnson used the first sentence of broadcast news abstracts about the 1988 primaries to determine the "dominant theme" of horse-race stories during that election. He identified seven different themes to characterize the election coverage.[8] Dearing used the *Vanderbilt Abstracts* to document aspect agenda setting on the AIDS issue in order to compare the media agenda to the polling agenda on AIDS.[9] What no one has yet documented is whether, and under what circumstances, the *Vanderbilt Abstracts* successfully represent the topical content of news.

Although certain alternatives to the *Vanderbilt Abstracts* are now available to scholars, understanding how the *Abstracts* represent content remains an important task. Riffe and Freitag have documented the rising number of articles published in this journal that rely upon quantitative content analysis,[10] suggesting the growing popularity of this research technique. The *Abstracts* will continue to be a popular resource

among content analysts for at least three reasons. First, full text transcripts have only been archived in databases such as Nexis since the early 1990s, so scholars interested in historical research or analysis of long-term broadcast news trends will continue to rely on the *Abstracts*. Second, access to the *Abstracts* is free, while transcription services like Nexis or the networks themselves demand payment for access. Third, the *Abstracts* are very brief, so if they do represent national broadcast news content accurately, they greatly simplify content analysis of large cases and long time periods. Despite the usefulness of media indexes as proxies for media content (and as tools for locating media content for analysis), little research has been done on the ways these indexes represent media content.[11] Much of this work has been aimed at international and comparative political scientists who attempt to use media content and media indexes to measure the frequency of real-world events,[12] an undertaking that most media scholars already recognize as dubious. In this study, we address ourselves to issues that centrally concern communication scholarship: how and whether the *Abstracts* represent the topical content of the broadcast news programs that are their source material.

After a brief discussion of the archiving and abstracting practices at the Vanderbilt Television News Archives that may influence the ability of the *Abstracts* to accurately represent the content of broadcast news, we turn to a quantitative comparison of abstracts and transcripts of news coverage of the policy debate that preceded the 1991 Persian Gulf War. While our case study cannot, of course, generalize to all uses of the *Abstracts*, our unitizing of the data and our use of inductively derived code categories are typical research techniques. Thus our findings should provide useful guidance about the possibilities and problems of using the *Vanderbilt Abstracts*.

Creating News Abstracts

While the quantitative case study presented below represents a starting point for further research, a description of archiving practices should be useful to scholars employing the *Abstracts* regardless of their case or time frame.[13] The Vanderbilt Television News Archive has recorded the ABC, NBC, and CBS evening news programs since 1968 and the weekend news since 1970. ABC's *Nightline* was added to the collection in 1988, and, in 1989, the Archives began taping CNN's evening news and scheduled network news specials. While archives like UCLA's simply warehouse tapes, the Vanderbilt Archive indexes and abstracts its video collection, thus providing a uniquely valuable service to scholars. All programs are taped in Nashville and, since 1978, at The George Washington University in Washington, D.C. Although this double taping should correct for occasional weekend news preemptions in Nashville, our research reveals that missing weekend news is still a problem for the Archive (see below).

Several elements of the abstracting process are likely to affect research that relies on the *Abstracts* as a proxy for full text or as a means of locating content for subsequent analysis. The director of the archive

described the six story characteristics customarily included in an abstract: "a) who reported the story and from where, b) story length, c) what it was about in general terms, d) who was mentioned, e) who was quoted, and f) what video or audio clips were used."[14] In addition, some abstracting processes have implications for the ability of the *Abstracts* to reflect topical content.

First, our earlier work describes how abstractors "reflect" rather than "quote" the words of sources in the news. This means that *Abstracts* may not contain the specific language used in a story. Topics become problematic because abstractors must sometimes choose among several subjects discussed by a news source and include what they judge to be most important. Thus, not only is the language of news imperfectly represented in the *Abstracts*, the topical content of news may be as well.

Second, because abstractors try to account for the video clips in a report, story topics with accompanying visuals might be more likely to appear in abstract entries than topics without special visuals. This abstracting practice has changed over time, which may complicate historical studies of the news agenda. In the early days of the *Abstracts*, video was only identified as "film shown" or "scenes shown." Over time, video description became more important to the abstract entry. According to Archive director John Lynch, if the content of the story makes the content of the clips obvious, they are not explained in any detail. If the abstractor judges that there is room for error, however, then the content of the clips is specified. For example, if a story about the bombing of Kosovo includes a clip of bombs falling, the entry would simply say "scenes shown." If the clip were of Red Cross workers handing out food, the entry might say "video shown of Red Cross workers."

We should also note that not every source of television news is captured by the Vanderbilt Archives, and therefore by the *Vanderbilt Abstracts*. Although the Archive records CNN's main evening news broadcast, it does not archive all of CNN's content, nor does it archive the content of other cable news stations or satellite news channels. Thus, while using the *Abstracts* to represent the full text transcripts they were drawn from has certain hazards, using them to represent the overall content of television news is likely deeply problematic. That said, broadcast television news continues to significantly outdraw cable news networks for audience share, so were one to think about the most widely shared universe of news discourse, the *Abstracts* may perform reasonably well where they capture broadcast news content reasonably well.

In short, the abstracting and archiving practices at the Vanderbilt Television News Archive can systematically affect the representation of broadcast news topics. In the case study below, we examine how those effects take shape when the *Abstracts* are directly compared to their transcript counterparts.

Method

The comparison between the Vanderbilt Television Abstracts and the transcripts of broadcast news is based on a case study of ABC News' coverage of the policy debate that preceded the 1991 Persian Gulf War. We suspect that given the current state of world affairs, studies of media

coverage of war are likely to proliferate over the next few years, making our choice of case immediately relevant for many researchers. However, like any case study, this one has its limitations. Although war news is a classic news story genre, we cannot assure the reader that the degree of accuracy of war news abstracting that we establish in this study is similar to the accuracy of abstracting for other kinds of news items. Nor can we say for certain that other broadcast networks or CNN are abstracted as accurately (or inaccurately) as the ABC news broadcasts are, though our interviews with the *Abstracts'* editor reveal that the Archive's abstracting practices are the same for all networks. Even less could we claim to ascertain whether the *Abstracts* are an adequate representation of the topical content of all television news, given that many more recently developed news outlets are not archived at Vanderbilt. Finally, it is possible that the quality of the *Abstracts* varies over time and that the *Abstracts* were (or are) more (or less) accurate in eras our case study does not cover. Nevertheless, our case study is quite adequate for our goal of providing evidence to answer a very basic question: Do the *Vanderbilt Abstracts* represent what they claim to represent, that is, the content of their specific source material?

Transcripts of ABC's national newscast, *World News Tonight* (WNT) that aired between 2 August 1990 (when Iraq invaded Kuwait) and 21 February 1991 (when the beginning of the ground war was reported to American audiences) were retrieved from Nexis. The corresponding abstracts were obtained from the Vanderbilt News Archives' Web site. There were 1,204 stories that appeared in both the transcripts and the *Abstracts*. Ninety-three stories appeared only in the transcripts, mostly because they aired on weekends when the newscasts were "bumped" or pre-empted at Vanderbilt's taping sites and thus never archived. Other stories could not be used because ABC extended the format for its broadcast during the Gulf crisis and the *Abstracts* treat many of these extended programs as news specials rather than as newscasts, making meaningful comparisons between transcripts and abstracts impossible.[15]

For each story, a coder assessed the presence or absence of nine topics chosen for their importance to the event: (1) the role of the U.S. or allied nations in providing aid to Iraq prior to the invasion of Kuwait; (2) Congress's power to declare war as it relates to U.S. policy in the Gulf crisis; (3) the likely monetary costs of military action and who will pay them; (4) the apparent effectiveness of economic and diplomatic sanctions against Iraq; (5) the hostages or prisoners held by Iraqi forces; (6) allied or Iraqi casualties; (7) the morale, training, or readiness of allied or Iraqi military forces; (8) tactical reports on clashes between allied and Iraqi forces; and (9) the price or potential scarcity of oil or gasoline. The same coder then applied the same coding scheme to *Abstracts* entries.

Any mention of these topics in a story or an abstract was noted, and as many as four story-level topics could be recorded for each story, although in practice no story was found to contain more than three. To test for reliability, a second coder independently coded 101 randomly selected transcripts. Intercoder reliability tests required that coders use

the same codes for the same stories, not just reach the same aggregate totals in their coding. Intercoder agreement was .960 (Brennan and Prediger's kappa = .957).[16]

When any of the nine topics appeared in news transcripts, they were usually treated at some length. The mean number of words devoted to a topic ranges from a low of 38 for reports of casualties to a high of 149 for stories mentioning the prewar buildup of Iraq's military. A total of 635 story-level topics were coded for the transcript data and 318 for the abstract data. In the analysis that follows, we test whether the *Abstracts* can successfully be used to represent the distribution of these topics in the newscast itself.

While the abstract data appear to capture about half the total number of story topics contained in the transcripts, the correspondence between the topics appearing in the abstracts and those appearing in the transcripts is actually quite low. Table 1 shows the numerical breakdown of topics within the transcript and abstract data. Of particular importance are columns four and five, which report how often abstracts underreport the presence of topics revealed in transcripts and vice versa. Take as an example the row labeled "progress of sanctions," which displays the number of stories in each data set that discuss whether economic sanctions appeared to be having a desired effect on Iraq. Of the 39 stories coded in the transcripts as mentioning this topic, 67% of the corresponding abstract entries (39-13=26, 26÷39=.67) contained no mention of whether sanctions were working. By the same token, 19% of the 16 stories containing this topic in the abstracts had corresponding transcripts with no mention of the topic (16-13=3, 3÷16=.19).

An examination of the fourth column reveals that the 50% of cases where a topic was coded in the transcript of a story but did not appear in the abstract of that story tend to be somewhat evenly distributed across the nine topics. For each topic category, at least half of the mentions of the topic found in the transcripts are missing from the corresponding abstracts, with the most severe underreporting occurring with the progress of sanctions, the readiness of Iraqi or allied troops, and reports of casualties and damage. In the most extreme case, researchers relying on abstract entries to identify stories that mention troop readiness would miss more than three-quarters of the cases contained in transcripts. The fifth column suggests that instances where the abstracts refer to topics not found in the transcripts are somewhat more variable, from none in the case of the prewar buildup of Iraq's military to more than a third of abstracts mentioning the preparedness of allied or Iraqi military forces. The presence of topics in the abstracts that are absent from the transcripts could occasionally be the result of coder error on our part, but they are more likely an outcome of the interpretive process of abstracting. In some cases, the *Abstracts* contain outright errors regarding the content of the news, though the *Vanderbilt Abstracts'* editorial processes make such errors unusual. The most plausible reason for this type of error is that in the process of condensing the transcripts, an abstractor's summary has explicitly referenced a topic that the transcript itself only implies.

TABLE 1

Distribution of Story-level Topics among Abstracts and Transcripts

| Topic	N of Transcript Stories Mentioning Topic	N of Abstract Stories Mentioning Topic	N of Stories with Matching Topics	% of Transcript Stories Missing in Abstracts	% of Abstract Stories Missing in Transcripts
Allied Prewar Buildup of Iraq's Military	9	4	4	.56	.00
Applicability of the War Powers Act	27	18	13	.52	.28
Monetary Costs of the Gulf Crisis	35	19	18	.49	.05
Progress of Sanctions	39	16	13	.67	.19
Hostages or Prisoners of War	169	96	87	.49	.09
Report of Allied or Iraqi Casualties	79	31	24	.70	.23
Readiness of Allied or Iraqi Troops	56	20	13	.77	.35
Damage Inflicted by Allied or Iraqi Forces	73	33	26	.64	.21
Impact of the Gulf Crisis on Oil Supply	149	81	71	.52	.12

While it would appear that the abstracts are randomly filled with significant errors when it comes to reporting the agenda of news, further analysis suggests otherwise. First, the coding of story topics followed a "mere mention" rule that should maximize the apparent discrepancies between abstracts and transcripts. If a long story on Iraqi hostages contains a single sentence about troop readiness, an abstract that fails to mention troop readiness would be counted as "missing" that topic. We might be more concerned with whether the abstracts do a good job of capturing what a story is "about" rather than conveying every mention of minor points that happen to be of concern to the content analyst. It seems likely that the more focused a story is on a single topic, the greater the likelihood that the topic will be mentioned in an abstract entry. Second, if entries tend to report primarily what a story is "about," there is reason to expect that stories appearing later in the broadcast should have fewer topical matches between abstracts and transcripts. To achieve thematic unity in a broadcast, stories appearing later in a newscast may mention topics raised in earlier stories as a way of transitioning to new material. Because of this tendency, topics from earlier in a newscast may appear briefly in later stories but may not be deemed especially significant by abstractors, who are more concerned with capturing dominant themes in a story. In this way, topics not covered previously in a broadcast may be deemed more worthy of inclusion in the abstract entry.

To test these possibilities, we selected stories for which either the transcript or the abstract was coded as containing one of the nine topics (n=683) and used logistic regression to estimate the likelihood that abstract entries report topics mentioned in transcripts. Two predictor variables were used in this analysis: the proportion of the story's running time devoted to the topic[17] and the relative order of the story within the broadcast. Total running time was also included as a control in this analysis. This regression produced significant coefficients in the

FIGURE 1

### Probabilities that Abstract Entries Will Report Topics Mentioned in Transcripts*
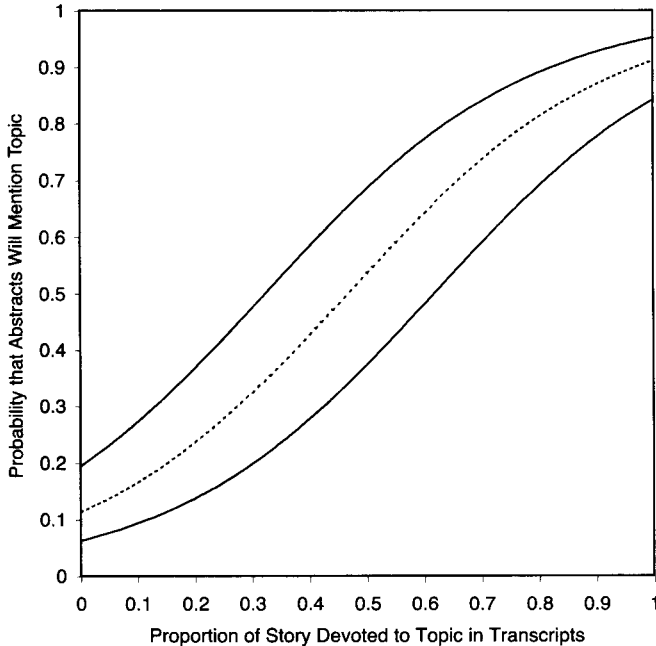
* The upper line is the probability distribution for lead stories in a broadcast, the lower line is the distribution for the closing stories in a broadcast, and the dotted line represents stories in the middle of a broadcast.

expected directions for both variables (3.95 for the ratio of time on topic, $p<.001$; -.05 for story order, $p=.01$), as well as a substantially improved fit to the data (model chi-square=183.67, $p<.001$). As logit coefficients are difficult to interpret on their own, we present the results of this analysis in Figure 1 as the probability distribution of a topical match between abstracts and transcripts. The upper line is the probability distribution for lead stories in a broadcast, the lower line is the distribution for the closing stories in a broadcast, and the dotted line represents stories in the middle of a broadcast.

Figure 1 confirms that the more attention devoted to a topic within a story, the greater the likelihood that the topic will be mentioned in its abstract entry. The probability distribution has a fairly uniform slope across the range of values, with probabilities tapering off only slightly near the upper and lower bounds of the distribution, indicating that the proportion of time devoted to a topic is roughly equivalent to the probability that the topic will be mentioned in a story's abstract. But the predicted probability for a topical match is much higher for stories appearing at the beginning of a broadcast than for those at the end of a broadcast. For example, an abstract entry for a lead story with half

FIGURE 2

Proportions of Story-Level Topics in Transcripts and Abstracts
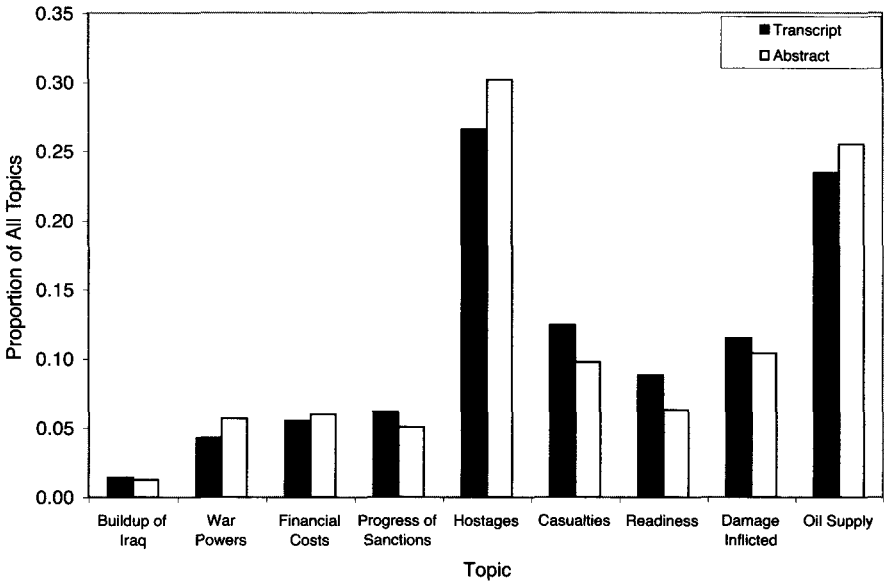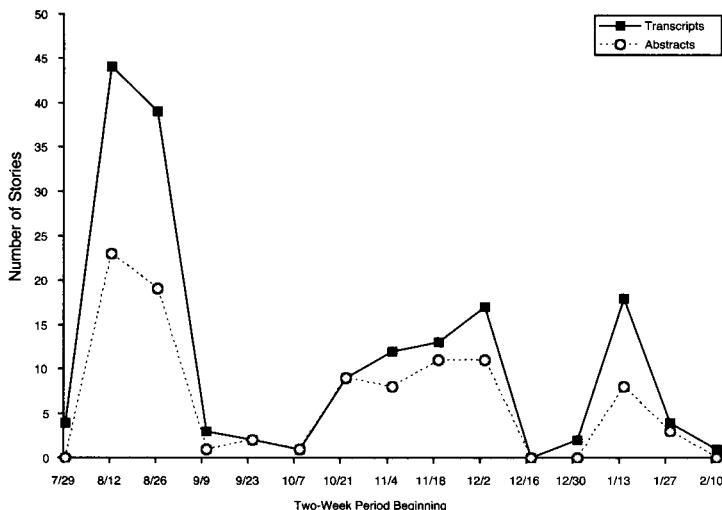

of its time devoted to a topic has a .68 probability of mentioning the topic, dropping to .54 for a middle story and .37 for a closing story. Even when a story is essentially about one topic, abstract entries often fail to convey the topical content of the story that is revealed in transcripts. The predicted probability of a topical match for a lead story with 90% of its time devoted to a single topic is .93, dropping to .87 for a middle story and .77 for a closing story.

Although Table 1 shows key discrepancies between the transcripts and abstracts, Figure 2, which displays the proportional breakdown of topics within the abstracts and transcripts (the proportions within each data set sum to unity), paints a brighter picture. While the abstracts devote slightly more attention to hostages and oil supply issues and slightly less to casualties and troop readiness than the transcripts, the overall distribution of topics is strikingly similar between the two data sources. Correlating the rank order of topics between the abstracts and the transcripts produces a Spearman's rho of .93 ($p<.001$), strong evidence of a near perfect correspondence between the topical distributions of these data sets. It would appear that the *Abstracts* can represent effectively the proportions of coverage devoted to specific topics within the context of a larger news story, a key finding for those studying aspect, or second-level, agenda setting in electronic news.

Since scholars often use abstract data to track the ebb and flow of news coverage over time, we tested whether the abstracts could effec-

tively represent such trends in our case. Our earlier work in this area has shown that as abstract data are disaggregated, their ability to represent the content of news itself declines sharply.[18] Here, we offer a "best case" test of the abstracts' representation of transcript content by examining over-time changes in the number of stories mentioning hostages or prisoners being held by Iraq, the story-level topic with the largest number of data points in both the transcripts (n=169) and the abstracts (n=96). Figure 3 displays these trends.

Although the abstracts register only 51% of the number of stories on hostages found in the transcripts (see Table 1), Figure 3 shows that these missing data do not prevent the abstracts from providing a fairly accurate representation of the trend found in transcript data. The correlation between the two series is an impressive .96. Moreover, the difference between the two trend lines at any given point in time is typically quite small. The close correspondence of these trend lines is also reflected in the slopes obtained by regressing each of the trends on the time variable, which produces a slope of -1.21 for transcript data and -.58 for abstract data. While this difference approaches conventional levels of significance,[19] it is one of magnitude rather than of kind; while the transcript series has a steeper slope than the abstract series, both consistently portray a general decline in the number of hostage stories over time. These data suggest that researchers would draw the same general conclusions about the waxing and waning of hostage stories on the nightly news regardless of whether they used abstract or transcript data.

Discussion   In contrast to our earlier research showing that the *Vanderbilt Abstracts* cannot be used to assess the evaluative tone of news discourse,[20] results of this analysis suggest that the *Abstracts* can be a reasonable proxy for studies of news agendas. The shape of the probability distribution in Figure 2 hints that the tendency of the abstracts to "miss" topical information need not entail serious problems. So long as the various topics of interest to researchers are sufficiently important or broad that a reasonable proportion of stories in which they appear will have them as their focus, the *Abstracts* may represent the news fairly accurately with regard to the proportions of various topics. Whether the *Abstracts* actually *will* give an accurate representation in any particular case is a question that our data cannot answer.

We would offer two major caveats to this general conclusion. First, scholars attempting to track the emergence of a particular topic on the news agenda probably cannot use the *Abstracts* to do so. As we noted, the abstracts have a tendency to miss minor topics in stories, which is conflated with a tendency to miss topics in stories that appear later in the broadcast. The emergence of a new topic in the news may have both of these characteristics: it may play a minor role at first, and if so, it is unlikely to lead the news.

Second, our findings on how the evaluative tone of news is represented in the *Abstracts*[21] suggest that opposition to Administration policy is understated, and we would expect that similarly, news topics not discussed by the administration might be systematically underreported in the *Abstracts*. Thus, scholars using the *Abstracts* as the basis of their analyses may overestimate the amount of control the Administration has over the news agenda. The only way to make sure that the topics of interest to the researcher are important and broad enough to be tracked successfully by the *Abstracts* is to examine a sample of transcripts or videotapes and compare them to corresponding abstract entries.

Our study represents only a place to start. Dozens of studies rely on a variety of news abstracts to represent news content in a variety of ways, and our results show only that this methodological approach has some validity under some circumstances. We encourage all researchers who use abstracts and indexes as proxies for full texts of news to test the quality of those proxies by directly comparing the abstracts and indexes to the full text of news. Not only will this confirm the validity of their own work, it will build up a critical mass of cases that can help us, as a field, determine the conditions under which proxies can and cannot be used as substitutes for the actual content of the news.

NOTES

1. Scott L. Althaus, Jill A. Edy, and Patricia F. Phalen, "Using the Vanderbilt Abstracts to Track Broadcast News Content: Possibilities and Pitfalls," *Journal of Broadcasting & Electronic Media* 46 (September 2003): 473-92.

2. Paul F. Lazarsfeld and Robert K. Merton, "Mass Communication, Popular Taste, and Organized Social Action" in *The Communication of Ideas*, ed. Lyman Bryson (New York: Harper, 1948), 95-118.

3. Ray Funkhouser, "The Issues of the Sixties: An Exploratory Study in the Dynamics of Public Opinion," *Public Opinion Quarterly* 37 (spring 1973): 62-75.

4. Jian-Hua Zhu, James H. Watt, Leslie B. Snyder, Jingtao Yan, and Jiang Yansong, "Public Issue Priority Formation: Media Agenda Setting and Social Interaction," *Journal of Communication* 43 (winter 1993): 8-29.

5. Pippa Norris, "The Restless Searchlight: Network News Framing of the Post-Cold War World," *Political Communication* 12 (October-December 1995): 357-70.

6. Frank R. Baumgartner and Bryan D. Jones, *Agendas and Instability in American Politics* (Chicago: University of Chicago Press, 1993).

7. Maxwell E. McCombs and Salma I. Ghanem, "The Convergence of Agenda Setting and Framing," in *Framing Public Life: Perspectives on Media and Our Understanding of the Social World*, ed. Stephen D. Reese and Oscar H. Gandy (Mahwah, NJ: Lawrence Erlbaum Associates, 2001), 67-82.

8. Thomas J. Johnson, "The Seven Dwarfs and Other Tales: How the Networks and Select Newspapers Covered the 1988 Democratic Primaries," *Journalism Quarterly* 70 (summer 1993): 311-21.

9. James W. Dearing, "Setting the Polling Agenda for the Issue of AIDS," *Public Opinion Quarterly* 53 (autumn 1989): 309-329.

10. Daniel Riffe and Alan Freitag, "A Content Analysis of Content Analyses: Twenty-Five Years of Journalism Quarterly," *Journalism & Mass Communication Quarterly* 74 (winter 1997): 873-83.

11. Scott L. Althaus, Jill A. Edy, and Patricia F. Phalen, "Using Substitutes for Full-Text News Stories in Content Analysis: Which Text Is Best?" *American Journal of Political Science* 45 (July 2001): 707-723; Althaus, Edy, and Phalen, "Using the Vanderbilt Abstracts"; Cheryl Zollars, "The Perils of Periodical Indexes: Some Problems in Constructing Samples for Content Analysis," *Communication Research* 21 (December 1994): 698-717.

12. See, for example, Roberto Franziosi, "The Press as a Source of Socio-Historic Data: Issues in the Methodology of Data Collection from Newspapers," *Historical Methods* 20 (winter 1987): 5-16; Robert Burrowes, "Mirror, Mirror on the Wall: A Comparison of Event Data Sources," in *Comparing Foreign Policies: Theories, Findings and Methods*, ed. James N. Rosenau (New York: John Wiley & Sons, 1974), 383-406; Gary Hoggard, "Differential Source Coverage in Foreign Policy Analysis," in *Comparing Foreign Policies: Theories, Findings and Methods*, ed. James N. Rosenau (New York: John Wiley & Sons, 1974), 353-81; John Woolley, "Using Media Based Data in Studies of Politics," *American Journal of Political Science* 44 (January 2000): 156-73.

13. For more information on archiving practices at Vanderbilt, see Althaus, Edy, and Phalen, "Using the Vanderbilt Abstracts."

14. Althaus, Edy, and Phalen, "Using the Vanderbilt Abstracts."

15. Althaus, Edy, and Phalen, "Using the Vanderbilt Abstracts."

16. Robert L. Brennan and Dale J. Prediger, "Coefficient Kappa: Some Uses, Misuses, and Alternatives," *Educational and Psychological Measurement* 41 (fall 1981): 687-699. This measure accounts for the possibility of coders agreeing by chance. Brennan and Prediger's kappa for individual coding categories ranges from 1.0 to .97, slightly better than the .957 figure for overall coding reliability because coding for presence versus absence of a particular topic is a simpler coding task than is choosing among several topical categories.

17. This measure was constructed by dividing the number of seconds devoted to a topic in a story by the total running time of the story (in seconds). Measures of total running time were derived from the *Vanderbilt Abstract* entries, which record time boundaries for each story. In cases where the division of stories in the abstracts did not match the procedure for unitizing being followed by the coder, the coder was instructed to estimate the story's running time by dividing the total number of words in the text defined in the abstract entry by the total number of words in the coder's story unit, which in discrepant cases was always a fraction of the abstract's story unit. This proportion was then multiplied by the abstract's time estimate to determine the coded story's running time. As the original videotapes of newscasts were unavailable for measuring the amount of time within a story that was devoted to a topic, estimates were calculated based on the number of words in a transcript that were devoted to a topic. A sample of 51 stories was selected to compare the number of transcript words contained in a story to the running time of the story in seconds. Dividing the story's total running time by the total number of words it contained and then averaging across the 51 cases produced an estimated mean of one word per .32 seconds (s.d. =.08). This time estimate was then multiplied by the number of words devoted to a topic to produce an estimate of the amount of time devoted to a topic.

18. Althaus, Edy and Phalen, "Using Substitutes for Full-Text News Stories in Content Analysis."

19. $t(26)=1.62$, $p=.12$; the t-test for the significance of differences between regression coefficients from independent samples is described in Jacob Cohen and Patricia Cohen, *Applied Multiple Regression/Correlation for the Behavioral Sciences* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1983), 55-56.

20. Althaus, Edy, and Phalen, "Using the Vanderbilt Abstracts."

21. Althaus, Edy, and Phalen, "Using the Vanderbilt Abstracts."