A study investigated the possibility that news sources account for the pattern of journalistic topic selection. It was hypothesized that the proportions of content in raw news materials made available to reporters by sources would be similar to the proportion of content found in published stories, and that the proportions of sources for these raw materials would be dissimilar to the proportion of sources found in published stories. The study focused on sources and media coverage for a midwestern statehouse. Press materials from the statehouse were analyzed and coded for a six-week period, as were statehouse stories published in the state's four major newspapers for a four-week period. The analyses indicated that proportions of content were similar across the raw materials and the published stories, supporting the first hypothesis. The results also showed that the proportion of specific sources for raw materials was different from the proportion of sources in the published stories, supporting the second hypothesis. The data suggest that journalists and their sources share notions about what aspects of a state bureaucracy are worth paying attention to. The disproportionate use of sources outside the statehouses cited in the published stories may represent journalistic attempts to achieve balance in reporting. (HTH)
Accounting for Patterns of Topic Selection In Statehouse Reporting

by

Sharon Dunwoody
Assistant Professor
School of Journalism and Mass Communication
University of Wisconsin
Madison, WI 53706

and

Steven Shields
Graduate Student
School of Journalism and Mass Communication
University of Wisconsin
Madison, WI 53706

Presented in an open paper session of the Communication Theory & Methodology Division, Association for Education in Journalism and Mass Communication annual convention, August 1984, Gainesville, FL.
Accounting for Patterns of Topic Selection
in Statehouse Reporting

Over the years, some very different kinds of research projects have all pointed to the notion that journalists do not select information idiosyncratically, that instead they pick topics in patterned ways.

But accounting for the patterns has been problematic. Traditional gatekeeping studies assumed that the best predictors were news values assimilated as part of the journalistic trade. More recently, attention has shifted to how organizational demands can determine patterns of information seeking.

Yet these studies have all kept the locus of attention on the reporter or media organization. Another group of researchers has shown—some of them inadvertently—that the locus might be better placed on information sources.

Perhaps most prominent among this group at present is sociologist Mark Fishman, who argues that patterns of news selection are best predicted by "phase structures." He defines a phase structure as "a very general scheme employed in everyday thought for picturing events in the context of successively developing phases." Everyone employs these frames of reference to make sense of his or her environment, but Fishman suggests that journalists don't utilize their own, personally developed phase structures to make information decisions; instead, he argues, they assimilate the frames of reference of their sources. "Journalists perceive events in substantially the same way that beat agency officials formulate their own and other persons' activities as events," he notes.

The upshot of such sharing is that journalists' and sources' notions of what about a bureaucratic structure are worth attending to should be very similar.
To be more specific, sources should be influential in defining the boundaries of possible story topics; they should set the journalistic agenda by, a priori, establishing what aspects of bureaucratic structure belong in the news columns.

This study investigates the possibility that sources indeed can account for patterns of journalistic information seeking by comparing information provided by sources with stories written in a specific setting: a statehouse.

A Look at the Literature

Many researchers have reacted cautiously to Fishman's findings because he bases his generalizations on a case study done at a single newspaper. And as an ethnomethodologist, his data are not amenable to statistical tests.

Yet, if one broadens one's notion of what constitutes a "source" of information for journalists, then one can find a number of empirical studies in the literature that support the argument that sources not only establish the boundaries of possible story topics but also govern the attention paid within those boundaries to varying categories of topics.

For example, when White examined his single wire service editor, he focussed on the idiosyncrasies of the person's selection process. Yet nearly 30 years later, when Hirsch reexamined the data from that study, he found that the types and proportions of stories chosen by this individual were nearly identical to the types and proportions of stories supplied by the wire services. The wire service editor had duplicated the proportionate distribution of stories offered.

Similarly, McCombs and Shaw reanalyzed the White data and calculated a Spearman's rho of .64 between ranks of seven news content categories supplied by the wire services and ranks of stories selected by Mr. Gates. They also
reanalyzed data from a replication of White's study\textsuperscript{7} and calculated a Spearman's rho of .80.\textsuperscript{8}

Gold and Simmons found a .915 coefficient of concordance between the ranks of proportions of content supplied by the AP and ranks of proportion of content utilized by 24 Iowa daily newspapers.\textsuperscript{9} And although Stempel found agreement between newspaper use of wire stories and the universe of available wire stories to be low in one study,\textsuperscript{10} in another his data indicate that newspaper use of wire stories was proportionate to the available topics generated by AP.\textsuperscript{11}

Most recently, Whitney and Becker, in a field experiment setting, found that both print and broadcast editors selected stories proportionately from an available pile; the authors calculated a Pearson correlation coefficient between number of items incoming and number selected in the seven topic categories of .71.\textsuperscript{12}

All these studies support the argument that the wires—as sources of stories—not only set the boundaries within which selections are made but also "cue" journalists as to the "proper" distribution of stories within those boundaries. In Fishman's term, the wires seem to be the source of the phase structure utilized by these editors.

But one might argue that the frames of reference of wire service "sources" are readily accepted by journalists while the frames of reference of other types of sources may not be. Journalists and traditional sources are supposed to have something of an adversarial relationship, after all. So one might expect to find the phase structure assimilation scenario falling apart in that relationship, Fishman's arguments notwithstanding.

It is here that empirical evidence becomes scant. Few studies have attempted to compare the raw input from sources with journalistic topic choices
in a way that allows any conclusions more detailed than general dependency
statements. But the few that have been published do find dependence by
journalists on source input.

For example, Sachsman monitored environmental press releases sent to Bay Area
reporters and then compared the releases with the stories published. Of the
stories produced locally, more than half had been influenced by public relations
efforts, he concluded.13

And Hale, in a study of newspaper coverage of state supreme court decisions
in California, found the media to be very dependent on press releases issued by
the court. Of the 88 court decisions mentioned in press releases, 66% were
reported by at least one of the 10 newspapers in this study. In contrast, of
the 51 court decisions that were not mentioned in press releases, only 8%
received any newspaper coverage.14

The study reported here was designed in part to fill in this empirical gap.
In the following pages we report on a comparison of raw input from statehouse
sources with the stories ultimately produced by statehouse reporters.

Hypotheses

If Fishman's notion of reporter assimilation of source frames of reference
is correct, then we should find similarities between the ranks of proportions of
content in source raw materials and the ranks of proportions of content in
published stories. We were able to find no studies that examined this
relationship for any type of reporting, much less for statehouse reporting. The
only study that even came close was a study by Baker and Walter, who content
analyzed coverage of state government in six Wyoming newspapers and found that
the newspapers gave proportionately the same amount of space to each legislator. But this study attempted no examination of the raw materials upon which the stories might have been based. Yet the consistent findings of the wire service studies, coupled with the small number of raw input studies that supported the notion of dependency of journalists on sources, prompted us to posit:

H1: The ranks of proportions of content in raw materials made available to reporters by statehouse sources will be similar to the ranks of proportions of content found in published stories.

But would a shared frame of reference also result in a use of similar sources? In our thinking, we returned to the traditional notion of the adversarial relationship between sources and journalists. Journalists may very well adopt the "phase structure" of their sources almost unconsciously. But such normative practices as the need for "balance" in stories might mandate a departure from the source's frame of reference. A journalist may readily accept a bureaucracy's notion of what is important and thus what needs to be communicated to the public. But such agreement on a topic may not limit the reporter's search for sources outside the bureaucracy deemed appropriate by journalistic norms.

At least one study indirectly documents this point by finding a disproportionate use of sources. Jones and Meadows, in a study of the raw information available to and utilized by a small group of science reporters in Great Britain, found that reporters were less suspicious of certain types of sources (the scientific community) than others (industry), and utilized information disproportionately on that basis. So we hypothesized:
H2: The ranks of proportions of sources in raw materials made available to reporters by statehouse sources will be dissimilar to the ranks of proportions of sources found in published stories.

Method

This study focussed on media coverage of a statehouse for a number of reasons. A primary one was that statehouse reporting was the focus of a larger research project being conducted by a graduate class at the School of Journalism and Mass Communication, making it possible to obtain the type of data needed for this analysis. But perhaps just as importantly, focussing on a single statehouse (Wisconsin) gave us a setting in which the relevant sources and reporters were clearly defined. And it also made possible the collection of raw source materials in an efficient manner.

This analysis required the collection of two types of materials: raw materials from sources and published stories from journalists. Collection took place during four weeks in spring 1982: two weeks when the state legislature was not in session and two weeks when it was in session.

Raw materials. Both Dunn and Nimmo found in studies of state and national government, respectively, that governmental officials rely heavily on written documents when communicating with journalists. So we obtained permission from the statehouse press room to, for all practical purposes, have our own mailbox there. Consequently, we gathered 698 pieces of printed materials during a six-week period (a four-week subset of these materials—those corresponding to the dates of the published stories—will be utilized in this paper). Four coders content analyzed the materials, identifying the type of document, its content and its source, among other variables.
Published stories. Because of the difficulties of retrieving and archiving broadcast material, we confined our analysis in this stage of the study to newspapers. More specifically, we concentrated on the stories published during the four-week period by the four major newspapers in Wisconsin: Milwaukee Journal, Milwaukee Sentinel, Wisconsin State Journal and The Capital Times. The latter two are published in Madison.

We concentrated on only four newspapers because these organizations maintained the most comprehensive bureaus at the statehouse, with individuals on hand on a daily basis. And these newspapers alone churned out such a high volume of stories that adding additional newspapers would have overwhelmed our available resources at the content analysis stage.

We relied on a statehouse clipping service, the Capitol Headliners, to collect the stories. While clipping services are sometimes unreliable, spot checks of this service indicated it was missing stories only intermittently.

Ultimately, 1246 published stories were collected from the four-week period in question and were analyzed by one of the authors. As with the raw materials, stories were coded for topic, type of story, general content, and sources.

Results

As indicated above, our collection procedures resulted in large numbers of both raw source materials and published stories. A brief descriptive analysis of the two collections follows:

Raw materials. Of the 698 pieces coded, approximately 2/3 fell into the four-week time period we wished to deal with in this paper. But almost a third of the 698 documents were undated, making it impossible to situate them reliably
in time. Once we eliminated the undated documents along with the dated documents from the two-week period not under study, we were left with 266 raw source items.\textsuperscript{21}

Not surprisingly, the vast majority of these items came from various individuals and components of the state government apparatus. However, 11.7\% of the pile came from persons or organizations outside state government. These groups included such organizations as the Sierra Club, citizen tax groups, etc.

The majority of materials (64.7\%) were designed for reporters; most of them were press releases of one sort or another. Most of the remaining raw items were paper communications generated by the state governmental bureaucracy. Reporters' mailboxes often were crammed with notices of hearings and with sundry official documents.

Approximately 26\% of the raw materials offered information about upcoming events such as meetings or hearings. Another 33\% provided followup information about events that had taken place; for example, a licensing board may have issued a press release detailing the actions at its last meeting. And 37\% of the raw materials dealt in a general way with issues. This last group included legislative position papers, newsletters, court and hearing testimony, etc.

\textbf{Newspaper stories}. Of the 1246 stories collected from the four newspapers during the four-week period, a majority (69.3\%) were datelined Madison. So it is with these 863 stories that our analysis will deal.\textsuperscript{22}

As indicated in Table 1, the stories were distributed fairly evenly among the four newspapers. Nearly 30\% of the stories carried either AP or UPI wire service datelines; the rest were generated by staff or stringers. The vast majority of stories (92.2\%) came from the news columns. The remaining handful
were divided among stories labeled "analysis" (2.9%), editorials (2.7%) and columns (1.9%). Three stories (0.3%) were unidentifiable.

--

TABLE 1 ABOUT HERE

--

The sources of information for these stories came almost entirely from within state government; in the median story, state government sources constituted 99.6% of all the sources mentioned. The average story utilized two sources.

Before discussing the hypothesis tests, we should note that, interestingly, there was little variation in the characteristics of either the raw source information or the published stories across the out-of-session/in-session weeks. We had expected to find some significant differences in raw material output and in story production, but we found none. Consequently, the two-week periods are aggregated in our analyses.

Hypothesis 1 suggested that we would find similar rankings of proportions of content across the raw materials and the published stories. We collapsed content into seven categories. And, as Table 2 indicates, the data support the hypothesis. Regardless of whether one compares the published stories with all raw materials or only with those raw materials designed for the press, Spearman's rho is high and statistically significant (in both cases, $r_s=.929$, $p=.01$).
Hypothesis 2, on the other hand, suggested that the ranking of sources utilized in published stories would be dissimilar to source rankings in the raw materials. Sources were collapsed into six categories, and, as Table 3 indicates, the rankings produce Spearman rho's that are not statistically significant. Consequently, hypothesis 2 is supported.

The major difference between rankings of sources in the raw and published materials can be found with one category: external sources. While external sources constituted only a little more than one in 10 of the raw materials available to statehouse reporters, they constituted nearly a quarter of the sources mentioned in the stories analyzed.

Discussion

These data lend empirical support to Fishman's suggestion that journalists and sources share notions about what aspects of a bureaucratic process are worth paying attention to. During the time period studied here, the "state government world" presented in stories by the four major newspapers in the state in a gross way resembled the world offered up to journalists in the plethora of print materials made available to them. The patterns in topic selection could be predicted by knowing the patterns of available information.
However, our finding of different patterns of source use between the raw and published materials suggests to us that the power of sources to set journalistic agendas can perhaps be limited by normative journalistic practices. While the stories analyzed here depended heavily on state government sources, they utilized sources outside government in a way that was disproportionate to external source availability in the raw information. One way to interpret such a pattern is to suggest that journalists were making decisions to step outside state government frames of reference in some source selections.

Earlier, we had suggested that disproportionate use of sources might be traceable to attempts by journalists to achieve "balance" in stories. One way of examining that suggestion on a post-hoc basis in these data was to hypothesize that external sources would be more likely to appear in stories containing conflict than in other types of stories. We had coded for the presence of conflict in the published stories, and Table 4 presents evidence that, indeed, the conflictive story was more likely to contain external sources than was the story that contained no conflict. The chi square test indicates

TABLE 4 ABOUT HERE

that the pattern of proportions in the table is not due to chance.

While these data do seem to support the argument that journalists select information in a patterned way and that a major determinant of that pattern may be the frames of reference utilized by sources, we must caution the reader that the simplicity of our analyses do not permit us to rule out the effects of confounding variables. For example, it could be the case that both state
governmental institutions and reporters are attending to "important events" in the environment quite independently of each other and that the existence of such events thus accounts for the apparent similarities in general content rankings. In this session, for example, the state's budget was a major issue. One might argue that the existence of a need to establish a budget established the agendas for both legislators and reporters.

Secondly, the reader must remember that patterns of ranks can change as one changes the number of rankings in a set. Our seven content categories in this study were the end products of a recoding process that made conceptual sense to us. But had we ended up with 10 categories, would we have found similarly high Spearman rho's? Such a question puts a burden on the researchers to conceptually defend the categories created; we must examine these categories more carefully as a result.

Still, the rankings unearthed in this analysis—combined with the findings of earlier studies mentioned above—point in a tantalizing way toward some level of similarity in the way that sources and reporters make judgments about what is important. If, in fact, journalists do rely on sources as general topic agenda-setters, then one might raise the argument that the world presented to us by the press indeed mirrors a reality "out there." It may be reality constructed by the providers of information.
Footnotes


3Mark Fishman, Manufacturing the News (Austin, TX: University of Texas Press, 1980), pg. 55.

4Fishman, op. cit., pg. 54.

5White, op. cit.


Whitney and Becker, op. cit.


These data were gathered as part of a larger research project that involved observation of reporters in the statehouse press room, content analysis of raw source materials available to the reporters, content analysis of the stories produced, and interviews with reporters. The last part of the design—the interviews—is still in the planning stages as of this writing. Individuals involved in the design and execution of the project included the authors, as well as graduate students Mark Doremus and Christina Pheley. We extend our thanks to these other individuals for their work.

Intercoder reliability was tested by examining the existence of agreement or disagreement between all possible pairs of coders across all variables for a random sample of the materials. Initially, we found that the proportion of paired coding agreements ranged from 100% to 61% over the 10 variables, with an average of 80% coder agreement overall. This prompted us to examine the variables that produced low coder agreement and to take two actions: (1) Coding schemes for some variables were reworked and the coding redone; and (2) Content variables were collapsed (see footnote #23), thus increasing coding agreement for those items.

At this stage, only one individual was involved in coding materials. Reliability was examined by having a second individual code a random sample of stories and then comparing the coding schemes of the two individuals. The two coded identically 80% of the time.

Although the 266 documents constitute only 38% of the original collection, we found the characteristics of this subset to be almost identical to the characteristics of the total n. For purposes of comparison, however, we will continue to utilize the 266 in our analysis.

We utilized only those stories generated in Madison in order to compare “like with like.” Although stories about state governmental activities were being generated from a variety of sites, including Madison, it was our view that the raw materials provided to Madison-based reporters could only be properly compared with those stories generated in Madison.

In the initial coding stages, the number of content categories for the raw materials reached 189. We did this in order to be as precise as possible at first. At the second stage, the categories were collapsed to 29. The final stage of recoding saw the collapse of the 29 categories to eight, one of which was an eclectic category that was eliminated in the subsequent analyses. The
final seven categories, and some examples of their contents, follow:

**Personnel regulation:** Information about decisions of such regulatory agencies or committees as the Pharmacy Examining Board, the Board of Examiners in Chiropractic, and the Real Estate Broker's Board.

**Social services and education:** Unemployment, aging, minorities, health, education.

**Economics and taxes:** Business and economic development issues, state tax and finance issues.

**Crime and legal matters:** Crime and prison issues. Most visible during the study period was debate on the location of a new state prison.

**Environmental issues:** Land use, foresting, energy and utility issues. Particularly visible during the study period was Project ELF, the federal government's attempt to build a radar system in parts of Wisconsin and the Upper Peninsula of Michigan.

**Political structure:** Internal legislative logistics, political party structure, election issues, city and town affairs.

**Transportation issues:** Roads, aviation.

24Conflict was operationally defined as "an explicit verbal, ideological or physical clash between or among persons, groups or institutions."

18
<table>
<thead>
<tr>
<th>Newspaper</th>
<th>Number of stories</th>
<th>%tage of all stories</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Capital Times</td>
<td>204</td>
<td>23.6%</td>
</tr>
<tr>
<td>Wisconsin State Journal</td>
<td>252</td>
<td>29.2%</td>
</tr>
<tr>
<td>Milwaukee Journal</td>
<td>233</td>
<td>27.0%</td>
</tr>
<tr>
<td>Milwaukee Sentinel</td>
<td>174</td>
<td>20.2%</td>
</tr>
</tbody>
</table>

| Total                  | 863               | 100.0%               |
Table 2
Comparison of Content Rankings
Between Raw Source Material and Published Stories

<table>
<thead>
<tr>
<th>Content area</th>
<th>All Raw Info %tage</th>
<th>Raw Info Designed for Press %tage</th>
<th>Published stories %tage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personnel regulation</td>
<td>21.4 1</td>
<td>27.4 1</td>
<td>21.3 2</td>
</tr>
<tr>
<td>Social services and education</td>
<td>18.0 2</td>
<td>17.3 2</td>
<td>21.6 1</td>
</tr>
<tr>
<td>Economics and taxes</td>
<td>13.9 3</td>
<td>14.9 4</td>
<td>13.9 4</td>
</tr>
<tr>
<td>Crime and legal matters</td>
<td>13.5 4</td>
<td>16.1 3</td>
<td>15.4 3</td>
</tr>
<tr>
<td>Environmental issues</td>
<td>11.7 5</td>
<td>10.7 5</td>
<td>13.5 5</td>
</tr>
<tr>
<td>Political structure</td>
<td>10.9 6</td>
<td>3.6 7</td>
<td>8.9 6</td>
</tr>
<tr>
<td>Transportation issues</td>
<td>8.6 7</td>
<td>10.1 6</td>
<td>3.8 7</td>
</tr>
<tr>
<td>Other issues</td>
<td>1.9</td>
<td>---</td>
<td>1.4</td>
</tr>
<tr>
<td>n=266</td>
<td>n=168</td>
<td>n=863</td>
<td></td>
</tr>
</tbody>
</table>

Comparison of all raw information ranking and published story ranking: \( r_s = .929, n=7, p=.01 \)

Comparison of ranking of press-designed raw information and ranking of published stories: \( r_s = .929, n=7, p=.01 \)
### Table 3
Comparison of Source Rankings
Between Raw Source Material and Published Stories

<table>
<thead>
<tr>
<th>Sources</th>
<th>All raw info</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%stage</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive branch (governor)</td>
<td>47.4</td>
</tr>
<tr>
<td>Assembly</td>
<td>23.3</td>
</tr>
<tr>
<td>Senate</td>
<td>13.9</td>
</tr>
<tr>
<td>External organizations</td>
<td>11.7</td>
</tr>
<tr>
<td>Joint committees</td>
<td>3.0</td>
</tr>
<tr>
<td>Judiciary</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>n=266</td>
</tr>
</tbody>
</table>

1For each story, the top three sources were identified. Thus, in the 863 stories with Madison datelines, we coded mentions of 1404 sources.

Comparison of all raw information ranking and published story ranking: $r_s=.773$, n=6, n.s.

Comparison of ranking of press-designed raw information and ranking of published stories: $r_s=.400$, n=4, n.s.
### Table 4
Relationship Between Presence of Conflict in Stories and External Source Use

<table>
<thead>
<tr>
<th>Presence of conflict?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>External</td>
<td>n=477</td>
<td>n=769</td>
</tr>
<tr>
<td>External</td>
<td>27%</td>
<td>16.8%</td>
</tr>
<tr>
<td>State govt.</td>
<td>73%</td>
<td>83.2%</td>
</tr>
</tbody>
</table>

100.0%                  100.0%

1All 1246 stories were utilized in this analysis.

Chi square = 18.288, 1 d.f., p<.001