The Choice Gap: The Divergent Online News Preferences of Journalists and Consumers

Pablo J. Boczkowski
Department of Communication Studies
Northwestern University

Limor Peer
Institution for Social and Policy Studies
Yale University

Working paper

Please do not quote or disseminate without authors’ permission.

Authors’ note: Research for this paper has been supported by a grant to both authors from the John S. and James L. Knight Foundation as part of the Carnegie-Knight Initiative on the Future of Journalism Education, an Innovation Fund award to the first author from Northwestern University’s School of Communication and additional funding from Northwestern University’s Media Management Center. Work on this project was conducted at Northwestern University. We would like to acknowledge the valuable contributions received from our research assistants: Dan Ellman, Elena Fields, Hilary Gifford and Matt Weir. We also thank Jim Ettema, Rich Gordon, Jay Hamilton, Eszter Hargittai, Bob Hariman, Eugenia Mitchelstein, Dan O’Keefe, Paul Karoff, Joe Turow, and seminar participants at Harvard University and the 2008 annual meeting of the International Communication Association.
Abstract

This study examines online news along the hard-soft content dimension, and determines whether there is a gap between the top news choices of journalists and the stories most often chosen by consumers on the same sites. The study’s design improves upon most of the relevant scholarship by using the story as the unit of analysis, examining stories both in terms of subject matter and story-telling format, and attempting to disentangle the influence of journalists’ and consumers’ choices on each other. Results of the analysis show that (a) journalists’ choices are substantively “soft” in terms of what the stories are about but not in terms of how they are told; and, (b) there is a gap between these choices and consumers’ choices that is uniform and relatively strong in terms of the subject of the stories, but localized and comparatively weaker in terms of how these stories are told. The paper discusses the implications of this choice gap on such issues as the dynamics of agenda setting, the prospects for consumer-authored content, and the watchdog function of the media in liberal democracies.
The Choice Gap:
The Divergent Online News Preferences of Journalists and Consumers

There has been growing agreement among media scholars about a trend toward the “softening of news,” for lack of a better phrase (Baum, 2002; Bennett, 2003a; Project for Excellence in Journalism, 1998; Patterson, 2000; Schudson, 2003; Zaller, 2003). Research has often identified market forces as a critical driver of this trend (Baum, 2003; Delli Carpini & Williams, 2001; Hamilton, 2004; Kalb, 1998; Prior, 2007). The reduced competition enjoyed by American media organizations during most of the twentieth century allowed them to combine strong economic performance with an editorial supply marked by a journalistic preference for hard news focused on public affairs topics (Fishman, 1980; Gans, 2004; Kaniss, 1991; Sigal, 1986; Tuchman, 1979). This situation persisted for decades, despite perception of a gap between this type of news and consumption patterns that expressed only moderate interest in it (Bogart, 1989). With recent increase in competition, companies have paid more attention to consumers’ preferences and softened their supply of news accordingly (Klinenberg, 2005; McManus, 1994; Mindich, 2005; Underwood, 1993). However, scholars have debated whether this trend decreases the gap by giving consumers more of what they supposedly want, or creates a new gap by underserving what could be seen as the main public for news—those highly interested in public affairs stories (Baum, 2002; Hamilton, 2004; Patterson, 2000, 2003; Prior, 2002).

Three limitations have prevented settlement of this ongoing debate. First, although consumer choices are increasingly disaggregated at the story level, most research still relies on survey, circulation, and ratings data that measure more aggregate choices. Second, although there is consensus that the news has gone soft, scholars have disagreed on the nature of the trend. Some focus on what stories are told, others on how they are told, and a third group on a
combination of these two foci. Third, studies have suffered from confounding factors related to the influence of journalists’ and consumers’ choices on each other.

This paper engages with this ongoing debate by examining the character of the top news choices of journalists on four web sites on the hard-soft dimension and by determining whether there is a gap between these choices and the top choices made by the consumers of these sites. Online news sites provide a fertile terrain to look at these issues because the Web is a highly competitive information environment that will likely have increased influence on the future of the industry. Furthermore, this paper presents results of a study designed to help overcome the three shortcomings identified above. First, the story is used as the unit of analysis to have a more fine-grained picture of consumers’ choices. Second, the analysis discriminates between what stories are chosen and how they are told to have a more precise characterization of the object of study. Third, it contrasts all the stories collected with a subset of stories in which the mutual influence of journalists’ and consumers’ choices is likely to be less pronounced. This subset includes stories that journalists display prominently even though they have low levels of demand among consumers and stories with high levels of demand among consumers despite the fact that journalists do not display them prominently. Results show a soft news supply in terms of the subject of news stories, but not in terms of how they are covered. They also show a uniform and sizeable gap between journalists’ and consumers’ choices when the subject of news is concerned, but a localized and lesser gap regarding the storytelling format.

The analysis presented in this paper makes methodological and conceptual contributions. The methodological contributions center on overcoming the limitations mentioned above. The conceptual benefits derive from a better understanding of what a supply of soft news consists of, and of exactly how the news choices of journalists and consumers diverge. In turn, these benefits
help make sense of transformations in the news industry and its publics, such as the dynamics of
agenda setting, the prospects for consumer-authored content, and the watchdog function of the
media in liberal democracies.

Theoretical Considerations

There is growing consensus among media scholars and analysts that the news has “gone
soft.” There were already signs of this trend in studies that were published decades ago. For
instance, the report, *The age of indifference* (Times Mirror Center, 1990), maintains that “even
some more traditional media have adopted a softer news focus in response to diminishing levels
of interest in hard news” (p. 2). In their study of television news from 1972 to 1987, Scott and
Gobetz (1992) argued that “there has been a slight tendency for television network news
programs to increase the amount of soft news presented mostly during the last one-third of the
newscast” (p. 406). The trend seems to have intensified recently, affecting both news products
and how journalists experience their production. Thus, to Bennett (2003b), “what has happened
to the news in the past twenty years is that it has shifted in the direction of soft news and
sensationalism” (p. 131). Moreover, in the preface to the twenty-fifth anniversary edition of
*Deciding what’s news*, Gans (2004) reflected on that quarter century and concluded that “the
change in news content that exercises journalists the most is a decrease in hard… news” (p. xiv).

Researchers have often regarded market dynamics as central in the evolution of this
trend. During most of the twentieth century, American news organizations enjoyed relatively low
levels of competition in their respective markets. This led to the coexistence of companies with
strong economic performance, journalists with professional values marked by a preference for
public affairs content, a style focused on issues and a disdain for their audiences (in Darnton’s
[1975, p. 176] apt formulation, journalists “really wrote for one another”), and a separation
between the editorial and commercial logics of the enterprise. Lack of market competition allowed this mode of production to survive despite the perception of a gap between the interests of journalists and consumers. For instance, Bogart (1989) examines data on newspaper readers’ ideal space allocation and interest in various topics and contrasts that with content analyses and editors’ importance ratings. He reports that 49% of the stories in a typical newspaper are public affairs stories whereas about 29% focus on non-public affairs content (see Table 6.1, p. 192), but that “readers want more of certain kinds of content than they presently get,” including primarily non-public affairs topics such as health, nutrition, environment, home, religion, and travel (p. 294). Thus, concluding that there is “a discrepancy between editors’ and readers’ ratings of the same subject.” (1989, p. 312). However, a growth in competition during recent decades, accelerated by the emergence of 24-hour cable television in the 1980s and the Web in the 1990s, has been eroding the market position of traditional players. In turn, this has led to a stronger impact of a commercial logic on editorial pursuits and more attention to the audience’s interests. Issue-based, public affairs news is more expensive to produce and seemingly less commercially appealing than its personality-based, nonpublic affairs counterpart. Thus, the trend towards the softening of news has been seen as an attempt by media firms to counter their market decline and narrow the perceived gap between what they produce and what their audience seems to want.¹

As Baum and Jamison (2006) have pointed out regarding political communication studies—although it applies more broadly to research on hard and soft news—“scholars have focused more on the supply of political information than on the nature of citizens’ demands” (p. 947, emphasis in the original). However, some studies have used surveys of consumers’ preferences, print circulation, broadcast ratings, and online-traffic data to ascertain whether softer news helps match the preferences of consumers or actually drives them away (Baum,
Two positions have emerged from these studies. One camp maintains that consumers want “softer” news, therefore the softening trend should help narrow the gap. For instance, commenting on his findings on website usage, Hamilton (2004) suggested that “the relative demand expressed for information on the Web shows patterns familiar from other media. Soft news and ‘news you can use’ often generate more interest than hard news” (p. 194). The opposing side contends that interest in “harder” content is paramount among consumers, thus a supply of softened news has created a new gap that undermines the market position of media firms. Thus, Patterson (2003) argued that “if we are to employ a standard based on the news that most consumers want and that most outlets provide, we would not want to have soft news in mind. We should be thinking in terms of hard news” (p. 140).

The attempt to determine whether the softening of news might have narrowed a gap between the news choices of journalists and consumers or created one where none existed has been further complicated by a lack of agreement among scholars about the nature of this trend. At least three views can be distinguished in the literature. First, to scholars such as Hamilton (2004), the trend entails a transformation in what news is reported: “I… use the term ‘soft news’ to refer to programs with low levels of public affairs information and ‘hard news’ to refer to shows with high levels of public affairs information” (p. 15). Second, to researchers such as Baum (2007), it is about how the news is covered instead of what its subject is, so even public affairs stories can be softened: “Soft news outlets emphasize different aspects of political issues than do their hard news counterparts. In war, for instance… where hard news focuses on geopolitics, soft news focuses on body bags” (p. 107). Finally, to analysts such as Gans (2003), the trend involves a change in the mix of what and how.
The change in the hard/soft news ratio has taken on several forms… [It] can be seen clearly by the expansion of the "back of the book" sections… about subjects as varied as science and gardening… Another version of this change has come in program formats… For example, the historic one-hour or half-hour news documentary series… which was often devoted to investigative reporting about a single subject has been replaced by the semiweekly… “news magazines” that supply a mix of hard and soft news. (p. 29)

The literature reviewed so far has advanced our understanding of the trend towards the softening of news and its consequences in terms of a possible gap between the choices of journalists and consumers. But it also has suffered from at least three limitations: (a) a relative mismatch between consumption data aggregated at the outlet or program level and consumer behavior that is increasingly disaggregated at the story level; (b) a lack of precision in the definition of the object of study; and, (c) the negative effects of confounding factors in ascertaining the news choices made by journalists and consumers.

First, in an increasingly competitive market where consumers have a growing array of unbundling techniques at their disposal—from remote controls to Really Simple Syndication (RSS) feeds—surveys of preferences or ratings and circulation data at the outlet or program level are too aggregate to assess the effectiveness of news supply on consumption choices often made at the story level.² Bogart (1989), who undertook one of the most sweeping attempts to “tackle the critical question of what content attracts [newspaper] readers” (p. 7), employed various methods but noted the difficulties of interpreting the resulting findings. Second, the lack of agreement about what the trend toward soft news consists of has made it difficult to have a fine-grained portrait of the phenomenon. Is it about the subject of the stories, how they are covered or both? Because different studies have defined the trend differently, this has also complicated a
comparison of findings across them and subsequent theory development efforts. Third, in addition to issues of newsworthiness that derive from professional judgment, the choices that journalists make have also been increasingly influenced by value signals conveyed by consumers in their demand patterns. Today’s journalists are more inclined to write for their audiences than Darnton and his colleagues did decades ago. By the same token, research on traditional (Graber 1988) and new media (Sundar, Knobloch-Westerwick, and Hastall 2007) has shown that consumers’ choices are partly influenced by value signals embedded in the supply of stories. For instance, all else being equal, an item displayed on a newspaper’s front page has a higher probability of being noticed and consumed than an item buried inside one of its sections.

Situated within this literature on the softening of news, the goal of this paper is to analyze the top news provided by journalists in four web sites on the hard-soft dimension, and to determine whether there is a gap—in terms of subject and format—between these stories and the stories most often chosen by consumers on these respective sites. Moreover, the research approach adopted here aims to help solve, at least in part, the limitations depicted above.

Methodology

Sample

Data were collected from four English language online news sites: chicagotribune.com (Chicago), seattlepi.newsource.com (Seattle), cnn.com (CNN), and news.yahoo.com (Yahoo). Chicago and Seattle are two of the most visited web sites of print newspapers in the United States, with almost 3 million and 3.3 million unique users for November 2006, respectively (Newspaper Association of America, 2006). These sites are the online counterparts of the Chicago Tribune and the Seattle Post-Intelligencer, with an average daily circulation of about 567,000 and 128,000, respectively, for the period ending in September 2006 (Audit Bureau of
CNN is the third most visited online news site, with a 2006 monthly average of almost 25 million monthly unique users, according to Nielsen measurements (Project for Excellence in Journalism, 2007). It is the online counterpart of the Cable News Network, which had a median prime time audience of 710,000 viewers during 2006 (Project for Excellence in Journalism, 2007). Yahoo is the most visited online news site, with a 2006 monthly average of over 28 million unique users, according to Nielsen measurements (Project for Excellence in Journalism, 2007). It is affiliated with Yahoo!, which runs the second largest web search engine with over 20% of market share for July 2007 (Nielsen/NetRatings, 2007).

These sites were also selected because they exhibit variance in terms of parent media company and geographic scope. Research has shown that the behavior of the parent media company can influence that of its online counterpart (Boczkowski, 2004; Chan-Olmstead & Ha, 2003; Paulussen, 2004; Singer, 2004), thus we chose cases that represent print—Chicago and Seattle, broadcast—CNN, and online—Yahoo, firms. Because scholars have shown that geographic matters variously affect media production (Baisnee & Marchetti, 2006; Hannerz, 2004; Kaniss, 1991; MacGregor, 1997; Picard, 2002), we selected two cases with a national-global reach—CNN and Yahoo—and two local-metropolitan cases—Chicago and Seattle.

We collected data on twenty-one randomly selected days—three for each day of the week—between April and June of 2007, for a total of three composite weeks. On each coding day, research assistants collected data at 10 a.m., 4 p.m., and 10 p.m. Central Time, yielding 63 distinct data collection shifts—21 days x 3 times per day—for each site. At each shift for each site, we identified the top ten stories selected by journalists and by consumers, respectively. We focused on these stories because, from the universe of possible stories, these are the ones that are deemed most relevant by journalists and also garner most attention from consumers. Thus they
represent a suitable approximation of each group’s key news preferences.

Online stories were defined as text-based packages that included a headline; a story might, but need not, have multi-media features—such as video or graphics—or links to related stories. In cases in which there was a main story accompanied by one or more sub-stories, only the main story was included. These criteria applied to all stories, regardless of whether they were breaking news, commentary, blogs and so on. The stories selected by journalists were the ten most prominently displayed stories on the home page; depending on the page's format, these were either listed in vertical order or were identified by counting from left to right and from the top down in a grid-like manner. Editors had a role in selecting the stories that appeared on their respective home pages of these four sites. CNN, Chicago, and Seattle are extensions of “traditional” news organizations and their online operations include a significant presence of editors. Yahoo also employs human editors (Glaser, 2005). The stories selected by consumers were identified by examining the ten most-clicked stories on each site—as indicated, variously, by sites’ “most viewed,” “most read,” or “most popular” lists; this information was publicly available on each site during our study. Across the 63 data collection shifts, each site yielded 1,260 stories—630 journalist-selected top stories and 630 consumer-selected top stories—for a total of 5,040 stories. The unit of analysis was the story. The following variables were coded: Intrasite news choice, Intrasite news choice overlap, Intersite convergent news choice, Intersite convergent news choice overlap, What, and How.

Analysis

We examined news choices by journalists and consumers at two levels of analysis: within site (intrasite) and across sites (intersite). Within each site, we looked at 630 stories selected by journalists and 630 chosen by consumers. Across sites, we focused on stories that covered the
same event and were selected by either journalists or consumers in more than one site in the same data collection shift. This analysis examined 539 such stories selected by journalists and 485 by consumers across the four sites. It enabled us to detect whether patterns of choice within one or more sites were also present across these sites.

Based on the issues identified in the Theoretical Considerations section, to help tease out the potential influence of stories that might have been chosen by journalists because they were generating substantive site traffic among consumers, or by consumers because they were saliently displayed by journalists, we conducted two separate analyses at each level of analysis. The first analysis examined the top choices of journalists and consumers in each site and also the convergent choices of journalists and consumers across sites. This analysis provides a general view of the choices made by journalists and consumers. The second analysis looks at the stories exclusively chosen by either journalists or consumers in each site but not by both. It also examines the convergent choices of journalists that were not shared by consumers and vice-versa. This analysis affords a distilled picture of these news choices by focusing on the subset of stories that were selected by one group but not by the other. Regarding the intrasite data, this second analysis examined 26% of the stories on CNN, 46% on Chicago, 60% on Seattle, and 67% on Yahoo. Concerning the intersite data, the second analysis looked at 58% of the convergent choices on the journalists’ side that were not selected by consumers, and 53% of the convergent choices on the consumers’ side that were not chosen by journalists.

Results

Patterns of Choices within the Sites

We first examine choices regarding what subjects the stories are about and then move to analyze issues concerning how they are told.
When all the stories are analyzed, in three of the four sites journalists select mostly nonpublic affairs news: 54% on CNN and 62% on Chicago and Seattle. Yahoo is the exception, but the proportion of this news is still high at 43% (see Table 1). When only the stories with no overlap are considered, public affairs stories are the majority on CNN, Chicago, and Yahoo, but the share of nonpublic affairs news remains substantial on these three sites—36%, 44%, and 48%, respectively—and is still dominant on Seattle at 59%.

A comparison of the choices made by journalists and consumers shows that when all stories are included in the analysis, on each one of the sites, journalists chose more public affairs stories than consumers (see Table 1). The difference between the choices of journalists and consumers is 13 percentage points on Seattle and Yahoo, 14 percentage points on CNN, and 17 percentage points on Chicago. Chi squares tests of homogeneity show that the distribution of public affairs and nonpublic affairs stories is not the same for journalists and consumers ($\chi^2 = 21.349$ to $46.298$, all significant at $p = .000$).

When we look at the stories chosen exclusively by either journalists or consumers—excluding stories with overlap within the respective sites—differences between journalists’ and consumers’ share of public affairs stories increased on all sites: from 14 to 51 percentage points on CNN, from 17 to 38 percentage points on Chicago, from 13 to 23 percentage points on Seattle, and from 13 to 19 percentage points on Yahoo (see Table 1). This widening of the gap between the choices of journalists and consumers is accompanied by increased $\chi^2$ values and remains significant ($p = .000$) on each one of the sites. On CNN, Chicago, and Seattle, this widening arises from increases in both journalists’ share of public affairs stories and consumers’ share of nonpublic affairs news. In Yahoo the proportion of public affairs stories decreases for both journalists and consumers, with a more pronounced drop for the latter. Furthermore, looking
across the sites, as the number of stories chosen exclusively by journalists and consumers decreased, the gap between them increased. For example, only 165 stories were chosen exclusively by journalists on CNN compared with 424 on Yahoo, and the growth in their respective gaps with consumer choices was 37 and 6 percentage points.

In terms of how the stories are told, an analysis of the journalists’ choices, either including all stories or only those without intrasite overlap, shows that straight news is a strong dominant option on all the sites. It ranges from 64% to 83% in the first approach and from 63% to 75% in the second one (see Table 2). When all the stories are analyzed in comparison with the choices of consumers, results show Chicago and Seattle journalists chose significantly more straight news than consumers: 77% versus 64% on Chicago and 64% versus 58% on Seattle (p = .000 and p = .034, respectively; see Table 2). There are no significant differences on CNN and Yahoo. When only the stories without overlap are analyzed, the differences between journalists and consumers in the share of straight news on Chicago and Seattle rose from 13 to 30 percentage points and from 6 to 11 percentage points, respectively. The decrease in straight news among consumers was tied to an increase in commentary on Chicago and Seattle—from 15% to 29% and from 11% to 16%, respectively.10

Patterns of Choices across the Sites

Public affairs stories occupied the majority position of convergent news choices among journalists. But nonpublic affairs news still represented an important share of the top news stories selected by journalists, both when all stories were analyzed and when those with overlap were excluded—44% and 37%, respectively (see Table 3). The picture is different regarding how the news gets told: straight news overshadowed the other options in stories with and without overlap with an 83% and 82% share, respectively.
Regarding the subject of the stories, an analysis of all the stories shows that there is a significant 21 percentage point difference between the choices made by journalists and consumers \((p= .000)\). The choices of the former converge more on public affairs news and those of the latter more on nonpublic affairs stories (see Table 3). After removing the stories with intersite overlap, the gap between the convergent choices of journalists and consumers increases from 21 to 28 percentage points. This gain results from changes on the journalists’ side because overlapped stories were concerned mostly with nonpublic affairs subjects. Thus, removing them from the analysis altered the thematic distribution of journalists’ convergent choices but not that of consumers. Concerning how news are reported, there were no significant differences between stories that converged on the journalists’ side and those that converged on the consumers’ side, either before or after removing stories chosen by both.

Discussion

This study has aimed to ascertain the character of the top story choices made by journalists of online news sites on the hard-soft dimension and to determine whether there is a gap between these choices and the top choices made by the consumers of these sites. The analysis has shown that (a) journalists’ choices on these sites are substantively “soft” in terms of what the stories are about but not in terms of how they are told; and, (b) there is a gap between these choices and consumers’ choices that is uniform and relatively strong in terms of the subject of the stories but localized and comparatively weaker in terms of how these stories are told.

There is no established threshold after which the proportion of a particular type of content or format allows the analyst to characterize a news supply as “soft,” but we relied on two indicators to suggest that most top journalistic choices on these sites were marked by soft subject matters and hard storytelling techniques. First, nonpublic affairs stories made up the majority on
three of the four sites in the intrasite analysis of all stories, and reached a 55% combined average among all sites. After removal of the stories with overlap, these stories still constituted the majority on one site and a sizeable minority on the remaining three, and reached a 47% combined average among all sites. In addition, the share of nonpublic affairs stories in the convergent choices of journalists was a minority but still high considering that the focus was on the top news of the day. Second, the preeminence of content (i.e., what) over format (i.e., how) while measuring the extent of a “soft” news supply also becomes evident if we consider that in each case—both intrasite and intersite and before and after removal of overlapped stories—the highest share of public affairs subjects was lower than the lowest share of straight news formats.

An examination of the contrast between journalists’ and consumers’ choices shows a sizeable and uniform gap in terms of what stories are selected—which is consistent with studies that have shown reduced interest in public affairs among online news consumers (Althaus & Tewksbury, 2002; Hamilton, 2004; Tewksbury, 2003; Schoenbach, de Waal & Lauf, 2005; Tewksbury & Althaus, 2000)—and a more variable situation in the case of how stories are told.

In the analysis of stories’ subject matter, three trends that illuminated the choice gap emerged after the removal of overlapped stories. First, in the intrasite analysis, on three of the four sites, journalists’ choices moved in the direction of public affairs stories, and, on all sites, consumers’ choices did so toward nonpublic affairs ones. Second and in relation to intrasite data, the smaller a post-overlapped, subsample became, the larger the gap became. This suggests that perhaps when journalists have fewer stories to choose regardless of consumers’ preferences they lean strongly towards public affairs news. Third, in the intersite analysis, the post-overlapped subsample of journalists’ convergent choices moved in the direction of more public affairs stories, whereas the consumers’ choices remained unchanged. This suggests that journalists’
penchant for public affairs news might not cross over to consumers and that consumers’ preference of nonpublic affairs news might be less dependent on journalists’ choices.

Regarding how stories are told, there was a significant gap on two sites and none in the intersite comparison. This signals greater variability in this dimension than in the case of the subject matter of stories. Differences between the what and how variables emerge not only when contrasting the uniformity of the former with the variability of the latter, but also when assessing their relative strengths. In all cases in which both what and how gaps were significant on a single site, the size of the gap in the first case was larger than in the second. Additional analyses suggest that only a fraction of the stories placed at the top of the homepage by journalists stay on a given site throughout the day (i.e., appear in all three data collection times). The small amount of stories that stay on journalists’ top ten list longer are slightly more likely to be non-public affairs stories, but there is no evidence that these stories are more popular with consumers than public affairs news. Moreover, analyses of the interactions between the what and how variables revealed no significant evidence that consumers’ choices mixed public affairs subject matters with nonstraight news formats.

Concerning issues of geographic scope, the analysis yielded no major variance in terms of the stories’ subject matter. The two global-national sites were at the two extremes of the gap continuum. However, there were differences in how the stories were told. The two local-metropolitan sites exhibited significant gaps in the how variable, partly driven by a lower prevalence of straight news chosen by consumers. Concerning issues of the type of media company, the study showed that Yahoo, a news aggregator, featured the highest share of public affairs news in the sample of all stories, the smallest gap in the what variable, and a strong hold on straight news. The highest share of public affairs news on Yahoo than on the other sites could
result, at least in part, from the higher proportion of wire service copy—often more focused on public than on nonpublic affairs topics.

The analytical approach adopted contributed greatly to the inquiry. First, using the story as the unit of analysis permitted a fine-grained depiction of the character of journalists’ and consumers’ choices. Second, discriminating between the subject matter and form of news has allowed us to address the multiple definitions of what soft news means and help settle empirical debates regarding its consequences for the gap between journalists and consumers. At least for the cases studied, a soft news supply and the gap with consumption patterns are strongly marked by the subjects covered in the news. Issues of storytelling format are either absent or take a back seat. Third, that patterns existing when all stories were included in the analysis became magnified after removing stories with overlap underscores the heuristic power of this strategy to disentangle the influences of journalists’ and consumer’s choices on each other.

Because this study examines online news, it is important to consider the implications of two medium-specific factors for generalizability issues. First, although some initial studies have indicated a higher prevalence of public affairs stories in the consumption of traditional over online news, the more recent research has shown that topical preferences are stable across media (D’Haenens, Jankowski & Heuvelman, 2004; Dutta-Bergman, 2004; Nguyen & Western, 2006; Pew, 2007; Tewksbury, 2003, 2005). Second, there has been disagreement about whether the information about popularity levels of news stories on online sites affects consumption patterns and, if so, how important this effect is (Knobloch-Westerclick, Sharma, Hansen & Alter, 2005; Sundar & Nass, 2001; Thorson, 2008). If such an effect existed and was strong, it could contribute to making the choice gap greater in online sites than in traditional media.11

It is worth considering four limitations of the research design. First, the researchers did
not have access to their respective measurement processes utilized for calculating the most popular stories in the sites studied. If these processes differed markedly, it might have affected the data in a way that could have altered the results. But another interpretation is also feasible. If there was major variance in these processes, that intra- and intersite patterns moved in similar directions could further strengthen the robustness of the findings. Second, headlines and leads on a homepage convey a better sense of the subject matter of a story than of its format. Although this discrepancy might have also affected the results, findings from a parallel ethnographic study of online news consumption conducted by the first author have also shown a preference for a straight news style (Author A, in press). Nevertheless, future studies should distinguish stories with various levels of information on the homepage about their storytelling techniques. Third, although the sample had variance in two important dimensions, it is comprised of only four sites which had a reputation for serious journalism and were based in the same country. Future research should examine sites of varying reputations and in different nations. Fourth, the study privileged order of appearance over other dimensions of story importance, such as use of visuals or headline size. Future scholarship on online news choices should try to take into account these dimensions.

This study raises implications for understanding at least three sets of developments at the intersection the choice gap and changes in the media industry and its broader social functions: namely, the potential consequences of an increasing supply of soft content for the news agenda, the growth of consumers’ role as producers of widely circulated news and information, and the future of the media’s watchdog function in liberal democracies.

As for the evolution of large companies, such as those examined here, the growth in market competition, the greater attention paid to consumers’ interests, and greater predisposition
to cater to them than when the industry was dominated by a few organizations with monopoly or oligopoly power, provide a reasonable basis to believe that the supply of news might move towards decreasing the share of public affairs stories in an attempt to close the choice gap. These phenomena transformations might, in turn, call for conceptual changes. For instance, agenda setting theory was premised on the idea that the media have a significant degree of autonomy in selecting what “are the most important topics of the moment” (McCombs, 2004, p. 20). However, if consumer behavior became a central factor shaping journalistic story selection, this premise would be challenged. This is consistent with Chaffee and Metzger’s suggestion that “the key problem for agenda-setting theory will change from what issues the media tell people to think about to what issues people tell the media they want to think about” (2001, p. 375).

Scholars have recently theorized on the beneficial possibilities of consumers’ growing involvement with the production of news (Benkler, 2006; Gillmor, 2004; Jenkins, 2006; Sunstein, 2006). Perhaps nowhere has this idea been more forcefully argued than in Benkler’s vision of a “networked public sphere,” which “allows all citizens to… no longer… be consumers and passive spectators. They can become creators and primary subjects. It is in this sense that the Internet democratizes” (2006, p. 272). However, the pervasiveness and strength of consumers’ preferences for nonpublic affairs stories revealed in this study belie notions about a well-informed citizenry that often accompany this kind of argument. They also raise concerns about what might happen if consumers’ choices shaped a large share of the media’s content.

Finally, one way in which media organizations contribute to the quality of democratic life is by playing a watchdog role over other powerful actors in society (Baker, 2002; Glasser & Ettema, 1998; Lazarsfeld & Merton, 1948; Waisbord, 2000). They do so mostly through public affairs stories, and their ability to be effective is dependent on spending considerable resources to
gather information about wrongdoing. The prevalence of soft news in terms of subject matter, the possibility that this is a trend that might intensify in the direction of closing the gap with dominant consumption patterns, and the potential of strong consumer involvement in the production of news raise questions about the ability of media companies to play a strong watchdog role in the future. Taken together, these findings and trends point to potentially significant changes in the capacity of the media to positively affect the balance of power in society.
References


The Choice Gap

Press.


Patterson, T. (2000). *Doing well and doing good: How soft news and critical journalism are shrinking the news audience and weakening democracy -- and what news outlets can do about it*. Unpublished manuscript, Harvard University, Cambridge, MA.


http://jcmc.indiana.edu/vol9/issue4/paulussen.html


Tewksbury, D., & Althaus, S. (2000). Differences in knowledge acquisition among readers of the paper and online version of a national newspaper. *Journalism and Mass Communication*


### Table 1

**Intrasite Comparison of “What” of News Choices**

<table>
<thead>
<tr>
<th></th>
<th>Intrasite news choice</th>
<th>Intrasite news choice overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Journalists (%)</td>
<td>Consumers (%)</td>
</tr>
<tr>
<td><strong>CNN</strong></td>
<td>n=630</td>
<td>n=629</td>
</tr>
<tr>
<td>Public affairs</td>
<td>46</td>
<td>32</td>
</tr>
<tr>
<td>Nonpublic affairs</td>
<td>54</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>$\chi^2=23.866, df=1, p=.000$</td>
<td>$\chi^2=88.180, df=1, p=.000$</td>
</tr>
<tr>
<td><strong>Chicago</strong></td>
<td>n=630</td>
<td>n=630</td>
</tr>
<tr>
<td>Public affairs</td>
<td>38</td>
<td>21</td>
</tr>
<tr>
<td>Nonpublic affairs</td>
<td>62</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>$\chi^2=46.298, df=1, p=.000$</td>
<td>$\chi^2=90.019, df=1, p=.000$</td>
</tr>
<tr>
<td><strong>Seattle</strong></td>
<td>n=630</td>
<td>n=630</td>
</tr>
<tr>
<td>Public affairs</td>
<td>38</td>
<td>25</td>
</tr>
<tr>
<td>Nonpublic affairs</td>
<td>62</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>$\chi^2=27.237, df=1, p=.000$</td>
<td>$\chi^2=47.984, df=1, p=.000$</td>
</tr>
<tr>
<td><strong>Yahoo</strong></td>
<td>n=630</td>
<td>n=630</td>
</tr>
<tr>
<td>Public affairs</td>
<td>57</td>
<td>44</td>
</tr>
<tr>
<td>Nonpublic affairs</td>
<td>43</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>$\chi^2=21.349, df=1, p=.000$</td>
<td>$\chi^2=32.410, df=1, p=.000$</td>
</tr>
</tbody>
</table>
Table 2

*Intrasite Comparison of “How” of News Choices*

<table>
<thead>
<tr>
<th></th>
<th>Intrasite news choice</th>
<th>Intrasite news choice overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Journalists (%)</td>
<td>Consumers (%)</td>
</tr>
<tr>
<td><strong>CNN</strong></td>
<td>n=630</td>
<td>n=630</td>
</tr>
<tr>
<td>Straight news</td>
<td>83</td>
<td>80</td>
</tr>
<tr>
<td>Feature style</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>Commentary</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Alternative</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>(\chi^2=1.536, df=1, \text{n.s.})</td>
<td>(\chi^2=3.673, df=1, \text{n.s.})</td>
<td></td>
</tr>
<tr>
<td><strong>Chicago</strong></td>
<td>n=630</td>
<td>n=630</td>
</tr>
<tr>
<td>Straight news</td>
<td>77</td>
<td>64</td>
</tr>
<tr>
<td>Feature style</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Commentary</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Alternative</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>(\chi^2=28.116, df=1, p=.000)</td>
<td>(\chi^2=52.260, df=1, p=.000)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 (cont’d)

*Intrasite Comparison of “How” of News Choices*

<table>
<thead>
<tr>
<th>Intrasite news Choice overlap</th>
<th>Intrasite news Choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Journalists (%)</td>
</tr>
<tr>
<td></td>
<td>n=630</td>
</tr>
<tr>
<td></td>
<td>Journalists (%)</td>
</tr>
<tr>
<td></td>
<td>n=381</td>
</tr>
</tbody>
</table>

**Seattle**

<table>
<thead>
<tr>
<th></th>
<th>n=630</th>
<th>n=630</th>
<th>n=381</th>
<th>n=381</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straight news</td>
<td>64</td>
<td>58</td>
<td>63</td>
<td>52</td>
</tr>
<tr>
<td>Feature style</td>
<td>26</td>
<td>29</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Commentary</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>Alternative</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 8.652, df=3, p=.034 \]

\[ \chi^2 = 12.727, df=3, p=.005 \]

**Yahoo**

<table>
<thead>
<tr>
<th></th>
<th>n=630</th>
<th>n=630</th>
<th>n=424</th>
<th>n=424</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straight news</td>
<td>75</td>
<td>78</td>
<td>68</td>
<td>73</td>
</tr>
<tr>
<td>Feature style</td>
<td>23</td>
<td>18</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>Commentary</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Alternative</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 4.548, df=3, \text{n.s.} \]

\[ \chi^2 = 5.449, df=3, \text{n.s.} \]

---

\( a \) To apply the \( \chi^2 \) test and eliminate expected cell frequencies <5, categories were collapsed to Straight News and Nonstraight News.
Table 3

*Intersite Comparison of “What” and “How” of News Choices*

<table>
<thead>
<tr>
<th></th>
<th>Journalists (%)</th>
<th>Consumers (%)</th>
<th>Journalists (%)</th>
<th>Consumers (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=539</td>
<td>n=485</td>
<td>n=311</td>
<td>n=255</td>
</tr>
<tr>
<td><strong>What</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public affairs</td>
<td>56</td>
<td>35</td>
<td>63</td>
<td>35</td>
</tr>
<tr>
<td>Nonpublic affairs</td>
<td>44</td>
<td>65</td>
<td>37</td>
<td>65</td>
</tr>
<tr>
<td><strong>χ²=42.622, df=1, p=.000</strong></td>
<td></td>
<td></td>
<td><strong>χ²=43.312, df=1, p=.000</strong></td>
<td></td>
</tr>
<tr>
<td><strong>How</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Straight news</td>
<td>83</td>
<td>80</td>
<td>82</td>
<td>78</td>
</tr>
<tr>
<td>Feature style</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Commentary</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Alternative</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>χ²=2.097, df=1, n.s.</strong></td>
<td></td>
<td></td>
<td><strong>χ²=1.350, df=1, n.s.</strong></td>
<td></td>
</tr>
</tbody>
</table>

*To apply the χ² test and eliminate expected cell frequencies <5, categories were collapsed to Straight News and Nonstraight News.*
Endnotes

1 This is not the only strategy pursued by news organizations to deal with a more competitive market and a perception of decrease market share. Other strategies include various forms of product differentiation, as it has been argued in scholarship on media fragmentation (Gandy, 2001; Jamieson & Cappella, 2008; Klinenberg, 2007; Neuman, 2001; Turow, 2005, 2006). However, production differentiation strategies need not be mutually exclusive with an overall softening of the news content—for instance, organizations have often softened their products in multiple, coexistent ways.

2 We partly build on Tewksbury’s (2003) strategy of using Nielsen//NetRatings page view consumption data, but we also extend it. Although his strategy moved towards analyzing stories, it did not examine them directly, but inferred their content from the information provided in their respective URLs—which also limited his sample selection.

3 This predisposition seems to have intensified with the availability of data on sites’ traffic in their respective newsrooms—including, for instance, at Yahoo News (Rosmarin 2007).

4 The Newspaper Association of America data combine the Seattle Times and Seattle Post-Intelligencer figures because the two companies have a joint operating agreement. Their web operations, however, are independent.

5 The print edition of the Seattle Post-Intelligencer folded in spring 2009.

6 Research on traditional media—for instance, Graber (1988)—has shown that visual cues such as headline size and location can influence consumption patterns. Because variance in homepage design of the sites studied for this project made comparison of such cues difficult, a decision was made to prioritize order of appearance in the operationalization of the top ten journalists stories.
During the period of study, in the four sites we studied, these stories were the most frequently clicked stories. These lists were updated every hour on Chicago, Seattle and Yahoo, and every twenty minutes on CNN. Furthermore, whereas some news sites made available information on other measures of consumer choices, such as “most e-mailed” and “most recommended,” we focused on the most clicked stories because they seemed the most straightforward expression of consumers’ preferences.

Four trained research assistants coded the 5,040 stories. Intercoder agreement was assessed a subset of 5% of the data. Regular intercoder agreement levels averaged 94% and ranged from 88% to 100%. Cohen’s Kappa intercoder agreement levels averaged 88% and ranged from 77% to 100%. Doubts that arose were subsequently consulted with the authors and resolved consensually.

The variables and categories are defined as follows:

1. **Intrasite news choice.** There were two categories:

   (a) **Journalists**, the top ten stories that appeared on the home page of each site in each shift.

   (b) **Consumers**, the top ten stories in the “most viewed,” “most read,” or “most popular” lists of each site in each shift.

2. **Intrasite news choice overlap.** There were three categories:

   (a) **Overlap**, stories that appear in both categories (a) and (b) of variable 1 in the same site during the same data collection shift.

   (b) **No overlap-journalists**, stories that appear in category (a) variable 1 but not in category (a) of variable 2, in the same site during the same shift.

   (c) **No overlap-consumers**, stories that appear in category (b) variable 1 but not in
category (a) of variable 2, in the same site during the same data collection shift.

3. Intersite convergent news choice. There were three categories:
   
   (a) Journalists, stories about the same event that are included in category (a) of variable 1 in more than one site in a given data collection shift.

   (b) Consumers, stories about the same event that are included in category (b) of variable 1 in more than one site in a given data collection shift.

   (c) No convergent choice, stories that do not fall under categories (a) or (b).

4. Intersite convergent news choice overlap. Only stories that displayed convergence among either journalists or consumers—i.e., only stories classified as (a) or (b) in variable 3—were coded for this variable. There were three categories:
   
   (a) Overlap, stories included in either category (a) or (b) of variable 3 that are about an event that is also reported in a story appearing in the opposite category of variable 1 in at least one of the convergent choice sites in the same shift.

   (b) No overlap-journalists, stories included in category (a) variable 3 that are about an event that is not reported in a story appearing in category (b) of variable 1 in any of the convergent choice sites in the same data collection shift.

   (c) No overlap-consumers, stories included in category (b) variable 3 that are about an event that is not reported in a story appearing in category (a) of variable 1 in any of the convergent choice sites in the same data collection shift.

5. What. The main topic addressed in the story. There are two categories:
   
   (a) Public affairs stories, dealing with news about politics, government, economics, business, international affairs, and “war on terror.” This includes stories about the activity of government, elected officials, political candidates; the economy and business
developments; and, events, happening in other countries, about the state or international organizations.

(b) Nonpublic affairs stories, addressing subjects such issues as sports, crime, entertainment, technology, and weather. This includes stories about sports teams and events; criminal activity at the local, national or international level; visual and performing arts, and literature; medical, scientific and technological matters; damage caused by disasters and large-scale accidents; and routine and nonroutine weather information.

6. How. The main story-telling technique. There are four categories:

(a) Straight news, which refers to a writing style that emphasizes the important facts and is often told in a nonpersonal, detached way. Often includes information about who, what, where and when up at the top—sometimes after an anecdote.

(b) Feature-style, alluding to stories that employ narrative or other literary devices and can include a more personal tone. It often has a beginning-middle-end narrative structure.

(c) Commentary, which includes all stories with a pronounced point of view about a topic or a product, such as columns, “op-eds,” and reviews.

(d) Alternative formats, consisting of all other formats not mentioned above.

10 Additional analyses show that most of these commentary stories are about movies on Chicago and about sports on Seattle.

11 In addition, it should be noted that this research does not ascertain whether consumers utilize online news as a primary or secondary—that is, supplementary to the consumption of traditional media—source of information. Future research would benefit from an examination of this matter by complementing the content analysis data with methods such as surveys or interviews.
Because this gap is only one factor that shapes the behavior of news organizations, this hypothetical scenario is contingent on an array of additional possible factors, including shifts in technology, ownership structure, regulatory environment, and public culture.

One alternative is that consumers of these sites studied chose a large proportion of nonpublic affairs stories partly because these sites already furnished a relatively large share of public affairs stories. Thus their behavior might have been different if their choices had been made in the absence of a substantive dose of news about public affairs. Future research should elucidate the character of consumers’ news choices in different circumstances of consumer-authored media.