Where is the revolution?

Health news on the Internet: Online user preferences and their contrasts with prevalence of private-sector originating sources

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Maureen Schriner

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Brian Southwell, PhD

January, 2011

Acknowledgements

Thank you to Brian Southwell, PhD, a brilliant scholar and teacher, for encouraging me to pursue my doctoral degree and supporting me in so many ways throughout these past six years toward graduation.

Thank you to Catherine Squires, PhD, for choosing me to be part of your research, for being a true colleague in sharing your ideas and opinions, and encouraging me to stay on track and complete my dissertation.

Thank you to John Eighmey, PhD, for your sense of humor and friendliness, for your innate sense of how to ask the right questions, and for providing advice and insights about developing a career in academia.

Thank you to Mark Snyder, PhD, for sharing your wealth of knowledge about the psychology of society. It taught me to see relationships with individuals and organizations in new ways, and to better understand others and myself.

Thanks to the consumer health organization that collaborated with me on this research. You once again demonstrated why your organization is a world leader in health, because you believe in, and invest in, research.

Finally, thanks to fellow researcher Rachel Armstrong for your excellent work with content coding.

Dedications

To my family – all of you – with love and devotion

To my friends – Hui-Hui Wang and Sue Klappa

Abstract

The public's access to health information on the Internet has been acclaimed as a revolution in health, empowering individuals to gain health knowledge by actively seeking information online. But recent scholarship has identified health scanning, a more passive exposure to health information on the Internet, as more common among online users than active health seeking. Health scanning suggests a different model of individuals gaining health knowledge through expert sources that provide health news. The present dissertation proposes a model of interactive agenda-building to explore the interaction of online health scanners, news media and originating sources of news. The theoretical basis of the model is agenda setting/agenda building and organizational social identity theory. Two studies test the interactive agenda-building model. Study 1 involved a survey of over 1,300 users of a consumer health Web site. Key findings were that online users prefer public, private and academic sources, rather than private-sector sources or personal, non-expert sources. Study 2 encompassed content analysis of articles from four mainstream news media Web sites that were listed on user-generated most-viewed and most-emailed lists. Originating sources within each article were categorized as personal or institutional, and institutional sources were further categorized as public, private, nonprofit or academic. Findings show online users consistently chose health news in which institutional sources were more prevalent than individual, and private-sector sources were more prevalent than public, nonprofit or academic sources. The contrast of online user preferences from Study 1 and online health news prevalence of private sources suggests the health consumer revolution has not yet evolved.

Table of Contents

List of Tables

List of Figures

Where is the revolution?

Health news on the Internet: Online user preferences and their contrasts with prevalence of private-sector originating sources

Chapter 1

INTRODUCTION

Health information has been transformed by the Internet, from concentration within circles of health experts to widespread availability to the general public (Anderson, 2004; Cullen, 2006; Rains, 2008). Over the past decade, as more adults began using the Internet, extensive research has accumulated about online users actively seeking health information (Brodie *et al*, 2000; Cline & Hayes, 2001; Cotten & Gupta, 2004; Finney Rutten, Squiers & Hesse, 2006; Goldner, 2006; Rains, 2007; Toms & Latter, 2007). About three in four U.S. adults are now online (Pew, 2009) and, among those online, over 80% report seeking health information on the Internet (Harris Interactive, 2010). Individuals' use of online health information is viewed as a driving force in the consumer health movement to shift health decision-making from medical experts to patients and their families (Hesse, 2009). At the same time, whether this shift, and the consumer health movement, is a wholesale revolution is an open question.

Health scanning: Conventional & new

Recent research suggests individuals may acquire more health information through what is called health scanning than health seeking. In contrast to what the literature refers to as health seeking, which is purposeful, active online searches for information on a specific topic, scholarship has developed the term health scanning as a

relatively passive activity in which individuals are incidentally exposed to health information (Niederdeppe *et al*, 2007; Shim, Kelly & Hornik, 2006; Tian, 2009). The experience of acquiring health information through scanning, or browsing, occurs during day-to-day routines of media use (Case, 2007) and typically involves use of news media (Dutta-Bergman, 2004; Johnson et al, 2001). Browsing through daily news, whether it is online or off, puts the individual in the traditional role of news consumer. It shifts control back to news media, and suggests a return to traditional agenda-setting effects. It also reinforces the influence of agenda-building and the power of organizations providing information subsidies (Gandy, 1982).

For all the bells and whistles that accompany new media technology, the structure of new media and its effects can mimic "old media" in many ways (Yzer & Southwell, 2008). Health scanning raises the question of what has actually changed, and what has remained the same, in the relationship of the public with journalists, and the public with originating sources of news. The purpose of this dissertation is to examine the interactions of online users, news media and originating sources in building an agenda of health news on the Internet. The contributions of the three entities are considered in terms of defining characteristics and preferences of online users and health journalists, and prevalence of originating sources in online health news.

There are two primary paths followed to explore who, and how, agenda-building occurs in online health news. The first path is to assess the online environment: How could agenda-setting, and more so agenda building, occur in online health news? Who creates the agenda? The second path is to assess sources of online health news. How

does interactivity of online users, news media and originating sources ultimately develop an agenda of health news?

The introduction provides a brief overview of the health news environment, introduces a new model of interactive agenda building and provides a rationale for examining agenda-building of health news on the Internet. The literature review then constructs a theoretical foundation for online agenda-building in health news, by examining scholarship regarding the interactions of online users, news media and originating sources of health news. Two studies then assess the online health news environment to determine the interactions involved in health scanning. Study 1 examines online user characteristics and preferences for online health news through survey research. Study 2 provides a contrast with observational data and content analysis of actual online user choices of health news, and the prevalence of originating sources in online health news. The discussion section summarizes the findings and implications for health news, and the individuals who scan health news.

The health news environment: Traditional and new

Researchers have long recognized the influence of news media in exposing individuals to new health information, and influencing the public and policymakers in their health beliefs and health behaviors (Davidson & Wallack, 2004; Meissner, Potosky & Convissor, 1992; Stryker, Moriarty & Jensen, 2008; Wanta & Elliot, 1995; Yanovitzky & Bennett, 1999; Yanovitsky & Stryker, 2001).

Even in the era of fragmented audiences and selective exposure (Bennet & Iyengar, 2008), there are key indicators that news media remain influential providers of health information to the online public (Mitchelstein & Boczkowski, 2010). Audience

size online is one indicator, with traditional news outlets attracting massive audiences and substantial growth in online Web traffic in recent years (NAA, 2009). Another indicator is significant inter-media agenda-setting online, with increasing influence of elite media in setting priorities for issues (Carlson, 2007; Meraz, 2009; Michelstein & Boczkowski, 2010). Additionally, traditional journalism appears to have ongoing normative influence on discourse of public issues, with Internet communication showing no significant difference from print news (Gerhards & Schafer, 2010). In sum, these indicators suggest the potential for agenda-setting effects on health scanners.

But Chaffee and Metzger (2001) raise an important question in how online interactivity may alter the agenda-setting equation. Introducing the element of interactivity to the agenda-setting process places an emphasis on relationships, in examining who is interacting to build an agenda. Rather than looking at agenda setting and its focus on issue salience (McCombs, 2004), the present dissertation takes a public relations perspective of examining agenda building and its focus on the salience of sources of online health news.

Health news sources: Traditional and new

Agenda building and the influence of sources is rooted in the history of public relations, in which relationships have been strategically cultivated in order to advance an organization or institution's agenda (Lamme & Russell, 2010). Recent literature has developed strong ties between agenda building and information subsidies, examining the public relations methods organizations use to advance their agenda through news media (Carroll & McCombs, 2003; Kiousis, Popescu & Mitrook, 2007; Kiousis & Wu, 2008). These studies develop agenda building as an extension, or secondary level, of agenda

setting (Ghanem, 1997) in measuring source salience and attributes, rather than issue salience. Alternatively, the field of political science has developed agenda building as an organic process in which the policy agenda, or public agenda, is constructed through the involvement and influence of various participants (Cobb & Elder, 1971; Gamson & Modigliani, 1989; Lang & Lang, 1980).

Agenda-building as interactive

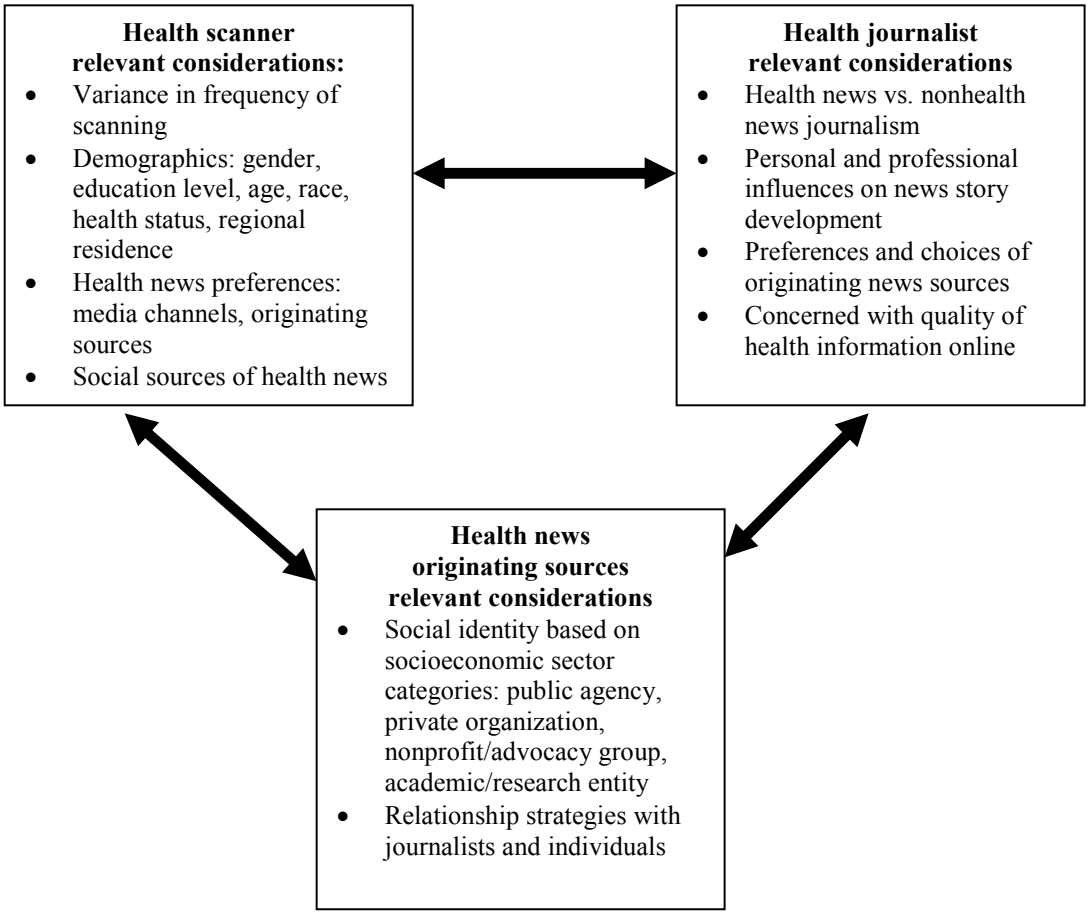
The model of agenda building proposed in this dissertation integrates agenda-building perspectives, and takes a more comprehensive view of sources. Under this model, interactive agenda builders include online users, who make choices of media sources and health news content in a competitive media environment; journalists who make choices as traditional gatekeepers and as information providers of health news in a dynamic, competitive media environment; and originating sources, organizations and individuals who interact with journalists and provide information for health news through strategic information subsidies (Gandy, 1982) and through media relations strategies (Zoch & Molleda, 2008). The interactive agenda-building model also takes a broader view of the concept of public agenda as one constructed based on the sources' strategic (or non-strategic) agendas rather than issue-specific agendas. It frames the transformation of health information not based on the technology of the Internet, but the agents who interact online to generate health news.

In proposing the interactive agenda-building model, this dissertation will examine two aspects of the model, the agenda builders' characteristics and their selection of health news, as users, producers and originating sources. Each group's

attributes and choices are tested to compare differences and identify what factors predict agenda building in online health news.

Figure 1 illustrates the interactive agenda-building model and the attributes to be examined in the literature review.

Figure 1. Interactive health news agenda-building: characteristics of agenda builders



The characteristics and choices of the agenda-builders are constructed through two studies. The first study is survey research of 1,600 users of a national, high-traffic

consumer health Web site. Questions included respondents' demographics, online health information acquisition and media habits, and their preferences for originating sources of health news online and off line. (See Appendix A for survey questions). Analysis focuses on differences among high, moderate and low health scanners. A second study involves content analysis of articles posted to most-viewed and most-emailed lists on Web sites for four mainstream news media, all of which have high volumes of Internet traffic. The content is categorized as health news or nonhealth news and coded for originating sources cited within the news. (See Appendix B for coding summary). Analyses assess the health news environment in various ways, and then assesses the originating sources within health news, in terms of the sources' influence in agenda building of online health news.

Rationale

Examining online agenda building, and proposing an interactive agenda-building mode, makes important advances in the literature of health communication, mass communication, computer-mediated communication and public relations in three areas.

The first advance in research is to establish agenda building, and the interaction of online users and sources of news, as a key measure of salience of online news. The focus on agenda building recognizes the weaknesses that have been identified in agenda setting (Pan & Kosicki, 1993; Scheufele, 2000) for being too broad and abstract, and biased toward issues at a superficial level. The shift from examining issues to examining relationships adds a new perspective in the development of agenda-building research that has emerged as a strong element in public relations literature.

The second advance this dissertation provides to the literature is analysis of a unique dataset, content analysis of articles from most-viewed and most-emailed news from mainstream news media Web sites. The four news media selected for data analysis are CNN, USA Today, The Washington Post, and the Los Angeles Times. These four media organizations are among the top 20 Web sites for online news (Table 1). Online use of news and information is heavily concentrated. Among 4,600 news and information Web sites, over 80% of Internet traffic was concentrated on the top 7% of Web sites (PEJ, 2010). The majority of the top Web sites are "legacy" media, traditional news organizations initially established as print or broadcast news media (Nielsen, 2010; PEJ, 2010). The concentration represents mass audiences, which addresses another point raised by Chaffee and Metzger (2001) that media will continue to unite millions of individuals. Consider the four news Web sites chosen for data analysis in this dissertation attract between 8.5 million to nearly 21 million unique visitors to their Web sites during an average month (See Table 1).

Table 1. Nielsen monthly rankings, 2009

Rank	Website	Unique Audience (000)
1	Yahoo News	40811
2	MSNBC Digital Network	35571
3	AOL News	24358
4	CNN.com	20739
5	NYTimes.com	18520
6	Google News	14737
7	Fox News	12650
8	ABCNEWS	10331
9	washingtonpost.com	9810
10	USATODAY.com	9311
11	TheHuffingtonPost.com	9073
12	LA Times	8522

(Nielsen, 2010; adapted by PEJ, 2010)

The most-viewed and most-emailed lists reflect choices of massive audiences. The lists are short-term snapshots of choices, in that the rankings are updated throughout each day, changing at about the same pace as the Web site's news content is updated. The benefit of most-viewed and most-emailed lists is the information captures, albeit in a brief period, the actual exposure of online users to news content through their actions when they visit a news site. Articles included on the most-viewed lists are those clicked on most often by online visitors during a given time period. Likewise, articles on the most-emailed lists are those in which the online users clicked the "e-mail" option offered for news articles. In ranking the clicks of online users, the most-viewed and most-emailed lists are observational data of aggregate choices, setting real-time priorities of news by online users. Observational data of this sort provides one approach to address the challenge frequently raised in health and mass communication research about the lack of means to measure individuals' actual exposure to news content (Niederdeppe *et al*, 2007; Slater *et al*, 2009). An additional strength of the observational data is it provides a comparison and contrast point to survey data, which involves self-reports by individuals of their media preferences and media use. The most-viewed and most-emailed lists provide transparency to the public's priorities for news within those four news media outlets. Whether the findings apply to the broader population is addressed in the Discussion section.

A third advance in this dissertation is to tie organizational identity to socioeconomic status – categorized by public, private, nonprofit or academic sector – to analyze originating sources of health news. Research in online health information has been lacking in categorization of sources, which is surprising given the extensive

concerns raised over the past two decades in the literature regarding the quality of online information. Another advance in public relations is the use of organizational identity in defining the characteristics of agenda builders.

Chapter 2

LITERATURE REVIEW

Agenda building

Agenda building has developed in the literature along two distinct tracks. One track, developed in media effects research, has been conceptualized as secondary-level agenda setting, or attribute agenda setting. The second track, evolving from political science studies, conceptualizes agenda building in terms of the agenda builders, the involvement and influence of participants who determine policymakers' agenda. An understanding of both tracks is useful in order to conceptualize and operationalize the interactive agenda-building model proposed for this dissertation.

Secondary-level agenda setting

Agenda building, as secondary-level agenda setting, adopts the basic premise of agenda-setting: the news media shapes the public's agenda through the transference of *issue* salience (McCombs & Shaw, 1972) or *object* salience (Dearing & Rogers, 1996). Agenda building also adds another layer of salience, attributes of the source or object of the media's attention, influencing two levels: what the public thinks about and *how* the public thinks about an object or a source (Ghanem, 1997). The importance of agenda building is the attributes provide contextual descriptives of the sources and objects within the news media's stories. The way the sources are characterized in the media are the same way those characteristics are reflected in public opinion (Carroll & McCombs, 2003; McCombs, Llamas, Lopez-Escobar & Rey, 1997; Walters, Walters & Gray, 1996).

Attribute agenda-setting has generated strong disagreements in the literature regarding sources and issues that comprise the news. Agenda-setting scholars argue

agenda building is interchangeable with agenda setting in the media's transference of salience to the public (McCombs, 2004). Alternatively, scholars for framing argue that source attributes and other contextual aspects of news content are not simply a matter of media influence on the public, but more deeply rooted in socioeconomic and cultural influences on news (Pan & Kosicki, 1993). One resolution to the scholarly debate, which has gained traction in the literature, has designated agenda building with the media agenda as the dependent variable and originating sources as the antecedents of news (Scheufele, 2000; Scheufele & Tewksbury, 2007). In turn, agenda setting retains the public agenda as the dependent variable, with the media agenda as an antecedent. The distinction follows the path of political science literature in conceptualizing agenda building as the agenda of policymakers, or policy agenda building (Dearing & Rogers, 1996; Scheufele, 2000).

Policy agenda building

Policy agenda building was proposed as a process in which participants interacted to influence issues considered by policymakers (Cobb & Elder, 1971; Elder & Cobb, 1984; Lang & Lang, 1980). The players could be lawmakers, news media, advocacy groups, and the public. The policy agenda is important in terms of what issues make the agenda, and which issues do not (Cobb & Elder, 1971). The presumption in policy agenda building has been that the interaction of agenda builders is behind the scenes of the public's purview until the issue bubbles to the conscious of the public, which typically has occurred once the issue receives some variation of combined attention from policymakers and journalists (Cook *et al*, 1983; Dearing & Rogers, 1996; Protess *et al*, 1997). Agenda building as a behind-the-scenes process is reinforced in the

concept of information subsidies (Gandy, 1982), in which news media receive interviews and public relations material from organized sources in the sources' attempts to influence news coverage.

The limitation of agenda building as an antecedent to agenda setting (Dearing & Rogers, 1996; Scheufele, 2000) is it ignores the interactive connections between originating sources and the public. Subsequent studies linking agenda building and information subsidies have characterized organizations and individuals, particularly political candidates, as acting strategically in trying to present themselves in a positive light to the news media *and to the public* (Golan & Wanta, 2001; Kiousis & McDevitt, 2008; Kiousis & Mitrook, 2002; Ohl, Pincus, Rimmer, & Harrison, 1995). From a public relations perspective, strategic positioning with key audiences, developing relationships with news media and key publics, is an essential function in constructing an organization's public identity and preserving its reputation (Fombrun & Shanley, 1990; Zoch & Molledo, 2006). Strategic communication, in which organizations communicate directly with the online public, is increasingly important as the public continues to transition to Internet-based communication (Kelleher, 2009; Spaeth, 2009; Sweetser & Metzgar, 2007).

An integrative approach to agenda building was proposed by Meijer & Kleinnijenhuis (2006). The two authors used issue ownership theory and agenda setting to theorize, and operationalize, the connections between originating sources who "owned" an issue, news media and the public. They found significant influence between an organization's public positioning strategies on a given issue, the amount specific news media covered the organization and the given issue, and the organization's

reputation among publics exposed to that specific news media. The limitation of the interactive agenda-setting study was that the agenda of media and public was focused on a single issue "owned" by the originating source.

Interactive agenda-building

The interactive agenda-building model is more purposeful than previous literature in focusing on the relationships of participants as the central elements in determining the news agenda, and the placement of priority on news issues. Interactive agenda building examines who are the participants in building an agenda of news in a comprehensive sense. The agenda builders are generally categorized into three groups: online health users, health journalists and originating sources of health. Each group is discussed in more detail below. Within each group, participants are making cognitive choices in a competitive media environment. Health news is the consensus point of the three groups. The influence is measured to the extent there is consistency within each agenda-building group, in terms of characteristics, and between the groups, in terms of priorities for health news.

The outcome for agenda builders is public news. The consensus of originating sources, news media and online users completes the transformation of health information into health news, as a means to build health knowledge, and to affect health beliefs and behaviors.

Health news agenda

To answer the question of who are the agenda builders contributing to online health news first requires a short explication of health news agenda. In agenda setting, the concept of public agenda was originally conceptualized by McCombs & Shaw

(1972) as public opinion about issue salience, or priorities for news issues. A shortfall of public opinion is it is not a direct measure of the public's exposure to news, nor per se, of individual use of news media. Public opinion is typically captured through surveys of individuals, which pressure individuals to have opinions and to recall what issues are important to them, both factors which are temporary and may not reflect individuals' actual beliefs (Zaller, 1992). The media agenda, conceptualized in agenda building as the antecedent to the public agenda (Dearing & Rogers, 1996; Scheufele, 2000), is comprised of news coverage, which varies in operationalization from one study to the next, depending on the content chosen for analysis. Agenda-building effects have been found in various media, from local news (Weaver & Elliott, 1985; Tan & Weaver, 2009) to international news (Kiousis & Wu, 2008), examining the effects on news media channels, rather than impact of specific issues, such as health news.

The present dissertation offers an alternative of agenda building as a process which generates an agenda of news, priorities within a specific issue, based on choices of online users, news media and originating sources. For the present dissertation, the news agenda of interest is health news. The news agenda is examined in terms of the choosers in generating the news and the extent there is a predictable pattern of choosers and choices, in contrast to agenda setting, which predicts a pattern of salient issues. There are indications of stability – and predictability – in the generation of a health news agenda. The routines of online users, in their health scanning habits and news preferences, may contribute stability. Another element of stability is mainstream news media, which produce professional news online and attract millions of online users. The extent originating sources within health news are stable, and predictable, is explored in

the framework of Organizational Social Identity Theory and analysis of sources categorized by organizational identity. Thus the focus of the news agenda is on the relationships of the agenda builders that compose priorities for health news. Health news is broadly defined as new information that conveys health information through news media channels. (See further explanation of health news in the Methods section and Appendix B coding summary.)

The literature provides insight into the characteristics of each of the groups of agenda builders: online users, described as health scanners; journalists, described as mainstream news media; and originating sources, described in categories of social identity.

Health scanners

Online users contribute as agenda builders to a health news agenda as consumers of news. Health scanning involves choices by online users, as part of news consumption, to be exposed to news that provides information about health (Niederdeppe et al, 2007; Shim, Kelly & Hornik, 2006).

There are various conceptualizations of health scanning that offer characteristics to describe online users as potential agenda builders of health news. Health scanning as a microlevel activity has been presented in the literature as both contrasting and complementary to online health seeking (Johnson *et al*, 2006; Niederdeppe *et al*, 2007; Shim, Kelly & Hornik, 2006; Slater *et al*, 2009). Both activities, whether contrasting or complementary, occur in tandem: Online users tend to engage in both health seeking and scanning to acquire health information (Johnson *et al*, 2006). Health seeking has been defined as online users purposefully seeking health information on a specific health

issue, such as a medical condition that they or others they know have been diagnosed with (Cline & Hayes, 2001; Cotten & Gupta, 2004; Goldner, 2006; Morahan-Martin, 2004). The planned activity of information seeking is driven by specific motivation, to reduce uncertainty about health status and increase confidence (Cotten & Gupta, 2004; Josefsson, 2006). Health scanning is less purposeful and more passive, although scanning is not an entirely passive process (Niederdeppe *et al*, 2007). Scanning has been referred to as incidental exposure to information "that a person just happens to 'run across' in his or her information travels" (Tian & Robinson, 2008).

Health seeking and health scanning are important avenues for individuals to acquire health information and build health knowledge. Scholarly work has identified a knowledge gap that has contributed to a health status gap between the "health haves" and the "health have-nots" (Brodie et al, 2003; Dutta Bergman, 2004; Stryker, 2008 Viswanath & Finnegan, 1996). Two key variations in health information acquisition via scanning that are considered here are media use and frequency of media use.

Health scanning frequency

Frequency of online health scanning is directly related to the online users' routines of media use (Althaus & Tewksbury, 2010; Dutta-Bergman, 2004). Scanning as an online activity can be described along two types (Case, 2007). One type of scanning is serendipity scanning, in which the online user, during an active health search on one user-defined topic, discovers valuable health information on another topic. The acquisition of health information is serendipitous. The other type of scanning conceptualized by Case (2007) is browsing, in which the online user is scanning general information and is incidentally exposed to health information. In this dissertation's

proposed model of interactive agenda-building, the serendipitous scanner is not directly engaged in news media as an agenda-builder of online health news agenda. The browsing health scanner is making active choices in terms of routine use of news media. For this dissertation, frequency of health scanning is defined by the online user's perceived exposure to online health news: high scanners are defined by a daily habit of reading health news online, or averaging about 30 times per month. Moderate scanners are considerably less active, with exposure to health news online two to four times per month (bi-weekly or weekly). Low scanners essentially have no habit of regular online scanning, in that they read health news online rarely or less than once each month. (See Appendix A for survey questions).

Health seeker and scanner demographics

The characteristics of health scanners have tended to follow health seeker characteristics, which is not surprising considering that high health scanners tend to be highly active health seekers as well (Tian & Robinson, 2008). Frequency of online health seeking and scanning has been consistently found to be higher among individuals based on two characteristics: gender, with health seeking and scanning more frequent among women than men; and education, with frequency of online health seeking and scanning increasing as education levels increase (Bansil & Keenan, 2006; Cotten & Gupta, 2004; Pew Internet, 2005; Rice, 2006; Shim, Kelly & Hornik, 2006). Gender has also been a strong predictor for online users who acquire health information for people other than themselves (Pew Internet, 2005), which complements research showing females take primary responsibility as caretakers and health decision-makers for their

family (Goldsmith, 2003) and may motivate females to pay more attention to health articles when they browse online for news.

Race has been identified as a predictor in health seeking, with whites more active seekers, (Bowen, 2003) and health scanning, specifically African Americans more frequent scanners (Shim, Kelly & Hornik, 2006). But race as a predictor of health scanning may be confounded by Internet experience (Rice, 2006), as more people of all races gain experience online. Similarly, age as a predictor may be confounded by Internet experience. Findings have often correlated older adults and those with fewer years of Internet use as less likely to use health information online (Anderson, 2004; Bansil & Keenan, 2006; Pew Internet, 2005). But given that more adults use the Internet, and have increased their experience online, age is not anticipated to be a predictor for health scanning frequency.

Predictions of health seeking and scanning based on self-reported health status have conflicting findings in the literature. Some studies have found individuals who self-reported fair to poor health, and those with chronic disease, were more likely to seek health information online than individuals reporting good to excellent health status (Baker *et al*, 2003; Bansil & Keenan, 2006; Pew Internet, 2000; Goldner, 2006). However, an analysis of a substantially larger population sample than previous studies found online health seekers, compared to the offline group, were more likely to report good health status (Cotten & Gupta, 2004). Specific to health scanning, Shim, Kelly & Hornik (2006) found high scanners were more likely than low scanners to engage in exercise, to eat fruit regularly, and to not smoke, activities highly correlated with people who self-report good to excellent health status.

An overriding concern is the means to build health knowledge through health information, including health news. Differences in health knowledge contribute to health disparities, a persistent gap, defined by socioeconomic status, between populations who experience good health outcomes and poor health outcomes (Viswanath, et al, 2006). Education and income are the strongest differentiators in the health knowledge gap, while race varies in its effect on health knowledge (Stryker *et al*, 2008; Viswanath, *et al*, 2006). Because education and income are highly correlated, the present dissertation assumes correlation between education and income, and addresses only education as a characteristic of health scanning frequency. Gender is not a specific factor in health disparities (Gorman & Read, 2006). Although women are more likely to seek health information than men at all socioeconomic levels, the knowledge gap in health information is still largely driven by differences in education (Pandey, Hart & Tiwary, 2003).

Drawing from previous findings, the characteristics of health scanners are hypothesized as follows:

Online health scanners will be similar to online health seekers in these attributes:

> *H1a: Scanning will occur at the same frequency as seeking: high health scanners will be high health seekers, moderate scanners will be moderate in health seeking; low scanners will be low in health seeking.*
>
> *H1b: The more educated the health scanner is, the more frequent the online user will scan health news.*
>
> *H1c: High health scanners will more likely be female.*
>
> *RQ1a: Are there significant differences in race or age between high scanners, compared to moderate and low scanners?*

RQ1b: How does self-reported health status vary among high, moderate and low health scanners?

Health scanner media use

In addition to frequency of online scanning, general media use is a critical variation in individuals acquiring health information (Johnson, Andrews & Allard, 2001). Health scanning differs from health seeking in that scanning is a more common activity, occurring in routine, day-to-day use of media (Niederdeppe *et al*, 2007; Shim, Kelly & Hornik, 2006). The possibility of an individual being incidentally exposed to health information varies based on the type of news media, whether it offers more health news. Some individuals choose to be embedded in media environments rich with health information (Johnson, *et al*, 2006). Individuals who have higher "health consciousness" or are more "health oriented" are more likely to choose health-oriented media and to pay greater attention to media health information (Dutta Bergman, 2004, 2006). Health-oriented individuals are more likely to engage in both health seeking and health scanning behaviors, and to acquire health information from multiple media channels, more so from print media than TV (Dutta Bergman, 2006).

Health scanner choices of news media

Scanners' uses of news media are not necessarily dependent on their uses of health information. To that end, Goldsmith (2003) raises an important distinction in defining the health "user" of information that applies to both health seeking and health scanning activities. He questions conventional terms of "patient" or "health consumer" in describing users of health information. The term "patient" represents only the small fraction of people (about 5 percent annually) that are actually using the health system at

a given time. Thus the "patient" perspective would not encompass the tremendous interest the public has expressed in online health information (Pew Internet, 2009; Rice, 2006). Moreover, the "patient" label ignores the extensive role of the family member who serves as a guardian and caretaker, particularly in the home setting, who takes responsibility for making health decisions. Further, Goldsmith (2003) explains, the term "health consumer" commercializes health decisions in a way that shifts power away from consumers. Alternatively, online "users" of health information are characterized by *actual* choices, rather than *available* choices online.

In health seeking, online users have access to exponential choices. But actual choices by online users are limited in practice. Online health seekers typically begin with a search engine, few look beyond the first Web page of results, and most rely on heuristic cues, rather than fact-checking, to determine the credibility of sources of information they find in their searches (Bates *et al*, 2006; Eysenbach & Kohler, 2002; Rieh, 2004; Toms & Latter, 2007). Thus the outcome of typical online health searches is the search engine's first-page results, along with accompanying advertisers, both of which draw the attention of online users (Bates et al, 2006; Mackert, Whitten & Garcia, 2008). In health scanning, online users may follow the same pattern of self-limited choices of Internet news, which comes from predictable choices of media use and media content.

Media channels are the first layer of predictability. Online user choices of media channels for news are increasingly concentrated within the top 7% of Web sites; the majority of news channels are online presences of traditional news media, accounting for 80% of Internet news web traffic (PEJ, 2010). Online users who routinely choose Web

sites of mainstream news are expressing a preference for an "authoritative voice" in health information (Duffy & Thorson, 2009).

Health scanner choices of health information

The next layer of choice is the users' selection of content on a given news Web site. Health news is an increasingly popular topic for online news content (Kaiser, 2010), which continues a trend from two decades ago, when news media began to significantly increase coverage of health issues (Wilkes, 1997). The list of articles on the user-generated most-viewed and most-emailed lists represent the consensus of an extemporaneous group of online users viewing news within a competitive media environment. It is a form of "crowd sourcing" in which online users determine what information to bring to prominence (Deshpande & Jadad, 2009). At any given time on each of the news Web sites, there are dozens of story choices on the home page and hundreds more stories accessible through clicks or searches. Selective perception comes into play at an individual level, drawing on findings that the user has more narrow selection of news online than in traditional print media formats (Tewksbury & Althaus, 2000). At the macro level, the most-viewed lists represent the group's public agenda of priorities for news. The duration of the agenda is short, lasting only until the next cycle of most-viewed articles is posted. But the impression is made. The most-viewed list confirms actual news exposure and actual online user preferences for news content.

The choices of news articles can be assessed in terms of range and frequency of topics of the news articles. For the present dissertation, the topic of interest is health news, in contrast to nonhealth news. Previous research has established public interest in health news. Through surveys conducted from 1996 to 2002, researchers found about 4

in 10 adults followed health news closely, with public health news and health policy news the two most important health topics (Brodie *et al*, 2003). More recently, survey research found one in five adults closely followed health news (Pews, 2008) The findings imply interest in health news is consistent for a certain group, which would likely be high health-oriented individuals (Dutta-Bergman, 2004). But consistency of interest in health news may not exist at the macro level of online news agenda-building. First, findings suggest there is only a minority who follow health news closely. Second, agenda building in online health news may follow the timeframe of agenda setting, in which issues are relatively short-lived. Summarizing research findings of agenda-setting, McCombs (2004) set a four- to eight-week timeframe for a typical issue to wax and wane with the media and public. This suggests mainstream news does not carry a consistent interest in health news, nor are originating sources able to gain attention of news media or online users consistently over time. In the online environment, with constant shifting of news focus on issues, there may be even lower expectation health would remain a constant issue on the online news agenda. Thus, the news agenda for health on most-viewed lists would be expected to fluctuate over time:

> *H2a: Frequency in online health news will not be consistent over time, from one month to the next, among articles from mainstream media most-viewed lists.*

Health scanner interactivity: online sharing of news

Sharing of online health news is another important characteristic to consider in the contributions of health scanners to interactive agenda-building. Online health scanners become intermediary sources of news by forwarding health news from an online news source to their social networks via email or social networks. The action adds

interpersonal communication to the interactive agenda-building process. The two step flow concept (Katz & Lazarsfeld, 1955) suggested the influence of individuals who distribute mass-mediated information to others. Recently, Southwell and Yzer (2007) established a framework for understanding the role of conversation in spreading mass-mediated information. Recent research findings have demonstrated conversation about health information, which was initiated as a result of exposure to mass media content, can affect health beliefs and behaviors (Hwang, 2009; Southwell, 2005). There is potential for the health news to be forwarded beyond the initial recipient of the sender, generating a viral distribution of the health information. But that potential is not tested in the present dissertation.

The action to forward health news suggests characteristics of the sender as well as the receiver. Katz & Lazarsfeld (1955) and subsequent studies about opinion leaders (Shah & Scheufele, 2006; Weimann, 1991) have found individuals who are active in sharing news information with others have several leadership characteristics, including early adopter of new ideas and technology and heavy consumer of news media. As such, high health scanners could be expected to email news more often than online users who scan for health news less often. Furthermore, sharing health news online requires the sender to have a deeper level of engagement with health news than mere exposure, in assessing the content and determining that the recipient would want to receive the information. The interactivity encompasses core elements of online exchange (Kiousis, 2002; McMillan, 2006): technically, with the element of emailing or sharing on social media; psychologically, with cognitive and affective assessment of content; and socially, through the sender's relationship with the recipient. The action of forwarding health

news suggests intention of the sender to influence the recipient. It is through intention that the sender builds the agenda of health news with the recipient. Thus, the sender would be expected to list some purpose in emailing health news.

Intention is one means that health news would differ from nonhealth news in what articles are likely to appear on most-emailed lists of news Web sites, in that the sender would have specific, distinctive incentives to e-mail health news that nonhealth news may not have. Caretakers, family members who care for others (Goldsmith, 2003), would also have incentive to email health news to others they care for or to those who help provide care. Additionally, health news carries attributes of news that is most likely to be emailed. In a study of New York Times most-emailed articles in comparison to all news articles the newspaper offered online that day, Bergman and Milkman (2010) found that awe-inspiring articles were the most likely to make the most-emailed list; practically useful, surprising, positive and affect-laden articles were also significantly likely to be on the most-emailed list. These common attributes of most-emailed news could be easily applied to health news. Collectively, the attributes of health news and incentives for senders to email health news would lead to predictions about sharing of health news online as follows:

> *H1d: High health scanners, compared to moderate and low health scanners, will have a higher frequency of emailing health information to others.*
>
> *H2b: Most-emailed news, compared to most-viewed news, will have higher frequency of news articles about health.*

A research question explores incentives for online users to email health news:

> *RQ1c: Why do online users email information about health to others?*

Health sources: News media & originating sources

The contrast of health news, to nonhealth news, extends beyond online user choices to the choices of news producers and originating sources providing journalists information for health news. The literature suggests two areas that health news, and health information, has traits distinguishing health from nonhealth news. The first is journalists' and online users' preferences for originating sources of health news to have expertise and, related to that, the expectation for high quality of information from originating sources. Another area of differences is elite and non-elite media coverage of health news.

Preference for experts

Recent studies on professional health journalists found journalists, which encompass both reporters and editors, had specific preferences for originating sources as providers of information to initiate story ideas and to develop stories in progress. The findings vary due to differences in methodologies, but there are common themes. Viswanath, *et al* (2008) found health and medical science journalists rely on public relations material and prefer scientists and researchers, or health care providers as story initiators, in other words experts in the health industry. While Len-Rios, Hinnant, and Park (2009) found health journalists favor "non-public relations" sources to generate story ideas, and resist using PR material, the most significant non-public source is other news media, an indication of intermedia agenda building. It also suggests journalists turn to other journalists as the "experts" in health news. Further, Len Rios *et al* found health journalists expect originating sources to follow standard public relations protocols, which requires originating sources to have expertise in media relations.

Another recent study paired the preferences of health journalists with preferences of public information officers at public health agencies (Lariscy, Avery & Sohn, 2010). The study found local journalists rely on local public health agencies for information, but believe those local public health sources are lacking in quality and expertise. Furthermore, journalists would prefer federal public health contacts as a means to obtain higher quality information, but those federal contacts have not been accessible to local journalists. In essence, local public health agency contacts are expert originating sources for journalists, but not expert enough for the standards of journalists reporting on health news.

The health journalists' high expectations for expert originating sources to provide information for health news appears related to a larger trend, stemming from health information on the Internet, to set high quality standards for health sources.

Online Health Information Quality & Expertise

Concerns from the health industry about the quality of online health information evolved during the 1990s as the public's use of the Internet grew (Seidman, 2006). Over the past decade, warnings have come from experts in the health industry, particularly physician groups (Silberg & Lundberg, 1997; Lindberg & Humphreys, 1998), about the dangers of Internet health information as being misleading and inaccurate (Impicciatorea *et al*, 1997; McClung, Murray & Heitlinger, 1998) and, at its worst, an unethical source touting fraudulent health cures to earn a profit (Eysenbach & Diepgen, 1998). Advocacy for a quality check of Internet health information was intended to protect online users from potential harm. To address concerns about quality of online information, a variety

of organizations developed tools to measure quality or standards to ensure quality of online information, which failed to be effective (Jadad & Gagliardi, 1998).

Health scanner choices: Originating source preferences

An implicit message to the public from advocates for online health information quality has been to emphasize the importance of expertise in health information, a message that endorses institutional sources rather than personal sources to provide information. Surveys have shown the public's highest preference for sources of health information is physicians, and that preference has increased from 2002 to 2008 (Hesse, Moser & Ruteen, 2010). Preference for physicians is further supported by the public's increasing confidence in the medical system as an institution (Saad, 2010). In a further preference for institutional over personal sources, the public cites health professionals, including doctors, as more common sources of health information than friends or family members (Pew Internet, 2009). One caveat on preference: While the public prefers physicians as information sources, the first place individuals would actually seek health information was to go on the Internet (Hesse, Nelson, *et al*, 2005). But the preference for physicians would translate to the public's expectations for online health information to be provided by medical expertise.

In sum, health care is not the place on the Internet for personal sources to have influence. This is not to infer that personal sources of online health information do not have a presence, or increasing potential, to provide the public with health information. Personal sources online, such as online health support groups, have been growing in popularity (Wright, 2006; Dutta & Feng, 2007). Furthermore, Web 2.0 applications to health information, dubbed Medicine 2.0 (Eysenbach, 2008), are directing individuals to

personal applications of online health information and a process of apomediation that replaces experts with decision-making tools and networks. Yet the current status of online health information, in context of the public's preferences, is that institutional sources of online health information clearly dominate over personal sources.

The public's preferences for institutional over personal health information sources would likely be reflected in their preferences for health news information online as well, and would be carried into journalists' preferences for originating sources who provide information on health. The expectation for expertise is an important element that differentiates health news from nonhealth news. The exception would be nonhealth science news; science and health are considered in the literature as comparable types of journalism (Fjaestad, 2007; Dunwoody, 1999; Viswanath, *et* al, 2008) In contrast, other areas of news, such as politics, fashion, religion and sports, have become crowded with personal sources who gain influence. One measure is the influence of personal blogs on mainstream news (Drezner & Farrell, 2004; Meraz, 2009), in terms of the agenda of news issues and influencing content, although blogs have also been shown to be influenced by mainstream news media (Reese, Rutigliano, Hyun & Jeong, 2008; Sweetzer, Golan & Wanta, 2008). The presumption for health news originating sources to have expertise is contrary to Internet trends for lay users, the new media's version of the armchair scholar, to have increased influence online. The contrast between institutional and personal sources provides a hypothesis for health news compared to nonhealth news:

> *H2c: Health news, compared to nonhealth news, will have significantly more institutional originating sources than personal originating sources in both most-viewed and most-emailed articles.*

H1e: Online users will have significantly greater preference for online health sources from institutional organizations, rather than individual sources representing personal views.

Health news choices: Elite and non-elite media

Differences in health news, compared to nonhealth news, appear in other aspects that involve preferences, including news channels. Knowledge gap research points to a differentiator in health news in terms of media type and health news content. Dutta-Bergman (2004) found high health-oriented individuals rely more on print news, while low health-oriented used broadcast media more often. A health knowledge gap has also been attributed to less attention to health news, and less news coverage of health by region (Slater et al, 2008). Considering that education and income are the strongest variables influencing health disparities, differences in health news would also have socioeconomic differences. In mass communication literature, differences in news based on socioeconomic status are conceptualized as elite and non-elite media.

Elite media is delineated from non-elite in several dimensions, including the news audience, the content of the news, the characteristics of the journalists themselves, and distribution of elite news on a national basis and news agenda (Gans, 1979; Lichter, Rothman & Lichter, 1990; West, 2001). Elite media distinguishes itself from non-elite media in its influence on national power elite: decision-makers in politics, business and other areas who are highly educated, relatively wealthy and in positions to influence others (Marchi, 2008) In agenda-setting research, findings have established intermedia effects by elite media, particularly the New York Times, in setting the agenda for non-elite media (Dearing & Rogers, 1996; Tewksbury & Althaus, 2000).

In interactive agenda building, elite and non-elite media differ in choices of health news based on news audience and content. Overall, individuals with higher education levels consume more news on a daily basis, on the Internet and other media channels, than those with lower education levels (Dutta-Bergman, 2005; PEJ, 2010; Stromback & Kiousis, 2010), thus for online news Web sites, online users are more educated than the general population, but average education levels of vary by news outlets. Elite media attract more highly educated audiences to their Web sites than non-elite media (Quantcast, 2010) and online users with higher educations choose a wider range of elite media for their news than online users with lower education levels (Kohut, Doherty, Dimock, & Keeter, 2008). The content of elite and non-elite is differentiated in story initiation and reporting, as journalists try to reflect the interests and information needs of their news audience (Len Rios et al, 2009; Whitney, 2009). Elite media, in efforts to match the higher health interests of its more highly educated audience, would presumably offer more health news and, in turn, the elite audience would more likely choose health news (Dutta-Bergman, 2004), leading to these hypotheses:

> *H2d: Elite media, compared to non-elite media, will have significantly higher frequency of health news in most-viewed and most-emailed lists.*
>
> *H2e: The differences in health news frequency between elite and non-elite media will be consistent over time, from one month to the next.*

Health news originating sources and social identity

Originating sources also contribute to differences in health and nonhealth news, in elite and non-elite media. Originating news sources have typically been studied in literature on agenda-building, as secondary agenda-setting, in terms of the information

the organizations provide to news media and the public, such as news releases, media materials, advertisements and Web sites (Curtin & Rhodenbaugh, 2001; Golan, Kiousis & McDaniel, 2007; Kiousis, Mitrook, Wu & Seltzer, 2006; Ku, Kaid & Pfau, 2003). Alternatively, agenda builders in political science literature are studied in terms of relationships, which addresses the agenda-builders' engagement in terms of power and influence on the policy agenda (Cobb, Ross & Ross, 1976; Lang & Lang, 1980), rather than specific issues or information. Influence is constructed around participation, with some groups more successful than others in participating and getting their voices heard by policymakers and the news media.

Gandy (1982) uses economic theory as a framework to conceptualize information subsidies as the relationships between issue advocates and the media. More than simply providing information, the exchange is relationship-based as a means to advance the organizations' agenda using power, resources and personal influence. Similarly, Gamson and Modigliani (1989), within the field of sociology, apply framing theory to explain the relationships among participants in the evolving issue of nuclear power, with "sponsor" organizations relying on sophisticated professionals to build relationships with media and various subgroups involved with the issue. In public relations literature, the relationships an organization builds with its publics are symbolic, based on images, and/or behavioral, based on actions (Grunig, 1993). Cumulatively, the literature from multiple disciplines indicates organizations develop relationships with journalists and the public that are developed based on the originating source's identity, a broader umbrella that includes the source's issues and information.

Enduring traits of social identity

The identity of a source can be understood within its social context, through Organizational Social Identity Theory. The social situation of interest is health news agenda-building, and the originating source's interactions with news media and online users. Social identity of organizations is theoretically connected to individual social identity, in that the organization develops a social identity based on interactions and perceptions with others (Albert & Whetten, 1985, in Whetten, 2006). An ongoing debate within the literature has struggled to clarify organizational identity and distinguish the concept from other organizational commitments or functions. Clarity has come through scholarly work to strengthen connections of organizational identity to the construct of individual self in social psychology literature (Albert, Ashfort, & Dutton, 2000; Brown, Dacin, Pratt & Whetten, 2006). The identity for organizations, as with individuals, is social in nature, in that identity is compared and contrasted by group characteristics (Ashfort, Harrison & Corley, 2008). An organization "identifies with" or, from the public's perspective, is "defined by," a characteristic common to other organizations, and contrasted to what the organization is not. Ashforth and Mael (1989) define identity characteristics as cognitive, and emphasize these characteristics do *not* (authors' emphasis) describe the effects or behaviors that result from the characteristics. In applying social identity parameters to agenda building, the organizations engaged in agenda building are distinctive in their identity characteristics; conversely, they may not be distinctive in what issues or information they present. This is does not diminish the importance of issues; rather, it elevates the salience of organizational identity.

In organizational communications, identity characteristics, defined by Albert and Whetten (1985), are "distinctive, central and enduring." In psychological literature, enduring characteristics of individuals are central traits (McAdams, 2008). Likewise, social groups can be categorized on a variety of factors (Tajfel, 1982; Baumeister, 2002). For online health news, the social group is defined as providers of new health information, or health news. However, sources may belong to more than one group of news providers, thus sources that provide health news may also provide nonhealth news. This reemphasizes the importance of organizational identity salience, rather than issue salience. An important constant characteristic for organizational sources is their category of economic sector. Organizations do not change their status as part of the public, private, nonprofit or academic sector, and their sector status determines their interaction in agenda building. The organization's economic sector status is inherent to the organization's group participation, as opposed to issues presented by an organization, which are not constant.

How the organization's identity is communicated with its publics is the path to connect the organization's social identity to its relational identity (Scott, 2008).

Health news sources: Journalist preferences

As participants in agenda building, organizations are judged based on their sector status. Scholarly work in health and mass communication has identified preferences journalists, online users and various publics express for public, nonprofit and academic organizations as sources of information. The preference is not dependent on any particular actions by organizations, but on the character constructed around the economic sector. The character of each economic sector is embedded in social

interactions of the organizations, historical and contemporary. Journalists and online users have expressed preferences for public and nonprofit organizations as sources of health information.

Journalists tend to suspect originating sources with profit-motivated agendas and, as a result, prefer news sources from public or non-profit organizations, rather than commercial, for-profit organizations (Curtain, 1999; Curtain & Gaither, 2008). In health news, the emphasis on quality and expertise of originating sources of health information, as discussed in the previous section, would suggest journalists would prefer public health agency sources. Online users express disdain for commercial entities (Eysenbach & Kohler, 2002; Rice, 2006).

Categorization of originating sources is based on their economic status as a public, private, nonprofit/advocacy or academic organization. Economic status aligns with Gandy's information subsidies concept in using economy theory to explain originating sources' capabilities to produce and distribute information. The broader question posed by the present dissertation is the critical analysis of how health news on the Internet empowers individuals or continues to support existing structures and influences on news production. If the health news environment online differs from the traditional news environment, the expressed preferences of online users for types of originating sources, based on economic status, would assume the following hypotheses:

> *H1f: Online users will have significantly greater preference for public, academic and nonprofit online sources of health news, rather than private sector or commercial sources.*
>
> *H2f: Health news, compared to nonhealth news, will be significantly more likely to rely on public, academic and nonprofit originating sources, rather than private sources, for most-viewed and for most-emailed news articles.*

RQ1d: How do high, moderate and low health scanners differ in preferences for originating sources of online health news?

Finally, applying the distinctions of elite and non-elite news media, the online user choices of health news among those with higher education, and elite media health news, would have greater concerns for non-commercialized health sources, and place a greater emphasis on non-commercial originating sources, as follows:

H2g: Health news in elite media, compared to non-elite media, for most-viewed and for most-emailed articles will have significantly greater reliance on public and academic sources. In contrast, non-elite media will have significantly greater reliance on private-sector sources for health news than elite media.

Chapter 3

SUMMARY OF HYPOTHESES

In summary, health news is assessed in terms of who is building an agenda of online health news and what is the interactive agenda-building environment for health news? The hypotheses address who is building online health news by analyzing health scanner characteristics, and health scanner preferences for sources of health news. The hypotheses then address the interactive agenda-building environment in assessing interactivity of online users, and the involvement of originating sources in health news.

Health scanners

Online user characteristics:
Online health scanners will be similar to online health seekers in these attributes:

H1a: Scanning will occur at the same frequency as seeking: high health scanners will be high health seekers, moderate scanners will be moderate in health seeking; low scanners will be low in health seeking.

H1b: The more educated the health scanner is, the more frequent the online user will scan health news.

H1c: High health scanners will more likely be female.

RQ1a: Are there significant differences in race or age between high scanners, compared to moderate and low scanners?

RQ1b: How does self-reported health status vary among high, moderate and low health scanners?

Online health news frequency over time
H2a: Frequency in online health news will not be consistent over time, from one month to the next, in articles from mainstream media most-viewed lists.

Online sharing of health news
H1d: High health scanners, compared to moderate and low health scanners, will have a higher frequency of emailing health information to others.

H2b: Most-emailed news, compared to most-viewed news, will have higher frequency of news articles about health.

RQ1c: Why do online users email information about health to others?

Differences between health news and nonhealth news

H1e: Online users will have significantly greater preference for online health sources from institutional organizations, rather than individual sources representing personal views.

H2c: Health news, compared to nonhealth news, will have significantly more institutional originating sources than personal originating sources in both most-viewed and most-emailed articles.

H2d: Elite media, compared to non-elite media, will have significantly higher frequency of health news in most-viewed and most-emailed lists.

H2e: The differences in health news frequency between elite and non-elite media will be consistent over time, from one month to the next.

Originating source preferences

H1f: Online users will have significantly greater preference for public, academic and nonprofit online sources of health news, rather than private sector or commercial sources.

H2f: Health news, compared to nonhealth news, will be significantly more likely to rely on public, academic and nonprofit originating sources, rather than private sources, for most-viewed and most-emailed news articles.

RQ1d. How do high, moderate and low health scanners differ in preferences for originating sources of online health news?

Elite vs. non-elite

H2g: Health news in elite media, compared to non-elite media, for most-viewed and most-emailed articles will have significantly greater reliance on public and academic sources.

In contrast, non-elite media will have significantly greater reliance on private sources for health news.

Chapter 4

METHODS

Overview

Two studies were conducted to examine the relationships among online scanners, news media and originating sources of health news. The first study involves a survey of online users recruited from a nationally known consumer health Web site. The second study involves content analysis of originating within-news sources in news articles that appeared on most-viewed and most-emailed lists of four mainstream news media Web sites. While the two studies did not attempt to pair the same online users of health information in collection of data, the studies were designed to provide complementary perspectives on a shared type of health communication: online health news. The mix of methodologies is intended to address the complexities of health communication and consider individual, organizational and societal influences that shape health communication (Kreps, 2008).

Study 1 – Survey of online users

Survey instrument

The survey was developed in collaboration with Web editors of a nationally known consumer health Web site to conduct a survey of online users. The purpose of the survey was to develop a profile of health scanner characteristics as agenda builders. The Web site is affiliated with an established medical organization and the Web site is ranked among the top 10 health Web sites internationally (Alexa.com, 2010). The research was reviewed and approved by the University of Minnesota Institutional Review Board.

The survey was comprised of 23 questions, including demographic questions. (See Appendix A for survey questionnaire.) Construction of the questions was guided by secondary and primary research. Questions, and ideas for questions, were drawn from the Pew Internet & American Life Project surveys of U.S. adults regarding online health seeking activities, taken periodically since 2000, and the Health Information National Trends Survey (HINTS) sponsored by the National Cancer Institute, which has sampled U.S. adults annually or biennially since 2003. Scholarly analyses of these data sets, particularly Rice (2006) and Tian and Robinson (2008, 2009), provided helpful insights. The present survey was also developed and revised based on responses to a pre-test conducted in November 2009 on the same consumer Web site, which indicated a need to reduce the overall length of the survey and to refine a number of questions.

Survey data

The survey was posted between March 12, 2010, and April 13, 2010, during which time the Web site attracted an average of over 530,000 unique visitors per day. On a random sample basis, visitors to the Web site were invited, through a pop-up invitation on their computer screen, to participate in a survey for purposes of research in online communication. The goal of the survey was to collect completed surveys from 1,500 U.S. adult participants. The sampling rate was initially set between March 12 and March 31 for every 1-in-500 site visitor to receive an invitation. The sampling rate was subsequently increased to a 1-in-200 rate in order to complete data collection within a one-month time period. No incentive was provided to participants in the survey. A total of 1,927 participants responded, for a response rate of 4.38 percent, which, based on the

Web editors' experiences with previous surveys conducted on the health Web site, would be considered typical.

The final sample was reduced by several factors. First, respondents who were not based in the U.S. were screened out because we presumed non-U.S. residents were not likely to use or be exposed to U.S. news media sources and would thus skew the results on sources of health news. Several responses were eliminated from the data because they were under age 18. Because the survey allowed subjects to complete the survey without responding to all of the questions, the sample size varies from one question to the next. Descriptives of data responses showed a slight drop-off in responses at the point in the survey where the questions became more complex. Thus, the sample size considered for analysis is 1,670 for the first question (minus non-U.S. residents) and declines by about 400 responses to demographic questions at the end of the survey (N = 1,282 to N = 1,366). A cross-tabs check of missing data showed the drop-outs were proportional across demographic categories. Moreover, the overall large size of the sample provides robustness to validate generalizing the findings to the larger population of online health users.

Measurements

The key variable of interest in Study 1 is health scanning. Health scanners are described in relation to health seeking and sharing of health information, in terms of demographic variables and health status, and in terms of preferences of health news sources. SPSS was used to analyze data.

Health scanning and health seeking activity

The health scanning variable was operationalized as the online user's frequency of using the Internet to read health news (Q1a, Appendix A). The question was posed as a passive position, "to read" the news, in contrast to a more active position "to seek" health information online. Health seeking was similarly operationalized by frequency of online searches for health information. To contrast health scanning from health seeking, the questionnaire was structured to ask subjects three levels of engagement with online health information within one question: "How often do you …" search for health information on the Internet, read health news on the Internet, or receive emails of health information (See Question 1A, 1B, 1C, Appendix A). The response options were scaled within four time periods: just about every day, about once each week, once or twice each month, rarely or never. The time periods were developed to delineate high-frequency online use compared to moderate- to low-frequency online use of health news.

In the data analysis, responses on health scanning and health seeking were converted, so high scanners were recorded as the highest number (Converting "Just about every day" to from =1 to =4) and so on. To determine the extent the two moderate levels of health scanning frequency represented true differences in Internet usage, and not differences related to non-Internet use variables, the two moderate levels were compared in t-tests. Demographics and health status showed no significant differences. But significant differences were found in frequency of online health seeking, $t(900) = -12.08$, $p < .001$, which validates important differences in Internet behaviors between scanners who are high-moderate ("about once a week" scanners) and low-moderate ("once or twice each month).

Demographics

Gender

Among survey subjects, 80.6% were female and 19.4% male (N = 1,360). While this is skewed in comparison to the general U.S. adult population, for the online health environment, an oversampling of females would be expected. Research in online health seeking has consistently found that women are significantly more likely than men to search for health information online (Bowen et al, 2003; Cotten & Gupta, 2004; Rice, 2006). A t-test comparison showed no significant variances between female and male in health status or any demographic variables except for age $t(1326) = 6.17$, $p < .001$. Males in the sample were older, at an average age of 59 (N = 260, SD = 14.12), compared to an average age of 53 for females (N = 1,068, SD 13.39). Overall, the age of subjects followed a normal curve, with the average age at 54 (N = 1,331; SD = 13.71, range age 18-90.)

Race

About 88% of subjects were white (N = 1,342). While the survey recorded responses for six non-white races and there were responses within each category, the counts were low, so they were collapsed into a single category coded as non-white for analysis.

Education

Education was measured as five categories: less than high school graduate, through graduate school. Due to low counts of subjects who had less than a high school education (N = 9), education was collapsed into four categories: high school or less, some college, college graduate (BA or associate degree), and graduate schooling.

Geographic residence

Among regions of the U.S. where subjects are residents (N = 1,630), about 20% of subjects were from the Northeast, 26% from the South, and 23% from the West. The Midwest region had the highest response rate, of about 32%, which would be expected in that the consumer health Web site is affiliated with a medical establishment whose primary location is in the Midwest. While geographic residence is not an element considered in the hypotheses, the geographic dispersion of the survey subjects provides additional validation of the sample.

Internet experience

Responses to the question of Internet experience ('About how many years have you been using the Internet?'), generated a mean of almost 14 years (N = 1,231; M = 13.74), with a standard deviation of about 5 years (SD = 5.43), with nearly all the survey subjects indicating they had at least two years of experience using the Internet.

Health status

Health status sought to record the subjects' perceptions of their health. It is important to note that as a self-report status, the individual's perceptions of his or her health measure the person's *attitude* about their health, not objective measurements of their actual health. Health status responses offer some verification of the sample's validity. The health status question asked subjects how they would rate their own health, as excellent, good, fair or poor, reflecting the same question asked in Pew Internet survey on health (Pew, 2009). In the 2009 Pew survey, the most common response (51%) was "good," followed by "excellent" (29%). In the present study, survey responses resulted in the same ranking, and similar percentages (N = 1,362, M = 2.08,

SD = .07), with the most common response of 'good' at 57%, followed by 20% ranking their health as "excellent."

Health information sharing

Sharing of health information was asked in several questions, to operationalize the concept as interactivity among individuals, as senders and recipients of information. Health sharing encompasses emailing health news, as posed in RQ1.03. The first question on sharing asked subjects whether they talked with family and friends about health issues, and then whether they shared health information they found on the Internet with a variety of personal contacts, including medical professionals, family, friends and coworkers. Face-to-face health information sharing segued to computer-mediated, with questions about sharing health news via email (Questions 9, 10 and 12), and via social networks (Question 14). Subjects were also asked about their technical use of technology, in how they emailed health news (Question 11).

In the data analysis, responses for time frequencies were converted so that the highest number was the lowest (Converting "Just about every day" to from =1 to =4 and so on.)

Preferences for health news sources

Online user preferences for originating sources of health news were operationalized in terms of the types of individuals or organizations that distribute or provide health news, both interpersonal and mass-mediated originating sources. The questions were constructed to separate health news *providers* from health news *distributors*. The delineation is intended to avoid a common confounding in mass communication research of interchanging channels of information with information

producers (Metzger, Flanagin, Eyal, Lemus, & McCann, 2003). While there is certainly overlap between health news distributors and health news providers, the survey results explain how individuals access health news (distribution points) and their preferences for sources within health news (originating sources). Health news access was the subject of two open-ended questions (Questions 2 and 3). A textual analysis of the responses is not part of the present study, but a qualitative view of the responses indicates a variety of sources.

Health news access was further explored through a multiple-select question, which listed both interpersonal and mass media sources of health news (Question 4). To obtain individuals' preferences for originating sources of online health news, three questions captured preferences in different ways. First, subjects were asked about types of health stories, with preferences recorded for seven types of originating sources (See question responses 15.1 – 15.7) on a four-point Likert scale of preference (Do not prefer at all, prefer somewhat, prefer a good amount, prefer very much). Subjects were then asked about their preferences for sources on the Internet, based on the URL suffix for the organizations, using the same four-point Likert scale of preference (See question responses 16.1 – 16.4) The third question asked subjects to think forward, about which sources on the Internet they would most prefer to receive health news or updates on health information from (Question 17).

Data analysis

A variety of analyses were conducted, using SPSS, to test hypotheses and explore research questions. Analysis of the survey data addressed the first set of hypotheses, to assess characteristics of online health scanners as the dependent variable,

and research questions regarding differentiators among health scanner groups (RQ1a and RQ1b). Survey data also addressed online emailing of health news (H1d), and online user preferences for health information sources on the Internet (H1e, H1f). Study 2 addresses the remaining hypotheses.

Study 1 addressed relationships in three aspects of health scanning characteristics. The first (H1a) is that health scanning will occur at the same frequency of health seeking. A paired-samples t-test was used to compare health scanning and health seeking activities. Crosstabulation and correlation for nonparametric ordinal data were used to test the relationship of health scanning with health seeking, education level (H1b), gender (H1c), race and age (RQ1a), health status (RQ1b). Kendall's tau-c, which is specific to nonparametric ordinal variables, was used in order to capture the differences among the health scanning groups, with sensitivity to rankings. Spearman rho is reported along with Kendall's tau-c as a reliability measure. Spearman rho is also used in bivariate analyses. A number of health information sharing variables were compared to provide context in explaining why online users email health information to others (RQ1c) and whether high health scanners email health news more often that moderate or low scanners (H1d). Specifically, crosstabulation and correlation analysis tested health scanning frequency with health-sharing variables: talking with family and friends about health issues; sharing online health information with medical professionals, family, friends or coworkers; emailing health news, receiving health news via email, and posting health information on a social network.

Preferences for originating sources of health news (H1e, H1f, H2f and RQ1d) are presented as descriptive data, then analyzed with t-tests and one-way ANOVA.

Results: Health scanner characteristics

Overall results show health scanners are similar to health seekers in their frequency accessing health news online. A key finding is that health scanning is a more common activity at all levels of frequency. Education and age define characteristics for high health scanners, but not other health scanning groups. And no correlations were found for self-reported health status and the frequency of health scanning among all four levels. Those who access information online more frequently are also more likely to share the information through interpersonal contacts and via computer-mediated communication.

Health scanning and health seeking similarities

Support for H1a was found in three statistical tests. Crosstabulation and nonparametric correlation of health scanning and health seeking showed that as frequency of health scanning increased, so did health seeking. Kendall's tau-c (1,521) = 0.526, $p < .001$, shows a relatively moderate, positive relationship. The progressive increase in scanning and seeking frequency is highlighted in Table 2. The paired-sample t test (Table 2) indicated significant differences between health scanning and health seeking, $t (1,520) = 4.539$, $p < .001$. The positive t statistic indicates individuals scan for health news more frequently than they actively seek health information. The finding supports the theoretical underpinnings of this dissertation, that online users access new health information more frequently by scanning for health news than actively seeking health information.

Table 2. Crosstabulation and correlation: Health seeking and health scanning

		Health scanning frequency				Total
		1 – l ow	2	3	4 – high	
Health	1 – low	143	62	17	6	228
seeking	2	122	327	179	46	674
frequency	3	25	48	205	111	389
	4- high	12	22	42	154	230
Total		302	459	443	317	1521

Kendall's tau-c (1,521) = 0.526, p < .001
Spearman's rho (1,521) = .612, p < .001

Demographics and health status

For demographic factors, correlations among the four groups of health scanners were not found significant for age, education or gender, applying Kendall's tau-C or Spearman's rho (Table 3). Race did measure as significant, but a correlation so close to zero (-0.051) indicates a weak relationship at best, particularly given the smaller sample size for non-whites (N=151). Since both research questions compared high-level health scanners to moderate and low, a bivariate regression was conducted between high scanning and all other health scanning frequency levels, for the demographic and health status variables. Because this analysis used a binomial variable (high scanner and not-high scanners), Spearman's rho was the statistical test. Under these conditions, significant differences were found for education, r (1,291) = 0.0527, p (< .05), and for race, in a negative direction, r (1,282) = -0.092, p (<.01). However, both relationships are relatively weak (less than 0.10). Age and gender did not meet significance levels. As such, there is partial, weak support for H1b, which proposed a relationship between health scanning frequency and education, but the support is weak. H1c, which proposed females would more likely be high health scanners, is rejected. The answer to RQ1a,

which asked whether high scanners differed with other health scanner groups by race or age, is no.

Regarding RQ1b, which sought to explore differences in health status among health scanner levels, neither correlations among all four groups, nor the bivariate correlation with high scanners to non-high scanners, were significant. It is worth noting the lack of relationship between self-reported health status and health scanning holds for health seeking as well, with no significant correlation between health status and the four levels of health seeking frequency.

Table 3. Correlations: Health scanning and demographics, health status

Health scanning (DV): High, moderate-high, moderate-low and low	Sample size	Kendall's tau-c		Spearman's rho	
		Value	Approx. sig.	Value	Approx. sig.
Age	1,270	- .027	.271	-. 031	.269
Education	1,291	.023	.309	.028	.308
Gender	1,295	.017	.491	.019	.495
Race (white/ non-white)	1,282 White = 1,131 Non-white = 151	-.051	.017	-.071	.011
Health status (1 = excellent; 4 = poor)	1,297	-.016	.460	-.021	.439

Health information sharing

For health information sharing, the hypothesis proposed high health scanners were more likely than moderate or low health scanners to email health news. The results of the bivariate correlation (high scanners and non-high scanners) showed a significant, positive relationship, r (1,379) = -0.331, $p < .001$, which supports H2b. The relationship is moderately significant. But the finding is worth considering in context of other health

information sharing activities. High health scanners are more interactive in sharing all types of health information, as summarized in Table 4. They are more active online, with sending and receiving health news via email and social network sites, they are more likely than non-high scanners to share Internet health information with doctors, nurses, family, friends and coworkers (Question 8), and they are more likely to talk with family members about a health issue.

The hyper-interactivity of the high health scanners supports scholarship on health-oriented individuals (Dutta-Bergman, 2004), and it advances the concept of high health scanners as being more influential contributors to agenda building in online health news. This adds an important element of comparison and context to discussing the types of originating sources preferred by online users.

For health information sharing, the hypothesis proposed high health scanners were more likely than moderate or low health scanners to email health news. The results of the bivariate correlation (high scanners and non-high scanners) showed a significant,

The question of why individuals email health news (RQ1c) provides further understanding of online users' contributions to online news agenda building. Descriptive data is provided in Table 5. Since subjects could choose any or all of the response options, the percentage sum does not equal 1005. The most common response, that they knew the person would be interested in the article (85%, N = 839), and the second most common response (55%) suggest it is primarily a matter of previous knowledge about the recipient that is motivating the email sender's actions. Given the findings on health information sharing, a bivariate correlation was conducted of high to non-high health scanners, which shows high health scanners are significantly more likely to look at

future interactions with the recipients of their health news email. High health scanners are more likely to indicate they intend to talk with the person about the issue in the future, r (801) = .159, $p < .001$, that they were trying to persuade the recipient to take action based on the article, r (801) = .133, $p < .001$. Although these relationships are relatively weak, the significance is in a positive direction and reveals the various inclinations of senders to email health news.

Table 4. Bivariate correlations: High health scanning and information sharing

Health scanning frequency: High, not high	Sample size	Spearman's rho	
		Value	Approx. sig.
Sent an email with health news	1,379	.331	< .001
Received an email with health news	1,365	.243	< .001
Shared health news on social network site	1,367	.155	< .001
Share Internet health information with	1,394		
… a doctor		.105	< .001
… a nurse		.095	< .001
…family members		.036	.185
…friends		.093	.001
…coworkers		.095	< .001
Talked with family members about health	1,390	.203	< .001

Table 5. Frequencies: Why online users email health articles

Why did you email a health article (or articles) to someone within the past six months? (Check all that apply.)	
N = 838	**Percentage frequency**
You sent the article because you knew the person would be interested in it.	85%
You intended to talk with the person later about the information you emailed.	28%
The person asked you to find information on the Internet about that issue.	24%
You had talked in the past with this person about a health issue.	55%
You were trying to persuade the person to take the action described in the health article.	20%
You thought the information would be useful for the person because he or she is taking care of someone with a health problem.	35%

Results: Health news preferences

Overall, hypotheses for online users' source preferences for health were confirmed. Online users expressed lower preference for personal sources than institutional, and lower preference for sources from commercial, for-profit organizations than for government or nonprofit sources of health news, or new health information online.

The first set of analyses examined responses to Question 15, what type of health stories online users prefer. Table 6 provides a summary of descriptive data. H1e was supported in finding online users prefer institutional organizations compared to individual sources presenting personal views. H1c comparisons were analyzed using paired-samples t tests. Individual/personal sources were compared to all other types of institutional sources from Question 15 (medical experts, university researchers, government regulators and government health campaigners, nonprofit organizations, and commercial businesses). All comparisons were significant, with personal preferences preferred less than medical experts, university researchers and government regulators at $p < .001$; and $p < .05$ for government health campaigns. The one exception was online users indicated they preferred personal sources more than commercial, for-profit sources, $t (1,340) = 7.703$, $p < .001$. That finding is consistent with paired-samples t test findings to test H1f, which proposed online users would have greater preference for public, academic and nonprofit online sources, rather than private sector or commercial sources.

H1f was supported in paired-sample t tests. Online users expressed significantly greater preference for any sources of health news than commercial, for-profit organizations at the $p < .001$ significance level.

Table 6 Online user source preferences based on type of news story

Q15: For news on health, to what extent do you prefer these types of health stories?						
			Percentage frequency			
	Sample size	Mean Scale 1-4 4=prefer much	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
Doctors or other medical experts in the news	N = 1,367	2.34	23%	36%	25%	16%
University or institute research	N = 1,366	2.98	5%	25%	37%	33%
Personal stories	N = 1,357	2.09	30%	40%	21%	9%
News from government regulators (crises news)	N = 1,366	2.66	11%	33%	35%	21%
Nonprofit organization campaigns	N = 1,358	2.22	22%	42%	28%	8%
News from government health campaigns	N = 1,343	2.17	25%	42%	24%	9%
Commercial, for-profit business news on health	N = 1,355	1.86	42%	36%	17%	6%

Table 7 Online user source preferences based on URL suffix

Q16: To what extent do you prefer these types of sources on the Internet that provide health information?						
			Percentage frequency			
	Sample size	Mean Scale 1-4 4=prefer much	Do not prefer At all	Prefer somewhat	Prefer a good amount	Prefer very much
Government Web sites (.gov)	N = 1,329	2.62	15%	31%	31%	23%
Commercial Web sites (.com)	N = 1,335	2.44	16%	37%	35%	12%
Nonprofit or association Web sites (.org)	N = 1,440	2.62	11%	34%	37%	18%
Personal Web sites/ blogs by individuals/groups	N = 1,334	1.76	47%	35%	13%	5%

Online user preferences were then tested using a second aspect of originating sources, the types of sources on the Internet based on URL suffix (Question 16). A summary of descriptive data is in Table 7. Government sites were defined as .gov, commercial sites as .com, and nonprofits as .org. Personal Web sites were described as being .com, .org or .net. In this paired-sample t test, online users expressed an even stronger rejection of personal Web sites as sources of health information, in support of H1e. All source types showed significant preference for .com, .gov and .org rather than personal Web sites at $p < .001$.

H1f was once again supported in that online users expressed greater preference for government (.gov) and nonprofit (.org) Web sites than commercial Web sites (.com). A third test, Spearman's rho correlation, compared ordinal rank preferences for health news stories from commercial, for-profit businesses (Question 15.7) and commercial Web sites (Question 16.2). The test found a significant relationship, r (1,309) = 0.409, p (< .001), that is relatively moderate, but positive, which shows online users were consistent in their lower ranking of preference for commercial entities.

A final set of analyses for Study 1 compared differences in health news source preferences among groups of health scanners. The first analyses focused on survey responses to Question 15, which asked subjects their preferences for types of health stories for seven different categories of sources. The findings from Question 16, which asked subjects their Internet health source preferences based on URL suffix, were then analyzed and compared to findings for Question 15. A third question (Question 17) addressed what would be the subjects' most preferred source for health information on the Internet.

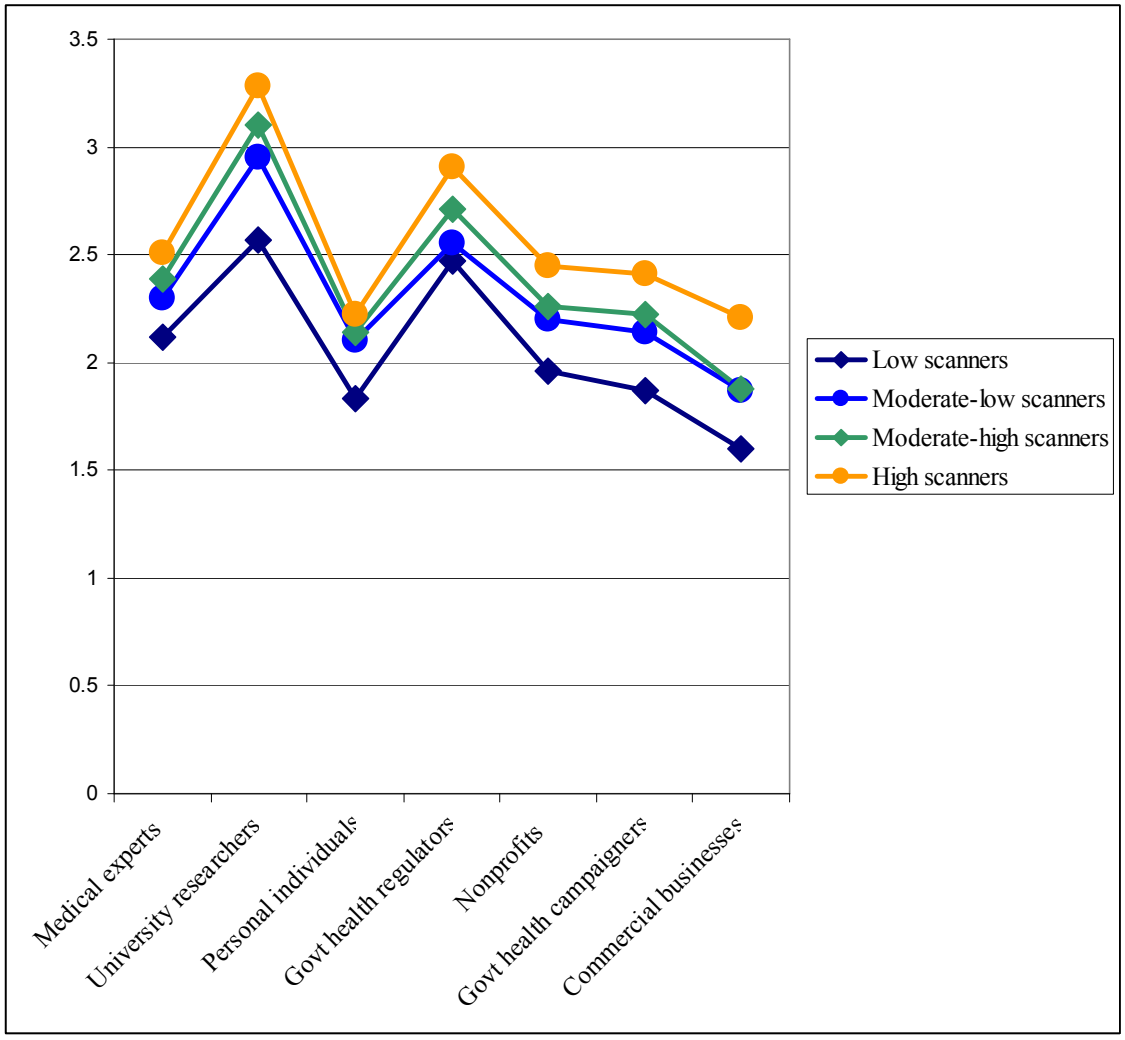
Descriptive data provides a starting point to explore RQ1d (Table 6). One important aspect of frequencies to note is that, in the combined four health scanner groups, the mean for six of the seven health news source categories were between 2 and 3, so health news sources averaged in the "prefer a good amount" response, an overall positive direction for health news sources. Commercial, for-profit businesses are the one health news source that averages below 2, with a mean of 1.86 for the combined health groups ("prefer somewhat").

Since crosstabulations and nonparametric correlations showed significant differences among the four health scanner groups for each type of health news source, at the $p < .001$ level, a one-way ANOVA tested preference differences among the four health scanner groups, with types of news sources. Tests for homogeneity of variance showed some significant groups, so statistics reported here are Brown-Forsythe statistics. Tukey post-hoc comparisons indicated significant differences between health scanner groups.

Low and high scanner differences

A clear pattern emerged for health news source preferences of low and high scanners (Summary in Table 8). Low health scanners, compared to other health scanner groups, choose a lower preference ranking for all types of health news sources except for one. The exception was low health scanners did not differ significantly with moderate-low health scanners in preference for government regulators who provide health news about crises, such as flu outbreaks. Conversely, high health scanners consistently gave the highest rankings for all health news sources among the health scanner groups.

Table 8. Means of health scanner source preferences



Moderate-low and moderate-high scanners

Moderate-low and moderate-high health scanners differ significantly in health news preferences from the low and high health scanner groups for most health news sources. But the two moderate groups share a common sense of source preferences. Moderate-low and moderate-high scanners did not differ significantly in preference for any of the sources. The one health news source where both moderate groups had no

significant difference with the high health scanners was personal/individual sources of news, which all three groups ranked relatively low.

Findings from Question 16 followed a similar pattern of significant differences as Question 15, this time regarding preferences for types of health Web site among health scanner groups. One-way ANOVA, with Tukey post-hoc comparisons, showed significant differences at all levels of health Web site types. High health scanners had significantly higher preferences for most source types, low health scanners had significantly lower preferences for most types, and moderate-low and moderate-high groups differed significantly only for one health news type, government Web sites. An additional finding in support of H1e is that each health scanner group ranked personal Web sites the lowest preferred type of health Web site.

Subjects could choose only one health source for a final question on sources (Question 17), asking which source they would most prefer for updated health information or health news. Based on frequencies (Table 10), about one in three online users would prefer a university research center for health (31.9%), followed by nonprofit health association (about 21.8%). The least preferred was commercial health business (3.4%), followed by individuals or groups (8.8%).

Table 9. Online user most preferred Internet source of health news

Most preferred Internet health source	Frequency	Percent
Doctor's office/ clinic	217	16.2
Nonprofit health association	292	21.8
Govt public health agency	137	10.2
Commercial health business	45	3.4
University research center	39.7	31.9
Individuals or groups	118	8.8
Total	1342	80.4

To see if most preferred Internet health source could explain health scanning frequency, a one-way ANOVA was conducted with health scanning as the dependent variable. But no significant differences were found among the Internet health sources.

Conclusion Study 1

In summary, findings from Study 1 construct a portrait of health scanners based not on demographics, but on online activities and health news preferences. The hypotheses and research questions that suggested differences in education level, gender, race or age could account for differences among groups based on frequency of health scanning were in general not supported. Conversely, hypotheses and research questions for online user sharing of information and health news preferences were supported, and did find a significant relationship based on frequency of health scanning.

Study 2. Content analysis of originating sources in health news

The second study involves content analysis of originating sources of health news, contrasting observational data with Study 1 survey data. It should be noted that these are two distinct sets of online users within the health news environment. The content analyzed in Study 2 encompasses articles from four mainstream news media sources that appeared on user-generated most-viewed and most-emailed lists of articles on the news media Web sites. This observational data provides insight into the choices of online users of their priorities for new information – including health news – at a given point in time. The actual choices of online users addresses a weakness of prior scholarly work in news content analysis in that there is no assumption about exposure to the news. Most-

viewed news is the most popular news online users are being exposed to. Most-emailed news is the most shared news online users are exposing others to.

The data also provides insight into the news media and originating sources found within the news stories and what their priority issues are. As agenda-building literature emphasizes, when one issue makes it on the news agenda, it is at the expense of another issue that is not on the agenda. The news reflects individual, organizational and societal choices of inclusion or exclusion. The contrast guides operationalization for Study 2 in examining what originating sources make the news, and which do not.

Data Collection

Four mainstream news media were selected for the study: CNN, USA Today, the Washington Post and the Los Angeles Times. There are several reasons these sites were chosen to operationalize agenda building in online health news. First, the news organizations have national news distributions, through either print or broadcast, and through high-volume Web sites. CNN has more international news distribution than the other three. The Web sites of each news media rank in the top 20 of all news media, with monthly visitors ranking from 8.5 million to 20.7 million (PEJ, 2010, Nielsen monthly rankings).

Another reason these four news media were chosen is each of the organizations post on their Web site's home page a list of the most popular news, as generated by online users, labeled as "most popular" or "most read" or "most-viewed" and "most-emailed." The term "most-viewed" is used in this study, as it conveys the actions of the online users. The home page access is important in that online users can be influenced by the choices of other readers. Although this study does not attempt to measure the

potential of a bandwagon effect among online users, the potential for a bandwagon effect is more likely by having the lists easily accessible. (CNN.com also lists the most-viewed video, but content for this study collected "most read" articles. Since the time data were collected in 2009, LATimes.com has dropped its "most-emailed" list and now has a "most commented" list.)

Elite and non-elite media

The news sites were also selected to provide contrasts between elite media and non-elite media. Elite media are operationalized in this study as the Washington Post and the Los Angeles Times, and non-elite media as CNN and USA Today. The operationalization is based on qualitative assessment of differences in two areas: audience demographics and lifestyle news content. A summary of key demographic differences in Table 10 (Quantcast, 2009) shows the Web sites for the two elite media, WashingtonPost.com and LATimes.com, consistently attract more online users with graduate schooling and with annual incomes above $100,000, and consistently have a lower percentage of online users with no college, than CNN.com or USAToday.com.

Table 10 Elite and non-elite news media: audience demographics

	CNN.com	USAToday.com	WashingtonPost.com	LATimes.com
No college	36%	41%	31%	32%
Graduate school	20%	17%	26%	22%
Income: $100k+	30%	28%	33%	32%

Source: Quantcast, 2009

In terms of content, the key differentiator among the news Web sites is lifestyle issue news. Lifestyle news for WashingtonPost.com and LATimes.com is typically focused on places of travel or dining, and food, issues of greater interest to individuals

with higher incomes who can more afford travel and dining experiences. Lifestyle news for CNN.com and USAToday.com is focused almost entirely on celebrity news, particularly in most-viewed articles, issues that appeal to a broader, more plebian, audience (Rodman, 2008).

For data collection, the most-viewed and most-emailed lists were copied from the news Web site into a Word document. Each list was limited to top 10 articles, because USA Today and the Los Angeles Times listed only their top 10 most-viewed or most-emailed stories. For each article on the list, the full story was copied into the Word document. In some cases, the articles were copied into the list several or more days after the list was obtained from the news Web site. The result was some stories from the lists were no longer available on the news Web site. If the stories involved content not generated by the news organization itself, such as Associated Press or other online news sources, the stories were obtained from those sources. Otherwise, the stories were dropped from the list. During certain days, the sites for the LATimes.com and WashingtonPost.com did not have updated lists throughout the day, so those dates were considered missing data and the dates were not included in the analysis.

Jason Erdahl[1], Star Tribune executive director of digital media, explained technical aspects of the user-generated lists (personal communication, April 23, 2009), as follows: The most-viewed and most-emailed lists are generated in real time, depending on the clicks of online users on a given news story. The viewer click transitions the screen from showing the story's headline and introductory paragraph to the full story. As such, the click-through action of the online user involves higher

[1] Erdahl is a member of the National Newspaper Association and, within NNA, the digital media editors group and presented common news industry practices.

engagement with the story than just viewing headlines on the screen. The typical daily pattern of traffic for online news Web sites is to experience three peaks of readership during weekdays: In the morning around 8 a.m., around noontime, and again at late afternoon, between 4 p.m. and 6 p.m. These peak traffic points correspond to conventional work schedules, when workers who are at desk jobs begin their work day, when they take lunch breaks, and before they head home for the evening. The weekend traffic to online news Web sites is typically much lower, indicating a different mix of audiences. A qualitative assessment found once stories made the top 10 list, they often remained on the top 10 throughout the day, particularly from late afternoon into the evening, indicating a bandwagon effect.

Data collection reflected these online viewer habits. The most-viewed and most-emailed lists were collected only during work weekdays, Monday through Friday, to encompass choices of typical online users. The time of day chosen for collection was afternoon, between 3:30 p.m. and 6 p.m. Central Time, to capture online user choices at the end of their work day. On two dates, collection was after 6 p.m. but a qualitative assessment found stories from those two dates were similar in types of stories to dates previous and following, so those two dates are included in the analysis. The late-afternoon time would also reduce the influence of other news media, such as a morning daily newspaper, or radio and TV morning news.

Most-viewed and most-emailed articles were collected from May through October 2009 during every weekday. The data was then narrowed to the first five-day work week in the month. (September dates included in the analysis were Sept. 14-18, due to the Sept. 7 Labor Day holiday.) These 30 dates in the analysis (five dates in each

of six months) included 20 stories for each date, the top 10 most-viewed and top 10 most-emailed, for each of the four news Web sites. However, this did not result in 80 different stories for each date. There were often duplications between most-viewed and most-emailed lists within a single date (overlap N=337), for a sum total of 28 percent most-emailed articles also found on the most-viewed lists. Duplications were also found on both lists from one date to the next. All duplicates were coded only once for consistency, so the sample size does not translate directly to number of stories coded.

Once the data were coded, the data were divided into two separate datasets, most-viewed articles (N=1,251; health news N = 266) and most-emailed articles (N = 1,184, health news = 330).

Data measurements

The unit of analysis for coding was the individual news article. The content of each article was examined for sources cited by the reporter as providers of information. Coders noted the news article date, time of collection, headline and, if available, the byline, including wire service or syndicated source. For example, the Associated Press was the most common wire service byline and People magazine was a frequent syndicated source for articles on CNN.com and USAToday.com. The article's rank in the top 10 list was recorded, although rankings are not a factor considered in this analysis. Content was then coded for originating sources. Because the purpose of coding was to identify originating sources of the information, the sources had to be specifically cited within the story. The approach was to consider where the news reporter obtained the information, by interviewing a person or accessing a document. This eliminated

counting people or organizations that were simply named in the story. Instead, coders focused on who or what provided reporters with the information.

Each originating source within the article was coded only once, so repeat mentions of the cited source were not counted. For organizations that had more than one spokesperson or source providing information in an article, each spokesperson was counted as a source. News articles that were repeated from a prior date copied the codes from that prior date.

News type categories: Health news, nonhealth news

News was categorized as health news and nonhealth news. The operationalization of health news was considered broadly in terms of a question: Does the reader gain health information from the story? Health news was further defined as health information that would leave some kind of impression on readers on an aspect of health. Once news was categorized as health news or nonhealth news, additional sub-categories were assigned to each article, such as business and health, mental health/psychological issues, policy and health, and so on. (See Appendix B for news sub-categories.) Some sub-categories were further refined after initial coding, particularly coding for entertainment health news. A particular issue involved coding a daily celebrity blog by a USA Today columnist, which usually appeared in the top 10 most-viewed list. The blog included numerous postings throughout the day. When a blog entry included celebrity health news, rather than code that date's entire blog as a health news article, the single blog entry was coded as an individual health news story. A clarification was added in the coding for entertainment health news to provide some type of health details. So, for example, news about a celebrity's pregnancy or birth of their

child, or the celebrity's death, was counted as health news only if there were some medical details provided. Other nonhealth entries in the celebrity blog were coded as one article per date, to avoid overweighting the dataset with nonhealth articles that were nearly all private sector (entertainment industry) sources.

Originating source categories

Originating sources were identified and counted, then categorized in two aspects, based on the dissertation hypotheses. First, the source was categorized as a personal source or institutional source. Personal sources were conceptualized as sources whose involvement with a news story is personal. These are individuals without an institutional affiliation or who were not considered to be representing an institutional affiliation. In the news media industry, they are considered "person on the street" or the individuals providing "human interest" perspectives of an issue. Personal sources included employees of organizations who were not representing the management perspective in a news article. Institutional sources are individuals affiliated with a specific organization or representing a specific industry. The professional standards of journalism, which require identifying sources by name and title, provided the information to categorize institutional sources in most cases. Anonymous or unnamed sources were coded only if there were clear indications by the reporter of the category of source, for example, a public official or a spokesperson for an industry. News media cited as sources of information within the article, such as "the Associated Press reported," were coded as news media sources and were not counted as originating sources.

The second level of categorization for originating sources coded institutional sources by socioeconomic type of organization. As conceptualized in the literature

review, organizations at a basic level are socially and economically identified as public, private, nonprofit/advocacy or academic. Descriptions of these four categories are included in Appendix B. There were several clarifications added to the code book to note. Sources from political parties were coded as nonprofit/advocacy sources. But elected officials, even those identified by party affiliation, were coded as public officials due to their decision-making power. Another clarification was when editorial writers or columnists provided their opinion, the originating sources cited in the commentary were coded. If the commentary referenced the author in first person (for example, "I think") then the editorial writer's organizational affiliation was counted as one originating source. But if no originating sources were cited, then the originating source was coded as the commentator, and was categorized based on his or her organizational affiliation. News columnists who were employees of the news media were coded as private sector (news media industry). It should also be noted academic sources were coded as academic, even if the topic involved collegiate sports, because the socioeconomic sector remained academic.

To analyze the originating sources within the unit of analysis, which is the individual news article, the originating source counts were transformed into proportions of totals. Counts for each type of originating source were divided by the sum total of institutional sources, so public originating sources divided by institutional sources is a new variable, percent public, and so on for the other three types of originating institutional sources. Percent personal and percent institutional were divided by total of personal plus institutional sources within each article.

Intercoder reliability

Coding of articles was initiated with three coders: two independent coders and the researcher. The code book was developed by the researcher, then reviewed in detail and discussed with the two coders for revisions, additions and clarifications. Articles from one date, in which the researcher had coded sources in the articles, provided a point of discuss and a guide for the independent coders. The coders had subsequent discussions with the researcher on areas of disagreement in coding. After an initial round of coding articles for June 2009 dates, intercoder reliability was checked and coders discussed differences. One of the independent coders was subsequently dropped from the project; two dates included in the dataset that had been coded by the dropped coder were recoded by the researcher. The final dataset was coded equally by one independent coder and the researcher (15 dates each).

In order to ensure intercoder reliability, over 10 percent of articles in the final dataset were coded by a second coder. The reliability of coding for health news/nonhealth news was relatively high (Cohen's Kappa = .872). In checking reliability of coding for personal or institutional, and the four organizational categories, SPSS reported Kappa statistics could not be computed, a recognized limitation of SPSS software that occurs when, within the entire dataset, one coder has not used a specific value used by the other coder (Lombard, Synder-Duch, Bracken, 2010). Sheskin (2004) states an alternative measure to Cohen's Kappa for measuring intercoder reliability is intraclass correlation coefficient, using Spearman's rho, Kendall's tau-b or Pearson's. One illustration of the close relationship of the correlation measures is that, for health news/nonhealth news, Spearman's, Kendall's tau-b and Pearson's were the same as

Cohen's Kappa (= 0.872). All variables other than health news/nonhealth news were correlated with the three alternatives, with all values relatively close in value. Since Spearman's measurements landed in the middle, reliability is reported for Spearman's as follows: personal/individual sources = 0.777; institutional sources = 0.792; public sources = 0.881; private sources = 0.755; nonprofit sources = 0.822; and academic sources = 0.827. Agreement of .80 or more is considered good, and between .70 and .80 is acceptable in research of exploratory nature (Lombard, Synder-Duch, Bracken, 2010), which the current study in categorization of originating sources would be considered.

It should be noted, in following standard intercoder reliability tests, the intercoder reliability comparisons were conducted with actual counts, not portions of the types of originating sources. Had reliability comparisons used source proportions of totals, reliability would likely be stronger, particularly for private sources.

RESULTS

Overall, analyses of the frequency of health news and comparisons with health news showed support for some hypotheses and rejection of other hypotheses. The rejected hypotheses provide compelling contradictions to Study 1 and to previous scholarship.

Frequency over time

Analysis of health news frequency over time showed no significant differences. This leads to rejecting H2a, which proposed differences from one month to the next for total most-viewed health news articles. Only most-viewed articles were considered for the analysis: health news articles N = 266; total articles N = 1,251. Crosstabulations and correlations found the month-to-month relationship for health news frequency was

significant, Kendall's tau-c (1,251) = .056, $p < .05$. However, a Kendall's tau value so close to zero (= .056) shows a very weak relationship. To further explore health news frequency from month to month, a one-way ANOVA was conducted. No significant differences were found. (Test of homogeneity of variances was significant, so Brown-Forsythe statistic was used, $F(5, 1223) = 2.023$, $p = .073$) Games-Howell post-hoc comparisons confirmed that there were no significant differences month-to-month. One-way ANOVA was also conducted with health news and health news article frequency date-to-date (N=30). This too found no significant differences. The findings suggest that rather than health news waxing and waning from the online users' agenda, health issues have a consistent presence from one month to the next, and one date to the next.

Online sharing of health news

In addition to examining health news frequency over time, the analysis examined health news sharing, in the form of emailing health news articles. In support of H2b, frequency of health news articles in both most-viewed and most-emailed lists (N=513) was significantly higher among most-emailed lists (M=0.64) than most-viewed lists (M=0.52), $t (2,097) = -4.261$, $p <.001$ in an independent-samples t-test. This complements a finding from Study 1 that high health scanners were significantly more likely than other groups of health scanners to send and receive emails about health news. It suggests the high health scanning group plays a more active role among online users in building an agenda of online health news, in this situation by sharing health information.

Differences between health news and nonhealth news

Differences in health news and nonhealth news were first analyzed by comparing most-viewed articles, with results of statistical analysis. Differences were then run for

most-emailed articles, with statistical results detailed only in tests when analyses of most-emailed results differed from most-viewed article analyses.

Institutional vs. personal originating sources

The first difference considered between health news and nonhealth news was institutional sources compared to personal sources. Because originating source percentages divided total sources between institutional or personal sources for each news article, only the percent institutional were compared with an independent t-test between health news and nonhealth news. For most-viewed articles. the mean for percent institutional was higher for health news articles (N = 266, M = 0.8974), than for nonhealth news (N= 985, M = 0.8875). However, the independent-samples t-test showed the difference between the means was not significant, which means H2c is rejected.

Comparisons of institutional sources for most-emailed news had the same result of no significant differences, rejecting H2c.

Elite vs. non-elite media

A second comparison of health news to nonhealth news proposed elite media would carry more health news than non-elite media, in both most-viewed and most-emailed lists (H2d). To conduct the analysis, news media were transformed into a binomial variable, with CNN.com and USAToday.com articles categorized as non-elite media (=0) and WashingtonPost.com and LATimes.com articles categorized as elite media (=1). An independent-samples t test was conducted with health news and binomial news. For most-viewed articles, the results were significant, but in the opposite direction as predicted, with non-elite media having more health news (N = 659, M = .24)

than elite media (N = 592, M = .19), t (1248.569) = 2.21, $p < .05$, equal variances not assumed. The findings resoundingly reject H2d.

For most-emailed articles, the same statistical test also had the same result, significant differences but with nonhealth news having more institutional sources than health news.

To ensure that coding for entertainment news, which added USAToday.com entertainment blog entries on health as separate health articles, did not introduce bias to the comparison of elite and non-elite media, all health news sub-coded as entertainment was filtered out, and the t test was run again. This time, there were no significant differences in health news between elite and non-elite, which would again lead to rejecting the hypothesis (H2d). Most-emailed articles were likewise tested and also showed no significant differences.

Elite and non-elite media were proposed to have consistent differences, from one month to the next, in frequency of health news (H2e), the supposition being that if there were significant differences between elite and non-elite news in frequency of health news for the overall sample, the differences would also be found between time periods. A new set of variables, a sum of elite and non-elite health news in most-viewed articles for each month, were created for this analysis. A paired-samples t-test did show significant differences over the six months, t (5) = 5.317, $p < .01$, with non-elite health news (N = 6, M = 52.5) carrying more health news than elite health news (N = 6, M=33). The comparison provided support for H2e, and was an additional measure in support of H2d. Because this hypothesis addressed agenda setting, the analysis comparing over

time was conducted only for most-viewed articles, as they would represent the larger public priority of health news.

Originating source preferences in health news

Within health news, the hypotheses compared types of institutional originating sources. In Study 1, H1f was supported in finding that online users had significantly higher preference for public, nonprofit and academic originating sources than for private-sector sources. In the present study, H2f examined how frequency of public, nonprofit and academic originating source compared with private-sector originating sources. The four types of institutional sources were computed as a portion of total institutional sources within each news article (the unit of analysis). The portions were the basis for generating the mean average for each source type. Paired-sample t-tests were conducted to compare means of private-sector sources to each of the other types of sources. Each of the comparisons showed significant differences, in favor of private-sector sources, considered first for most-viewed articles, then most-emailed articles.

Prevalence of private-sector sources

In comparing most-viewed articles, for all comparisons, $N = 265$ and $p < .001$: private ($M = .5346$) was significantly more frequent than public originating sources ($M = .3082$), than nonprofit sources ($M = .0919$), and than academic sources ($M = .0714$). The findings lead to rejecting H2f, and pointing to a contradiction with online users' self-express preferences for sources found in Study 1, which ranked public, academic and nonprofit sources all higher in preference than private-sector, commercial sources of online health information.

To further explore differences among the originating institutional sources, a one-way ANOVA was conducted comparing the four types of originating sources for health news and nonhealth news. Homogeneity of variances tests were significant for public, nonprofit and academic groups, so results are reported using the Brown-Forsythe statistic.

There were significant differences between health and nonhealth news for all four source types, with public and private sources being more frequent in nonhealth news, and nonprofit and academic sources more frequent in health news compared to nonhealth. However, despite the significant differences between health and nonhealth news, the overall result was frequency of private sources was highest. A summary of the means plots, presented in Table 11, confirmed findings of the paired-samples t-tests: Private originating sources were highest, followed by public sources. Nonprofit and academic were the least frequent sources.

In comparing most-emailed articles, results for originating source reflected the same preference of online users: private sources in health news ($N = 326$, $M = .3549$) were significantly more frequent than public sources ($M = .2376$), $t(325) = 3.347$, $p < .01$; than nonprofit sources ($M = .1801$), $t(325) = 5.559$, $p < .001$, or academic sources ($M = .2347$), $t(325) = 3.604$, $p < .001$. Means plots from one-way ANOVA, summarized on Table 12, show most-emailed does have a pattern of originating sources that differs from most-viewed articles. Differences between health and nonhealth news sources were all significantly different, but private-sector sources still dominant in frequency. In health news, public and academic sources were similar in frequency, but neither surpass

private sources in frequency. (Post-hoc comparisons were not possible in SPSS because health news is a binomial variable.)

Table 11. Most-viewed articles: Originating source types

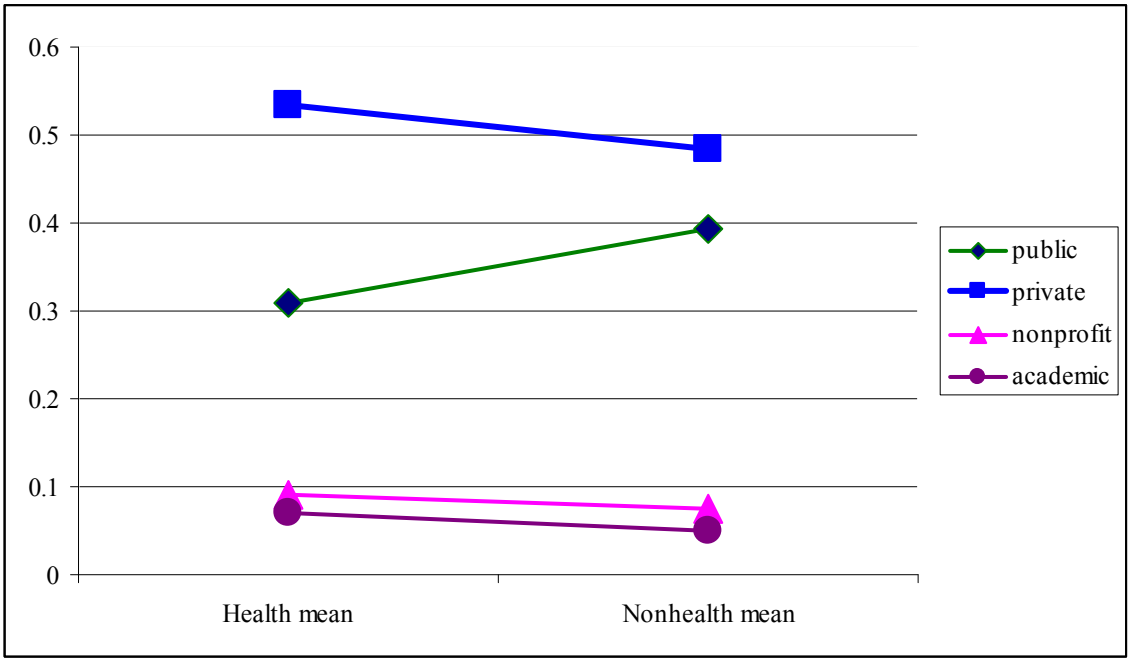
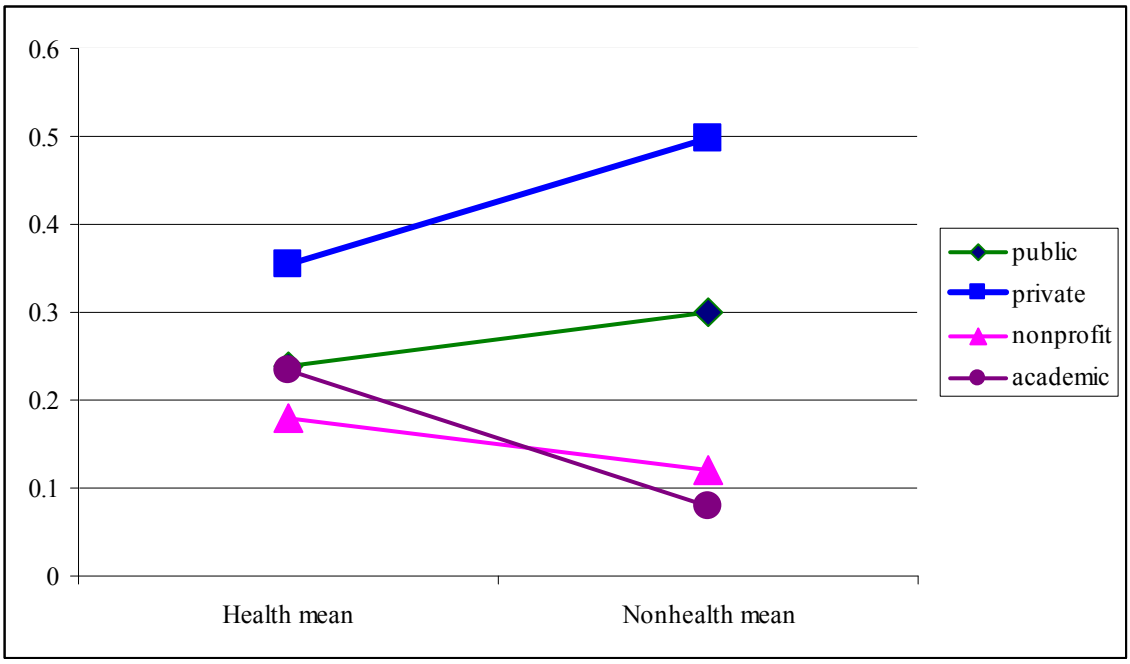


Table 12 Most-emailed articles: Originating source types



Elite and non-elite lookalike sources

To test H2g, which proposed differences in originating sources in elite and non-elite health news, the statistical tests for H2f were repeated for elite and then non-elite media. The hypotheses were that elite media health news, compared to non-elite health news, would have higher frequency of public, academic and nonprofit sources than private sources. Conversely, non-elite health news would have higher frequency of private originating sources.

For elite media health news (N = 109), there was no significant difference found between public (M = .4042) and private (M = .4272) originating sources in a paired-samples t-test, but public sources were significantly more frequent than nonprofit (M =.1123), t (108) = 6.369, p < .001; and more frequent than academic sources (M = .0714), t (108) = 7.494, p <.001. In other comparisons with academic sources, private sources had higher frequency than academic sources in health news, t (108) = -7.163, p <.001. No significant differences were found between academic and nonprofit sources. One-way ANOVA comparing elite media health news with nonhealth news sources found no significant differences between health and nonhealth news for public, private, nonprofit or academic originating sources. Means plots for elite media articles showed the same pattern of source frequencies as all most-viewed article sources, with private sources ranking first, then public sources, with nonprofit and academic sources ranking far below. This would lead to rejecting H2f in part, for elite health news originating sources.

Table 13. Health news elite and non-elite media: originating source means

Paired-sample t-tests	Most-viewed elite health news N = 109	Most-viewed non-elite health news N = 313	Most-emailed Elite N = 133	Most-emailed non-elite ** N = 193
	MEAN	MEAN	MEAN	MEAN
Private	.4272	.4795	.3519	.3569
Public	.4042	.1889**	.3571	.1553
Academic	.0714**	.1910**	.1517**	.2918**
Nonprofit	.1123**	.1376**	.1640*	.1912**

** In comparisons with private sources, significant differences at p < .01.*
*** In comparisons with private sources, significant differences at p < .001.*

For non-elite media health news (N = 313), paired-sample t-tests revealed significant differences, with higher preference for private sources (M = .4795) than any other originating source: public (M = .1889), t (312) = 8.273, p < .001; nonprofit (M = .1376), t (312) = 10.726, p < .001; or academic (M =.1910), t (312) = 7.993, p <.001. This provides partial support for H2g. The one-way ANOVA for non-elite media health and nonhealth news sources showed significant differences for all four types. Again, Brown-Forsythe statistic is reported here, as tests of homogeneity of variances were significant for public, nonprofit and academic at p <.001. Nonhealth news compared to health news has more public originating sources, F(1, 727) = 35.740, p <.001; and more private sources, F(1, 583) = 4.208, p < .05. On the other hand, health news compared to nonhealth news had higher frequency of nonprofit sources, F(1, 478) = 16.290, p <.001; and academic sources, F(1, 414) = 47.7, p <.001.

However, the means plots confirmed findings of the t-tests in that the ranking of originating sources for non-elite news were identical to elite news. Originating sources were significantly more likely to be from the private sector, as a proportion of institutional sources, than any of the other three types. An unexpected finding was that

academic and nonprofit sources ranked lowest in frequency in health news for both elite and non-elite media. In sum, the findings would lead to rejecting H2g.

In most-emailed elite media articles (N = 587), patterns similar to most-viewed articles were found in frequency of originating source types. In paired-samples t-tests of health news (N = 133), there were no significant differences in frequencies between public sources (M = .3571) and private sources (M = .3519), but private sources were more frequent than nonprofit (M = .164), $t(132) = 3.506$, $p < .01$; and more frequent than academic sources (M = .1517), $t(132) = 4.159$, $p < .001$.

For sources in elite most-emailed articles, there were no significant differences between health news (N = 195) and nonhealth news (N = 402) in frequency of public sources (M = .3571) or private sources (M = .3519). Private sources were significantly more frequent than nonprofit sources (M = .1640), $t(132) = 3.506$, $p < .01$; and more frequent than academic sources (M = .1517), $t(132) = 4.159$, $p < .001$. This leads once again to rejecting H2g. In comparisons of health news to non-health news, significant differences were not found for public or nonprofit sources, but were found for private sources, $F(1, 192) = 4.436$, $p < .05$; and for academic sources, $F(1, 188) = 9.836$, $p < .01$.

Surprisingly, non-elite most-emailed health news articles (N = 193) showed a higher mean frequency of academic originating sources than any of the previous analyses (M = .2918), although paired-samples t-tests indicated no significant difference between frequency of academic and private sources (M = .3569). Private sources were significantly more frequent than public (M = .1553), $t(192) = 5.384$, $p < .001$; and more frequent than nonprofit (M = .1912), $t(192) = 4.332$, $p < .001$. One-way ANOVA

comparisons of non-elite health news to nonhealth news sources showed significant differences in frequencies among all four source types, but again means plots confirmed the rank order of private sources as most frequent, followed by academic, public and nonprofit.

Findings of the hypotheses are summarized in Table 14.

Table 14 Summary of hypotheses results

STUDY 1

Online user characteristics			
Hypothesis		**Results**	
H1a	Online health scanners will be similar to online health seekers in these attributes: Scanning will occur at the same frequency as seeking: high health scanners will be high health seekers, moderate scanners will be moderate in health seeking; low scanners will be low in health seeking	Supported	
H1b	The more educated the health scanner is, the more frequent the online user will scan health news.		Rejected
H1c	High health scanners will more likely be female.		Rejected
RQ1a	Are there significant differences in race or age between high scanners, compared to moderate and low scanners?	Race: weak relationship	Age: no
RQ1b	How does self-reported health status vary among high, moderate and low health scanners?		Does not vary
Online sharing of health news			
H1d	High health scanners, compared to moderate and low health scanners, will have a higher frequency of emailing health information to others.	Supported	
RQ1c	Why do online health users email information about health to others?	Varied	
Differences between health news and nonhealth news			
H1e	Online users will have significantly greater preference for online health sources from institutional organizations, rather than individual sources representing personal views.	Supported	
Originating source preferences			
H1f	Online users will have significantly greater preference for public, academic and nonprofit online sources of health news, rather than private sector or commercial sources.	Supported	
RQ1d	How do high, moderate and low health scanners differ in preferences for originating sources of online health news?	Varied	

STUDY 2

Online health news frequency over time

Hypothesis		Results	
H2a	Frequency in online health news will not be consistent over time, from one month to the next, in articles from mainstream media most-viewed lists.		Rejected
H2b	Most-emailed news, compared to most-viewed news, will have a higher frequency of news articles about health.	Supported	
Differences between health news and nonhealth news, elite and non-elite			
		Most-viewed news	Most-emailed news
H2c	Health news, compared to nonhealth news, will have significantly more institutional originating sources than personal originating sources in both most-viewed and most-emailed articles	Rejected	Rejected
H2d	Elite media, compared to non-elite media, will have significantly higher frequency of health news in most-viewed and most-emailed lists.	Rejected	Rejected
H2e	The differences in health news frequency between elite and non-elite media will be consistent over time, from one month to the next.	Supported	(not tested)
Originating source preferences health news and nonhealth news, elite and non-elite			
		Most-viewed news	Most-emailed news
H2f	Health news, compared to nonhealth news, will be significantly more likely to rely on public, academic and nonprofit originating sources, rather than private sources, for most-viewed and most-emailed articles.	Rejected	Rejected
H2g	Health news in elite media, compared to non-elite, for most-viewed and most-emailed articles, will have significantly greater reliance on public and academic sources. In contrast, non-elite media will have significantly greater reliance on private sources for health news.	Elite differences rejected Non-elite differences supported	Elite differences rejected Non-elite differences supported

Chapter 5

DISCUSSION

There are several findings of the research that propose a different kind of interactive agenda-building than was initially proposed in the review of the literature. The findings suggest online health news is developing new ways of interactivity for online users and, in other ways, reinforcing conventional agenda-building influences on health news.

The findings of online health scanner characteristics indicate it is the level of individuals' engagement online that determines their contribution to building an agenda of online health news. Health scanners were not distinguished by demographics or health status. Rather, the more frequently they search for health information, the more they scan health news. High health scanners were also more likely to share health information they find on line. Moreover, high health scanners have higher rankings than low- or moderate-level health scanners, in their preference for all types of health news sources. The conventional forces of agenda-building were seen in the originating source types in health news.

Three important findings explain the interactions and environment of online health news.

Health scanners' constant agenda

The first key finding is that online health scanning is a more common activity than online health seeking. This reinforces a traditional news agenda-setting model. Individuals who scan for information are allowing sources of news to set the agenda for them. On the other hand, a primary presumption of agenda-setting, that issues wax and

wan on the public's agenda, does not appear to apply to online health news. Health scanning as a constant activity is complemented by findings of health news frequency over time. The consistent portion of health news to nonhealth news from one month to the next, and from one day to the next, could be possible only through constant, and consistent, interactions of news media with originating sources, and the interactions of health scanners to set health news as a priority. It suggests originating sources on a daily basis are able to set an agenda of health news. The daily health news agenda is sustained with news media coverage, and online user consumption of that health news.

Extending the news media agenda via social networks

A second key finding is the active role of high health scanners in sharing health information from the Internet via emailing, and through interpersonal contacts. It extends the influence of news media, as the high health scanners adopt the agenda of health news constructed by news media, and the originating sources within the news stories, and share the information within their social networks. Whether online users are moderating or mediating news information in their messages when they share online health information, or whether their information influences recipients, are questions to be explored in the future.

Contradictions in source preference and source presence

A third, and perhaps most, important finding is the contradiction between online user preferences for originating sources and the frequency of presence types of originating sources have in mainstream health news. Online health users, for all levels of health scanners, gave consistent low rankings to private organizations. However, analysis of originating sources in health news articles, among articles on most-viewed

and most-emailed lists for a given date, showed private sources are more frequent than any other type of source. There may be numerous explanations for the contradiction. As noted previously, the two studies involve distinct data, one of survey research of online users and the other of observational data of most-viewed and most-emailed article originating sources of health news. While these are not intended to present the online users as the same, the contrast in findings in the results presents compelling comparisons and contrasts.

A first consideration is that health scanning, incidental exposure to new health information, should not be narrowly defined in the conventional sense of public health news or academic research studies on health. Individuals obtain health news from a broader scope of sources than just health providers, which results in the higher prevalence of private-sector sources. As such, a likely reason for the contrast in preferences is online users are not aware they are choosing health news that so often has private sources providing the information. This elevates the importance of news media coverage for private sources, as indirect communication via news media provides private sources access to individuals.

The commercial nature of news media provides another explanation for the contrast in preferences. The prevalence of private-sector sources in health news may be a symptom of news that is derived from a business structure based on an advertising revenue model. The advertising model also supports the notion of public relations effectiveness, in that private-sector organizations would have more resources. From a public relations perspective, the ability for organizations to communicate indirectly to publics via the news media remains an important part of their communication strategies.

The wide distribution of mass media allows organizations to leverage their communication resources to reach a broader audience at a lower cost than advertising or communicating directly to those publics. At the same time, private organizations presumably have greater access to resources and more resources to invest in developing relationships with news media to obtain news media coverage. A disparity in resources used for public relations could explain why nonprofit and academic sources ranked so low in frequency of originating sources in health news. The contradiction between online users' expressed preferences in Study 1 and the findings in Study 2 of private-sector source prevalence in health news indicate the potential that information subsidies, and the economic resources required by information subsidies (Gandy, 1982), continues to have effectiveness particularly for private-sector organizations.

The dominance of private sources could also be explained in terms of the sheer number of private institutional sources that provide information on health. One area that has been explored in the literature, and appears to find support here as well, is the health information provided by the entertainment industry can result in strong presence of private-sector sources for health news. One example is the death of entertainer Michael Jackson, which triggered numerous news articles about a range of health issues, from prescription drug overdoses to skin conditions, during mid-year 2009 when content was collected.

An additional explanation of the contradiction is online users are expressing socially desirable responses in describing their preferences for health news. Individuals would not want to appear overly influenced by commercial, for-profit entities for an

issue as personal and sensitive as their health. But observational data provides a contrast in actual online user preferences for health news.

The low ranking of academic sources in health news articles also is directly contradictory of online users' expressed preferences for where they would most prefer to obtain health information: from a university center. This disparity suggests an opportunity for university health centers to consider how to deliver more health news more effectively to online users, so the online users do not have their health news agenda set by private organizations, which may not have the best interests of the public's health.

The consistency of findings in Study 2 of originating sources of health news appear to provide initial validation of Organizational Social Identity Theory, and the use of socioeconomic types to categorize the social identity of organizations.

Personal sources low rankings

In constructing the online health news agenda, there is a unifying element for online users, news media and originating institutional sources: their lack of interest in personal stories. While a plethora of scholarship has touted the Internet as an opportunity for individuals to share their health stories, to connect online, and to find support for their health issue, the personal stories of individuals are not a preferred or chosen source for health news. The definitive dismissal of personal sources reinforces a traditional aspect of health news: The emphasis on expertise.

Given the past decade of medical experts warning the public about poor quality of information on the Internet, the finding should come as no surprise that the public does not choose to use personal health stories. At the same time, the lack of interest in

personal stories suggests a limitation to interactivity of online health news as a product of professional journalism and health expertise.

Limitations and direction for future research

The present studies in this dissertation present several limitations. The first limitation, as previously noted, is that Study 1 survey data and Study 2 content analysis of observational data, represent two data sets that may or may not overlap. Thus a direct comparison of findings from Study 1 and Study 2 are not possible. But this does not negate the importance of considering contrasts in findings between the two studies.

Another limitation specific to Study 1 is the sample population's dominance of females. While this is representative of the online health-seeking public, it has limited applications to the general public. A further limitation of Study 1 is the lack of information about non-respondents, those who received an invitation to take the survey but refused, thus there is no ability to test for possible non-response bias.

For Study 2, primary limitations are around the sample and the coding. The sample represents a snapshot of most-viewed and most-emailed articles at a specific time of day, over six weeks in six separate months, and for just four news media. Had the sample been expanded to more types of news media, and been randomly collected at various times, it would likely have affected the results. Another limitation was decisions made in coding health news and, in particular, decisions for coding originating sources within each news article. While intercoder reliability was above .70, this was sufficient for tentative conclusions on the research findings. Reliability closer to .90 would have allowed for stronger conclusions.

A final limitation in the study is it provides only an introductory assessment of agenda-building of online health news. Each entity contributing to agenda building, the online users, the news media and the originating sources, would benefit from further indepth research and analysis to determine their influences on the online health news agenda.

There are a number of areas that warrant more attention in future research. One area in particular is the actions of online users to forward health news to other online users. This study identified high health scanners as active emailers of health news and preliminary findings in the senders' motives for emailing health news. Given the prevalence of health news in user-generated most-emailed news lists, health news emailing is worth further investigation, particularly a qualitative investigation of what types of news is being emailed.

Conclusion

The purpose of this dissertation was to examine who builds the agenda of online health news and to examine the interactive environment of online health news. Through survey research of online users and content analysis of health news, a picture of online health news agenda-building emerged that does not suggest radical empowerment of individuals. Rather, it highlights the importance of public relations efforts in shaping health news that is viewed and redistributed by online users.

Those online users vary not by demographics but by frequency of interaction online and by sharing health news. While online users express preference for health news that is comprised of public, academic and nonprofit sources, observational data of health news content is contradictory. Instead, originating sources of health news are

more frequently private-sector sources. This demonstrates the influence of private sources, and indicates the public relations capabilities of the private sector in setting the health news agenda through use of information subsidies (Gandy, 1982).

In sum, those who are incidentally exposed to health news, through health scanning, express preferences for particular types of health news sources. But an examination of online user choices of health news presents a contrast, with prevalence of private-sector sources. This pattern suggests online health scanning does not necessarily change the agenda-setting environment, and that the information subsidies in traditional agenda building continue. Private-sector organizations with resources to promote ideas through information subsidies (Gandy, 1982) are outpacing academic and nonprofit and, in some cases, public sources, in generating health news despite online users' preference for academic sources. The findings collectively suggest a role for public relations in health news agenda building that represents both opportunity and challenges in providing new health information to a public with a voracious appetite for health information online.

REFERENCES

Albert, S., Ashforth, B.E., & Dutton, J.E. (2000). Organizational identity and identification: Charting new waters and building new bridges. *Academy of Management Review, 25*(1), 13-17.

Anderson, J.G. (2004). Consumers of e-Health: Patterns of use and barriers. *Social Science Computer Review*, *22*(2), 242-248.

Ashfort, B.E., Harrison, S.H., & Corley, K.G. (2008)..Identification in Organizations: An examination of four fundamental questions. *Journal of Management*, 34(3), 325-374.

Ashforth, B.E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, 14(1), 20-39.

Bates, B.R., Romina, S., Ahmed, R., & Hopson, D. (2006). The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical Informatics and the Internet in Medicine, 31*(1), 45-52.

Baumeister, R.F. (2002). Ego depletion and self-control failure: An energy model of the self's execution function. *Self and Identity, 1*, 129-136.

Bennett, W.L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication, 58*, 707-731.

Bergman, J.A., & Milkman, K.L. (2009). Social transmission, emotion, and the virality of online content. *Social Science Research Network*, Dec. 25, 2009. Available at: http://ssrn.com/abstract=1528077.

Brodie, M., Hamel, E.C., Altman, D.E., Blendon, R.J., & Benson, J.M. (2003). Health news and the American public, 1996-2002. *Journal of Health Politics, Policy and Law, 28*(5), 927-950.

Brown, T.J., Dacin, P.A., Pratt, M.G., & Whetten, D.A. (2006). Identity, intended image, construed image, and reputation: An interdisciplinary framework and suggested terminology. *Journal of the Academy of Marketing Science, 34*(2), 99-106.

Carlson, M. (2007). Order verses access: News search engines and the challenge to traditional journalistic roles. *Media, Culture & Society, 29*(6), 1014-1030.

Carroll, C.E., & McCombs, M. (2003). Agenda-setting effects of business news on the public's images and opinions about major corporations *Corporate Reputation Review, 6*(1), 36-46.

Case, D.O. (2007). Looking for information: A survey of research on information seeking, needs, and behavior, 2nd edition. Burlington, MA. Elsevier, pp. 89-96.

Chaffee, S.H., & Metzger, M.J. (2001). The end of mass communication? *Mass Communication & Society, 4*(6), 365-379.

Cline, R.J.W., & Haynes, K.M. (2001). Consumer health information seeking on the Internet: The state of the art. *Health Education Research, 16*(6), 671-692.

Cobb, R.W., & Elder, C.D. (1971). The politics of agenda-building: An alternative perspective for modern democratic theory. *The Journal of Politics, 33*(4), 892-915.

Cook, F.L., Tyler, T.R., Goetz, E.G., Gordon, M.T., Protess, D., Leff, D.R., & Molotch, H.L. (1983). Media and agenda setting: Effects on the public, interest group leaders, policy makers, and policy. *Public Opinion Quarterly, 47*, 16-35.

Cotten, S.R., & Gupta, S.S. (2004).Characteristics of online and offline health information seekers and factors that discriminate between them. *Social Science & Medicine, 59*, 1795-1806.

Cullen, R. (2006). *Health information on the Internet*. Westport, CT, Praeger Publishers.

Curtin, P.A., & Gaither, T.K. (2008). Crisis response and agenda building during the spinach e.coli crises: A mixed-methods analysis. *Presented at the International Communication Association 58th Annual Conference*, May, 2008, Montreal, Quebec, Canada.

Curtin, P.A., & Rhodenbaugh, E. (2001). Building the news media agenda on the environment: A comparison of public relations and journalistic sources. *Public Relations Review, 27*, 179-195.

Curtin, P.A. (1999). Reevaluating public relations information subsidies: Market-driven journalism and agenda-building theory and practice. *Journal of Public Relations Research, 11*(1), 53-90.

Davidson, A.E., & Wallack, L. (2004). A content analysis of sexually transmitted diseases in the print news media. *Journal of Health Communication, 9*, 111-117.

Deshpande, A., & Jadad, A.R. (2009). Editorial: Trying to measure the quality of health information on the Internet: Is it time to move on? *The Journal of Rheumatology, 36*(1), 1-2.

Drezner, D.W., & Farrell, H.(2004). Web of influence. *Foreign Policy*, November/December, 32-40.

Duffy, M.E., & Thorson, E. (2006). Emerging trends in the new media landscape. In (Eds.) J.C. Parker, & E.E. Thorson (pp. 93-116). *Health communication in the new media landscape.* New York, NY, Springer Publishing Company.

Dunwoody, S. (1999). Scientists, journalists, and the meaning of uncertainty. In (Eds.) S.M. Friedman, S. Dunwoody, S., & C.L. Rogers, C. L. *Communicating uncertainty: Media coverage of new and controversial science* (pp. 59-80). Mahwah, NJ, Erlbaum Associates.

Dutta, M.J., & Feng, H. (2007). Health orientation and disease state as predictors of online health support group use. *Health Communication, 22*(2), 181-189.

Dutta-Bergman, M.J. (2004). Primary sources of health information: Comparisons in the domain of health attitudes, health cognitions, and health behaviors. *Health Communication, 16*(3), 273-288.

Dutta-Bergman, M.J. (2006). Media use theory and Internet use for health care. In (Eds.) M. Murero & R.E. Rice, *The internet and health care: theory, research, and practice* (pp. 83-106.) Mahway, NJ, Erlbaum Associates.

Elder, C.D., & Cobb, R.W. (1984). Agenda-building and the politics of aging. *Policy Studies Journal, 13*(1), 115-129.

Eysenbach, G., & Diepgen, T.L. (1998). Evaluation of cyberdocs. *The Lancet, 352*(November 7, 1998), 1526.

Eysenbach, G., & Kohler, C. (2002) How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal, 324*, 573-577.

Eysenbach, G. (2008). Editorial: Medicine 2.0: Social networking, collaboration, participation, apomediation, and openness. *Journal of Medical Internet Research, 10*(3), e22.

Finney Rutten, L.J., Squiers, L., & Hesse, B. (2006). Cancer-related information seeking: Hints from the 2003 Health Information National Trends Survey (HINTS). *Journal of Health Communication, 11*, 147-156.

Fjaestad, B. (2007). Why journalists report science as they do. In (Eds.) M.W. Bauer, & M. Bucchi, *Journalism, science and society: science communication between news and public relations* (pp.123-131). New York, NY, Routledge.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal, 33*(2), 233-258.

Gamson, W.A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *The American Journal of Sociology, 95*(1), 1-37.

Gandy, O. H. (1982). Beyond agenda setting: Information subsidies and public policy. Norwood, NJ, Ablex Publishing.

Gerhards, J., & Schafer, M.S. (2010). Is the internet a better public sphere? Comparing old and new media in the USA and Germany. *New Media & Society, 12*(1), 143-160.

Ghanem, S. (1997). Filling in the tapestry: The second level of agenda setting. In M.E. McCombs, D.L. Shaw, & D.H. Weaver (1997). *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory.* Mahwah, NJ, Erlbaum Associates.

Golan, G.J., Kiousis, S.K., & McDaniel, M. (2004). Second-level agenda setting and political advertising: Investigating the transfer of issue and attribute saliency during the 2004 US presidential election. *Journalism Studies, 8*(3), 432-443.

Golan, G., & Wanta, W. (2001). Second-level agenda setting in the New Hampshire primary: A comparison of coverage in three newspapers and public perceptions of candidates. *Journalism & Mass Communication Quarterly, 78*(2), 247-259.

Goldner, M. (2006). How health status impacts the types of information consumers seek online. *Information, Communication & Society, 9*(6), 693-713.

Goldsmith, J. C. (2003). The consumer: Patient no more. *Digital medicine: Implications for healthcare leaders* (pp. 93-115). Chicago, IL, Health Administration Press.

Grunig, J.E. (1993). Image and substance: From symbolic to behavioral relationships. *Public Relations Review, 19*(2), 121-139.

Hesse, B.W. (2008). Enhancing consumer involvement in health care. In (Eds.) J.C. Parker, & E.E. Thorson. *Health communication in the new media landscape* (pp. 119-150). New York, NY, Springer Publishing Company.

Hesse, B.W., Moser, R.P., & Rutten, L.J. (2010). Surveys of physicians and electronic health information. *New England Journal of Medicine, 362*(9), 859-860.

Hwang, Y. (2009) *Social diffusion of campaign effects: Campaign-generated interpersonal communication as a polarizing mediator of anti-tobacco campaign effects.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

Impicciatorea, P., Pandolfini, C., Casella, N., Bonati, M. (1997). Reliability of health information for the public on the world wide web: Systematic survey of advice on managing fever in children at home. *British Medical Journal, 314*, 1875.

Jadad, A.R., & Gagliardi, A. (1998). Rating health information on the Internet: Navigating to knowledge or to Babel? *Journal of the American Medical Association, 279*(8), 611-614.

Johnson, J.D., Andrews, J.E., & Allard, S. (2001). A model for understanding and affecting cancer genetics information seeking. *Library & Information Science Research, 23*, 335-349.

Johnson, J.D.E., Case, D.O, Andrews, J., Allard, S.L., & Johnson, N.E. (2006). Fields and pathways: Contrasting or complementary views of information seeking. *Information Processing and Management, 42*, 569-582.

Josefsson, U. (2006). Patients' online information-seeking behavior. In (Eds.) M. Murero & R.E. Rice, *The internet and health care: theory, research, and practice,* (pp. 127-149.) Mahway, NJ, Erlbaum Associates.

Kaiser Family Foundation. (2010). *Kaiser health tracking poll*, April 2010. Menlo Park, CA The Henry J. Kaiser Family Foundation.

Katz, E., & Lazarsfeld, P.F. (2006) Personal influence: The part played by people in the flow of mass communications. New Brunswick, NJ, Transaction Publishers. Originally published in 1955 by The Free Press.

Kelleher, T. (2009). Conversational voice, communicated commitment, and public relations outcomes in interactive online communication. *Journal of Communication, 59*, 172-188.

Kiousis, S., & McDevitt, M. (2008). Agenda setting in civic development: Effects of curricula and issue importance on youth voter turnout. *Communication Research, 35*(4), 481-502.

Kiousis, S., Mitrook, M., Wu, X., & Seltzer, T. (2006).First-and second-level agenda-building and agenda-setting effects: Exploring the linkages among candidate news releases, media coverage, and public opinion during the 2002 Florida gubernatorial election. *Journal of Public Relations Research, 18*(3), 265-285.

Kiousis, S., Popescu, C., & Mitrook, M. (2007). Understanding influence on corporate reputation: An examination of public relations efforts, media coverage, public opinion, and financial performance from an agenda-building and agenda-setting perspective. *Journal of Public Relations Research, 19*(2), 147-165.

Kiousis, S., & Wu, X. (2008). International agenda-building and agenda-setting. *The International Communication Gazette, 70*(1), 58-75.

Kohut, A., Doherty, C., Dimock, M., & Keeter, S. (2008). Audience segments in a changing news environment: Key news audiences now blend online and traditional sources. *Pew Research Center Biennial News Consumption Survey*, Washington, DC. The Pew Research Center for The People & The Press.

Kreps, G.L. (2008). Qualitative inquiry and the future of health communication research. *Qualitative research reports in communication, 9*(1), 2-12.

Ku, G., Kaid, L.L., & Pfau, M. (2003). The impact of web site campaigning on traditional news media and public information processing. *Journalism & Mass Communication Quarterly, 80*(3), 528-547.

Lamme, M.O., & Russell, K.M. (2010). Removing the spin: Toward a new theory of public relations history. *Journalism & Communication Monographs, 11*(4), 281-362.

Lombard, M., Synder-Duch, Bracken, C.C. (2010) Practical resources for assessing and reporting intercoder reliability content analysis research projects. Last updated on June 1, 2010. Available at: http://astro.temple.edu/~lombard/reliability/.

Lang, G.E., & Lang, K. (1980). Polling on Watergate: The battle for public opinion. *Public Opinion Quarterly, 445*, 530-547.

Lariscy, R.W., Avery, E.J., Johnson, E., & Youngju, S. (2010). Health journalists and three levels of public information use: Issue and agenda disparities? *Journal of Public Relations Research, 22*, 2, 113-153.

Len-Rios, M.E., Hinnant, A., & Park, S. (2009).Understanding how health journalists judge public relations sources: A rules theory approach. *Public Relations Review, 35*, 56-65.

Lichter, S. R., Rothman, S., & Lichter, L. S. (1990). Group portrait. *The media elite: America's new powerbrokers* (pp. 20-63). New York, NY. Hastings House.

Mackert, M., Whitten, P., & Garcia, A. (2008). Interventions for low health literate audiences. *Journal of Computer-Mediated Communication, 13*, 504-515.

Marchi, R.M. (2008). Race and the news: Coverage of Martin Luther King Day and Dia de Los Muertos in two California dailies. *Journalism Studies, 9*(6), 925-944.

McAdams, D. P. (2008). The person: An introduction to the science of personality psychology. (pp.105-204.) Hoboken. NJ, John Wiley & Sons.

McClung, H.J., Murray, R.D., & Heitlinger, L.A. (1998). The Internet as a source for current patient information. *Pediatrics, 101*(6), e2, 1-4.

McCombs, M. E. (2004). *Setting the agenda: The mass media and public opinion*. Cambridge, UK. Polity Press.

McCombs, M.E., & Shaw, D.L. (1972) The agenda-setting function of the mass media. *Public Opinion Quarterly, 36*(2), 176-187.

McCombs, M., Llamas, J.P., Lopez-Escobar, E., & Rey, F. (1997) Candidate images in Spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly, 74*(4), 703-717.

McMillan, S.J. (2006). Exploring models of interactivity from multiple research traditions: Users, documents and systems. In (Eds.) L.A. Lievrouw & Livingstone, S., *Handbook of new media: Social shaping and social consequences of ICTs* (pp. 205-228), Thousand Oaks, CA, Sage Publications.

Meijer, M., & Kleinnijenhuis, J. (2006). Issue news and corporate reputation: Applying the theories of agenda setting and issue ownership in the field of business communication. *Journal of Communication,56*, 543-559.

Meissner, H.I., Ches, S., Potosky, A.L., & Convissor, R. (1992). How sources of health information relate to knowledge and use of cancer screening exams. *17*(3), 153-165.

Meraz, S. (2009) Is there an elite hold? Traditional media to social media agena setting influence in blog networks. *Journal of Computer-Mediated Communication, 14*, 682-707.

Metzgar, M.J., Flanagin, A.J., Eyal, K., Lemus, D.R., & McCann, R.M. (2003). Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. In (Ed.) P.J. Kalbfleisch, *Communication Yearbook 27* (pp. 293-336), Mahwah, NJ, Erlbaum Associates.

Mitchelstein, E., & Boczkowski, P.J. (2009) Between tradition and change: A review of recent research on online news production. *Journalism, 10*(5), 562-586.

Morahan-Martin, J.M. (2004). How Internet users find, evaluate and use online health information: A cross-cultural review. *CyberPsychology & Behavior*, *7*(5), 497-510.

National Newspaper Association. (2009). Newspaper Web site audience rises twelve percent in 2008. *NAA press release,* J. Sigmund.

Niederdeppe, J., Hornik, R., Kelly, B. J., Frosch, D., Romantan, A., Stevens, R.S., Barg, F.K., Weiner, J.L., & Schwartz, J.S. (2007). Examining dimensions of cancer-related information seeking and scanning behavior. *Health Communication, 22*(2), 1953-1967.

O'Connor, Amy. (2004). Reputation management. *Encyclopedia of Public Relations*. Sage Publications. Available at <http://www.sage-ereference.com.floyd.lib.umn.edu/publicrelations/Article_n372.html>.

Ogan, C.L. (2009). News audiences, decline of. *Encyclopedia of Journalism*. Sage Publications. Available at <http://www.sage-ereference.com.floyd.lib.umn.edu/journalism/Article_n260.html>.

Ohl, C.M., Pincus, J.D., Rimmer, T., & Harrison, D. (1995). Agenda building role of news releases in corporate takeovers. *Public Relations Review, 21*(2), 89-101.

Pan, Z., & Kosicki, G.M., (1993). Framing analysis: An approach to news discourse. *Political Communication, 10*, 55-75.

Pew Project for Excellence in Journalism (2010). The state of the news media: An annual report on American journalism. Washington, DC, Pew Research Center. Available at: http://www.stateofthemedia.org/2010/.

Pew Internet & American Life Project. (2005). *Health information online.* Pew Research Center. Washington, DC, S. Fox.

Pew Internet & American Life Project. (2008). *The engaged E-patient.* Pew Research Center. Washington, DC, S. Fox.

Pew Internet & American Life Project. (2009). *The social life of health information.* Pew Research Center. Washington, DC, S. Fox & S. Jones.

Protess, D.L., Cook, F.L., Curtin, T.R., Gordon, M.T., Leff, D.R., McCombs, M.E., & Miller, P. (1997). The impact of investigative reporting on public opinion and policymaking targeting toxic waste. *Public Opinion Quarterly, 51*, 166-185.

Quantcast (2009). Quantcast profiles [search sites]. Available at: http://www.quantcast.com.

Quantcast (2010). Quantcast measurement. Available at: http://www.quantcast.com/measurement.

Rains, S.A. (2007). Perceptions of traditional information sources and use of the World Wide Web to seek health information: Findings from the Health Information National Trends Survey. *Journal of Health Communication, 12*, 667-680.

Reese, S.D., Rutigliano, L., Hyun, K., & Jeong, J. (2007). Mapping the blogosphere: Professional and citizen-based media in the global news arena. *Journalism, 8*(3), 235-261.

Rice, R.E. (2006). Influences, usage, and outcomes of Internet health information searching: Multivariate results from the Pew surveys. *International Journal of Medical Informatics, 75*, 8-28.

Rieh, S.Y. (2004). On the Web at home: Information seeking and Web searching in the home environment. *Journal of the American Society for Information Science and Technology, 55*(8), 743-753.

Rodman, G. (2008). Mass media in a changing world, 2nd edition. Boston, MA, McGraw Hill.

Saad, L. (2010). Congress ranks last in confidence in institutions. Gallup's 2010 Confidence in Institutions Poll, July 22, 2010. Princeton, NJ, Gallup, Inc. Available at: http://www.gallup.com/poll/141512/Congress-Ranks-Last-Confidence-Institutions.aspx?version=print.

Scheufele, D.A. (2000). Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass Communication & Society, 3*(2), 297-316.

Scheufele, D.A., & Tewksbury, D. (2007). Framing, agenda setting, and priming: The evolution of three media effects models. *Journal of Communication, 57*, 9-20.

Scott, C.R. (2007). Communication and social identity theory: Existing and potential connections in organizational identification research. *Communication Studies, 58*(2), 123-138.

Seidman, J.J. (2006). The mysterious maze of the World Wide Web: How can we guide consumers to high-quality health information on the Internet? In (Eds.) M. Murero & R.E. Rice, *The internet and health care: theory, research, and practice,* (pp. 195-214.) Mahway, NJ, Erlbaum Associates.

Shah, D.V., & Scheufele, D.A. (2006). Explicating opinion leadership: Nonpolitical dispositions, information consumption, and civic participation. *Political Communication, 23*, 1-22.

Sheskin, D. (2004). *Handbook of parametric and nonparametric statistical procedures.* Boca Raton, FL, Chapman & Hall/CRC Press.

Shim, M., Kelly, B., & Hornik, R. (2006). Cancer information scanning and seeking behavior is associated with knowledge, lifestyle choices, and screening. *Journal of Health Communication, 11*, 157-172.

Silberg, W.M., Lundberg, G.D., Musacchio, R.A. (1997). Editorial: Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor – let the reader and viewer beware. *Journal of the American Medical Association, 277*(15), 1244-1245.

Slater, M.D., Hayes, A.F., Reineke, J.B., Long, M., & Bettinghaus, E.P. (2009). Newspaper coverage of cancer prevention: Multilevel evidence for knowledge-gap effects. *Journal of Communication, 59*, 514-533.

Southwell, B.G., & Yzer, M.C. (2007). The roles of interpersonal communication in mass media campaigns. *Communication Yearbook, 31*, 420-462.

Southwell, B.G. (2005). Between messages and people: A multilevel model of memory for television content. *Communication Research, 32*(1), 112-140.

Spaeth, M. (2009). Presidential politics and public relations in 2008: Marshall McLuhan 2.0. *Journalism Studies, 10*(3), 438-443.

Stromback, J., & Kiousis, S. (2010). A new look at agenda-setting effects – comparing the predictive power of overall political news consumption and specific news media consumption across different media channels and media types. *Journal of Communication, 60*(2010), 271-292.

Stryker, J.E., Moriarty, C.M., & Jensen, J.D. (2008). Effects of newspaper coverage on public knowledge about modifiable cancer risks. *Health Communication, 23*, 380-390.

Sweetser, K.D., & Metzgar, E. (2007). Communicating during crisis: Use of blogs as a relationship management tool. *Public Relations Review, 33*(2007), 340-342.

Sweetser, K.D., Golan, G.J., & Wanta, W. (2008). Intermedia agenda setting in television, advertising, and blogs during the 2004 election. *Mass Communication & Society, 11*, 197-216.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33*, 1-39.

Tan, Y., & Weaver, D.H. (2009). Local media, public opinion, and state legislative policies: Agenda setting at the state level. *The International Journal of Press/Politics, 14*(4), 454-476.

Tewksbury, D., & Althaus, S.L. (2000). Differences in knowledge acquisition among readers of the paper and online versions of a national newspaper. *Journalism and Mass Communication Quarterly, 77*(3), 457-479.

Tian, Y., & Robinson, J.D. (2008). Media use and health information seeking: An empirical test of complementarity theory. *Health Communication, 23*(2008), 184-190.

Tian, Y. (2009). Incidental health information use on the Internet. *Health Communication, 24*(2009), 41-49.

Toms, E.G., & Latter, C. (2007). How consumers search for health information. *Health Informatics Journal, 13*(3), 223-235.

Viswanath, K., Blake, K.D., Meissner, H.I., Saiontz, N.G., Mull, C., Freeman, C. S., Hesse, B. & Croyle, R.T. (2008). Occupational practices and the making of health news: A national survey of U.S. health and medical science journalists. *Journal of Health Communication, 13*, 759-777.

Viswanath, K., Breen, N., Meissner, H., Moser, R.P., Hesse, B., Steele, W.R., & Rakowski, W. (2006). Cancer knowledge and disparities in the information age. *Journal of Health Communication, 11*, 1-17.

Walterns, T.N., Walters, L.M., & Gray, R. (1996). Agenda building in the 1992 presidential campaign. *Public Relations Review, 22*(1), 9-24.

Wanta, W., & Elliott, W.R. (1995). Did the "Magic" work? Knowledge of HIV/AIDS and the knowledge gap hypothesis. *Journalism & Mass Communication Quarterly, Vol. 72*(2), 312-321.

West, D.M. (2001). *The rise and fall of the media establishment*. Boston, MA, Bedford/St. Martin's.

Whetten, D. (2009) Albert and Whetten revisited: Strengthening the concept of organizational identity. *Journal of Management Inquiry, 15*(3), 219-234.

Wilkes, M.S. (1997). The public dissemination of medical research: Problems and solutions. *Journal of Health Communication, 2*, 3-16.

Wright, K.B. (2009). Increasing computer-mediated social support. In (Eds.) J.C. Parker, & E.E. Thorson. *Health communication in the new media landscape* (pp. 243-266). New York, NY, Springer Publishing Company.

Yanovitzky, I., & Bennett, C. (1999). Media attention, institutional response, and health behavior change: The case of drunk driving, 1978-1996. *Communication Research, 26*(4), 429-453.

Yanovitzky, I., & Stryker, J. (2001). Mass media, social norms, and health promotion efforts: A longitudinal study of media effects on youth binge drinking. *Communication Research, 28*(2), 208-239.

Yzer, M.C., & Southwell, B.G. (2008). New communication technologies, old questions. *American Behavioral Scientist, 52*(1), 8-20.

Zaller, J. (1992). The nature and origins of mass opinion. Cambridge, University Press.

Zoch, L.M., & Molleda, J. (2006). Building a theoretical model of media relations using framing, information subsidies, and agenda-building. In C.H. Botan & V. Hazleton (Eds.) Public relations theory II (pp. 245-272). Mahway, NJ, Erlbaum Associates.

Appendix A: Survey

First pop-up page:

[Organization], along with a doctoral student from the University of Minnesota, is conducting research about people seeking health information on the Internet. ***The survey is estimated to take about 2 minutes***. The survey responses will be completely anonymous and used for research purposes only in congregate — as a group of answers, with no individual answers identified.

Yes, I want to take the survey

No thanks

Second page:

We thank you for participating in this research.
Please click on the "Next" button to begin the survey.

Next *Screen out non-US residents

Screen

In which region of the U.S. do you live?

- ☐ Northeast
- ☐ South
- ☐ Midwest
- ☐ West
- ☐ I do not live in the U.S.

*Non-US: Thanks and end survey

Question 1

How often do you …

A_1 **Search for health information on the Internet?**	Just about every day	About once each week	Once or twice each month	Rarely/ never
A_2 **Read health news on the Internet?**	Just about every day	About once each week	Once or twice each month	Rarely/ never
A_3 **Receive health information by email, such as e-newsletters or email updates you signed up for?**	Just about every day	About once each week	Once or twice each month	Rarely/ never

Question 2

When you want health news, or updated information about a particular health issue, which HEALTH web site do you use most often?

[open - comments]

Question 3

Which NEWS web site do you use most often for health news?

[open - comments]

Question 4

Thinking about the health issue most important to you, in the past six months, how did you usually receive news or updated information on that important health issue?

(Check any source of news you use often.)

- ☐ Your doctor, nurse or clinic
- ☐ Your insurance company
- ☐ Newspapers or magazines
- ☐ TV or radio
- ☐ A news Web site
- ☐ A health organization Web site
- ☐ Received an email from a family member or friend about the news

Question 5

In the past six months, did the information you got from the Internet change how you care for your own health or the health of others?

☐ **Yes**

☐ **No**

Question 6

In the past 6 months, what was the most typical health information you got from the Internet? (Choose one.)

- Important information on one health issue
- Information on several important health issues
- Information on many interesting health issues
- A large amount of "junk" information on health issues of no interest to you

Question 7

In the past six months, how often have you talked with family and friends about any health issues?	Daily	Weekly	Monthly	Rarely/ never

Question 8

Please indicate if you have shared health information you found on the Internet with any of the following people. (Choose all that apply.)

- Doctor
- Nurse
- Family (spouse, children, parents)
- Friends
- Co-workers

Question 9

In the past six months, how many times have you emailed articles about health to people you know?	More than 10 times	About 5 to 10 times	About 1 to 4 times	Never

Question 10

Why did you email a health article (or articles) to someone within the past year? (Choose all that apply.)

- ☐ You sent the article because you knew the person would be interested in it.
- ☐ You intended to talk with the person later about the information you emailed.
- ☐ The person asked you to find information on the Internet about that health issue.
- ☐ You had talked in the past with this person about a health issue.
- ☐ You were trying to persuade the person to take the action described in the health article.
- ☐ You thought the information would be useful for the person because he or she is taking care of someone with a health problem.

Question 11

Which of these options did you use to send an email to someone about health information you found on the Internet? (Choose any or all that apply.)

- ☐ Clicking on the "email article" option.
- ☐ Copying and pasting the URL (Web site address) in your email message.
- ☐ Copying the entire text of the article and pasting it directly into your email.
- ☐ Copying the entire text and creating a document to email..

Question 12

In the past six months, how many times have you received emails of health articles from people that you know?	More than 10 times	About 5 to 10 times	About 1 to 4 times	Never

Question 13

How often do you use social networking Web sites (such as Facebook or MySpace)	Daily	Weekly	Monthly	Rarely/ never

Question 14

Have you used a social networking Web site to share health information?

☐ **Yes**

☐ **No**

Question 15

For news on health, to what extent do you prefer these types of health stories?

Health news stories featuring …	**I prefer this type of story …**			
Doctors or other medical experts in the news who give advice for caring for your health (for example, Dr. Oz on TV and radio)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
News from a university or other research institute about new research on causes or cures for illness or health conditions (for example, a University professor's research)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
Personal stories from individuals about their experience with health issues and their personal advice on caring for your health (for example, a cancer patient's personal web site)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
Government public health agencies' news about a contagious illness or food contamination spreading around the country (for example, the FDA or CDC)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
News from nonprofit organizations about promotions to fight specific diseases or illnesses (for example, American Heart Association "Wear Red" to promote women's heart health)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
News from government public health agencies to promote ways to help people with their health (for example, a state web site on flu shots or help quitting to smoke)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
News from commercial organizations or for-profit businesses (for example, a new health product on WebMD.com)	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much

Question 16

To what extent do you prefer these types of sources on the Internet that provide health information?

Type of health source	**I prefer this type of health source …**			
Government Web sites (ending with **.gov**) For example, the Centers for Disease Control, the NIH, the FDA, or other public health agencies	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
Commercial Web sites (ending with **.com**) For example, WebMD.com or Drugs.com or Web sites of clinics, hospitals or insurance companies	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
Nonprofit or association Web sites (ending with **.org**) For example, FamilyDoctor.org (American Academy of Family Physicians Web site) or the American Lung Association's Web site	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much
Personal Web sites or blogs by individuals or groups who share the same medical problem (these may end with **.com** or **.org** or **.net**) For example, a Web site where people with arthritis share their experience with the disease and treatments they have had, or a Web site of a person who is trying to lose weight	Do not prefer at all	Prefer somewhat	Prefer a good amount	Prefer very much

Question 17

In order for you to receive important health news or updates on health issues, which one of these sources would you most prefer from the Internet? (Choose only one.)

- ☐ The Web site for your doctor's office or clinic
- ☐ A nonprofit health association or organization's Web site
- ☐ A government public health agency's Web site
- ☐ The Web site of a commercial business that specializes in health information
- ☐ A Web site for a university research center for health
- ☐ Web sites by individuals or groups who have experience with a particular health issue

Question 18

How would you rate your own personal health?

- ☐ Excellent
- ☐ Good
- ☐ Fair
- ☐ Poor

Question 19

Your gender:

- ☐ Female
- ☐ Male

Question 20

Your race

- ☐ White alone
- ☐ Black or African American alone
- ☐ Asian alone
- ☐ Some other race alone
- ☐ Two or more races

Question 21

In what year were you born? _____

Question 22

About how many years have you been using the Internet? _____

Question 23

Your education

- ☐ Less than high school
- ☐ High school graduate
- ☐ Some college
- ☐ College graduate (bachelor's or associates degree)
- ☐ Graduate school

Appendix B: Coding summary

News articles were coded as follows:

1. Articles were categorized on most-viewed or most-emailed list, along with rank order on the list.
2. News was categorized as health news or nonhealth news. Health news was broadly defined to answer a question: Does the reader gain health information from the story? Health information would leave some kind of impression on readers on an aspect of health in the article. Another consideration was whether the reader would form an attitude or opinion about the person in the story based on health information provided. If so, the story was coded as health news. In nonhealth news, the reader did not gain health information from the story. If the reader would be influenced by the main topic of the story, which was not health related, the article was coded as nonhealth news.
3. Health news was categorized into sub-codes, as follows:

- Business and health, including health activities by businesses, employee health benefits issues, medical insurance
- Crime and legal issues dealing with health. Crime was categorized as health news if there were health details in the story dealing with the victims or the perpetrator.
- Crisis and health involved public health issues, such as contaminated food, flu epidemic issues, or other community-based health crises.
- Education and health encompassed all levels of education (pre-school through college) and included school-based efforts to improve health, community programs for health education, schools or students involved in health events or efforts.

- Entertainment and health: Included celebrity health issues, such as a celebrity in a hospital or receiving medical treatment, celebrity deaths only if health details are provided on cause of death; celebrity pregnancy or birth only if health details are provided.
- Mental health/psychology: Included personal advice columnists, articles dealing with human emotional behavior, such as stress reduction
- Physical health/illness included health issues affecting the human body
- Prevention health: Included exercise, wellness efforts, immunization, nutritional values
- Policy/ politics and health included laws or public policies dealing with health issues, such as health insurance, discussions on health reform, advocacy for health improvements
- Sports and health: Included amateur sports if athletes play for fitness; pro sports when athlete's injury involves medical details; kids' sports when kids' health is detailed
- Technology and health: Included new medical treatments involving new types of treatment, research on new devices or equipment, new drugs, scientific breakthroughs in health
- Travel and health

Nonhealth news was also subcoded based on topic. Because the subcodes were not used in analysis, they are not included here.

4. Originating sources within news were defined as sources providing information to the reporter. The coders were specific in attempting to identify originating sources by

attribution, and to not simply coding any individual listed in the story. The attribution-based examination was intended to examine the stories to determine which sources were actually interviewed or provided information for the news report. For example, a criminal who was charged with a crime was not coded as an individual if the story came from a police department. But if the criminal was quoted in the story, that was coded as an individual source. A health news example would involve a physician discussing treatment for a patient. If the physician referenced the patient, only the physician was coded as a source. If the patient was cited as providing information to the reporter, or was quoted in the story, then the patient was coded as a source.

5. Sources within a story were identified and coded only one time, regardless of the number of times the source was quoted or cited. If the organization was listed as a source, and then a spokesperson from the organization was quoted, the organization and individual were coded as one source. If an organization had more than one spokesperson identified as a source in the story, the additional spokesperson(s) was also coded.
6. Originating sources with news were first categorized as personal or institutional. The title and descriptions of the person provided the differentiator. A person identified by title as affiliated with an organization was coded as institutional source.

Individual/ personal sources were coded as follows:

- Individuals involved in the story for personal reasons. These were individuals who were not identified by title or organizational affiliation.
- Persons on the street interviewed incidentally by a reporter.

- Individuals randomly asked for their opinions, who had no apparent organizational affiliation or expertise, were coded as individual.
- Individuals in the workplace who were not spokespersons representing their employer, who were incidental to the story, were coded as individuals. For example, a waitress at a restaurant who is quoted about an incident at the restaurant would be coded as individual.
- People identified by where they work, but cited in the story based on their personal involvement (not work-related involvement) were coded as individual sources.
- Citizen journalists, such as CNN's "iReporter"
- Criminals who were not celebrities, and their family and friends cited in the story.
- Readers who post questions in interactive blogs

7. Institutional sources were further categorized within socioeconomic identities, based on their social and economic function, as follows:

Public sources:

- Public officials are individuals affiliated with government agencies, elected officials.
- Candidates running for public office
- Experts who were former elected or administrative officials, but are currently identified in news media as affiliated with a business, institute or university, were coded by their current role. If the former official is identified only in their former role, they were coded as public. For example, a former president or former secretary of state.
- International agencies such as World Health Organization, International Monetary Fund.

- Court hearings, court documents, court evidence
- Judges, prosecutors, public defenders
- Legislative hearings
- Public schools, public school administrators, teachers acting as spokespersons on behalf of a school

Private sources:

Sources were coded as private if they had commercial interests or a private agenda, in that their intentions are not directed toward the public.

- Celebrities and individuals from the entertainment industry, and the elebrity's friends and family
- Book authors who were not affiliated with a public, nonprofit or academic organization
- Organizations involved with sales of a product or service, unless affiliated with a public, nonprofit or academic organization
- Experts identified by their profession, unless affiliated with a public, nonprofit or academic organization, such as a physician, attorney, airline expert, financial advisor, etc.
- Advice columnists who were not affiliated with a public, nonprofit or academic organization
- Polls/ polling are private sources unless the poll sponsors are specifically listed as public or nonprofit organizations
- Twitter and Facebook were categorized as private-sector organizations; if a person sending the message was an individual without organizational affiliation, then the

source was coded as individual. If the news story only cited the information as coming from a Twitter or Facebook account, the source was coded as private.

Nonprofit sources

Nonprofits encompassed charity organizations and advocacy organizations as follows:

- Political parties were coded as advocacy, except when the source was an elected public official, who would then be coded as a public source.
- Associations of individuals or organizations were coded as advocacy, including industry associations, such as chambers of commerce, and trade associations.
- Foundations
- Institutes or think tanks, unless they were specifically referenced as affiliated with a public or academic organization
- Religious organizations
- Unions
- Grassroots political organizations and lobbying groups

Academic sources

- Academic sources were broadly defined as any source affiliated with a university or college, research institute or academic/scientific journal. Academic sources included:
- Academic/scientific journals such as JAMA or Nature
- Professors or university faculty or staff
- Students identified by university affiliation
- University administrators
- University sports organizations, or sports administrators

8. News media who were cited as sources in the story as intermediaries to the story, attributed as "according to …" or "as told to…" were coded as news media (not individuals or originating institutional sources). If a news organization or individual affiliated with the news organization was an active part of a news story, or reporting an opinion about a news story, the individual was coded as a private source.
9. Anonymous organizational sources were coded by organizational source only if their organizational affiliation was identified. Anonymous individual sources, without organizational affiliation, were not coded.
10. Columnists themselves were coded as sources if they were actively part of their commentary, or part of the events in their commentary, or who referenced themselves in the editorial as a source such as "I think" or "we think." Columnists were coded by organizational identification as public, private, nonprofit or academic. Columnists without an identified organizational affiliation were coded as private sector. If the commentary referenced other sector sources, and the columnist was an employee of the news media, or not identified by organizational affiliation, the sources cited in the commentary were coded, but the columnist was not coded. Blog entries are categorized as columnists