

# HOW QUESTION WORDING IMPACTS NEWSPAPERS ESTIMATES: TRADITIONAL vs. ONLINE

**Gregg Lindner and Jane Traub, Scarborough Research**

---

## Background

The ability of newspaper reading to take on forms other than the traditional printed version of the paper has added a level of complexity to the measurement of newspaper readership. Historically, newspaper readership has been defined in most readership studies as “reading or looking into” the printed, or “hard-copy” version of the newspaper. With the launch of Internet versions of newspaper vehicles, the issue of what defines readership has been made less clear. The majority of newspapers, both local and national, now publish online versions of their daily media vehicles. Currently, almost 1,500 daily newspapers in North America publish an Internet based version of their paper.<sup>i</sup>

Online newspaper versions are now commonplace and they come in a variety of formats. The convergence of the content of these papers with their printed counterparts varies widely. Some newspapers have nearly identical content on their web site and printed page, while others have radically different content in each version. The way the content is organized and presented varies as well. Some newspaper websites look fairly similar to their print counterparts. Other papers appear quite different from their printed versions. The ubiquitous nature of online papers and their similarity or dissimilarity to their print counterparts may contribute to confusion among readers when asked to report their readership of a specific newspaper media vehicle. Should they report both their print and online consumption? Print only? Is the respondent capable of considering each readership experience separately, or have they blended and merged in the reader’s consciousness?

Previous research suggests that the main reason people read a specific online newspaper is that they are readers of the print version of that newspaper.<sup>ii</sup> One recent survey found that 50% of online newspaper readers said they not only read but subscribe to a printed newspaper.<sup>iii</sup> This may result in uncertainty among those who read both the print and online versions of a specific paper, when trying to recall during readership research whether or not they have read an issue of the newspaper. The time period that readership studies use for screening, typically the last week for daily newspapers or last four weeks for Sunday newspapers, may also cloud the picture. Did they read the printed paper on Wednesday or did they read it online on Tuesday? Certainly the opportunity for respondent confusion exists.

Indeed, the very concept of a Daily versus Sunday newspaper becomes muddled when considering the online version of a newspaper. Before Internet newspapers were made available, the traditional Sunday newspaper was available only on Sunday (or perhaps Saturday, in major cities where local publishers print early copies). The experience of purchasing or receiving the Sunday newspaper was once totally unique from the daily newspaper. The Sunday version was historically large and heavy or bulky, with special sections that only appeared weekly. It contained a magazine and printed advertising inserts. It always looked and felt different from the daily paper.

Contrast this experience with the reader of the “Sunday” version of an Internet newspaper. It often looks nearly identical to the daily version. Some or all of the content may continue to be available on the newspaper’s website until the next Sunday. This experience is quite different for the reader than the once-unique aspect of the traditional Sunday paper, and provides an example of how the Internet has changed the readership experience for the consumer of the paper.

The questions of respondent confusion with regard to print and Internet newspaper reading are relevant to Scarborough Research. We are the primary provider of newspaper audience data to the U.S. market. Scarborough’s measurement of printed newspaper audiences is based on a series of questions asked during the telephone interview phase of our syndicated survey. The questions follow the ARF published guidelines for newspaper audience data collection.

Several industry leaders raised the concern that, in responding to these readership measurement questions, some respondents may believe that the question seeks their behavior with respect to virtual (electronic) Internet-published newspapers as well as printed (hard-copy) newspapers, or that respondents may be confusing these two formats. The Scarborough newspaper readership questions are intended, of course, to refer only to traditional hard-copy printed newspapers.

To explore the question of the impact of explicitly excluding Internet newspapers during measurement of traditional Print newspaper readership, a quantitative field test was executed. The fieldwork was conducted by Scarborough calling centers from October 10<sup>th</sup> – December 1<sup>st</sup>, 2002. Three alternate question wordings, intended to clarify the necessary distinction between newspaper formats, were added to the standard Reader Profile readership question. The objective of the test was to determine the impact each of these question variations had on newspaper readership levels.

Prior to fielding the quantitative test, a separate, *qualitative* test was conducted. The purpose of this test was to identify the descriptive term or terms that consumers commonly use in distinguishing between physical printed newspapers and electronic Internet newspapers. Consumers might use various language to describe the electronic medium (Internet newspapers) that we were attempting to exclude from the readership measurement. Moreover, it was important that wording be identified that worked for experienced Internet users, non-users, and everyone in between.

The study design for the qualitative test was a random digit dial telephone sample of respondents across the U.S. A random respondent was selected within each household contacted, using the “most recent birthday” method. The interview began with a description of the survey topic: newspaper reading. The interviewer asked the respondent to describe in their own words how they would define electronic newspapers, without specifically naming them in the question, to avoid bias. Verbatim responses were captured and tabulated. 314 interviews were completed in September, 2002. The survey found that “Internet newspapers” and “Online newspapers” were the most commonly used terms to describe the medium in question. The interview questions are shown below:

1. “Today’s survey is about reading newspapers. Newspapers are now available in many formats. Do you currently read any type of newspapers?”
2. “Besides the printed newspaper, have you ever read or looked into any other kinds of newspaper?”
3. “What kind or type would that be?”
4. “Now, I’d like to ask you about your readership of traditional printed newspapers. First I’d like to know about your readership of weekday newspapers, that is, newspapers published every day Monday through Friday. In the last seven days, have you read or looked at any printed Daily newspaper? “

The remaining questions captured Sunday newspaper readership, Internet usage and demographics. The frequency distribution of answers to question 3 above is shown in Table 1 below:

**Table1: Verbatim Responses to Alternative Newspaper Sources**

<b>Coded Responses</b>	<b>Percent Answering (n=314)</b>
Internet/The Internet/On the Internet	48%
Online/From Online	20%
Other*	15%
Computer/Home Computer/On the Computer	11%
Web/The Web/Website/On the Web	5%
Electronic/Electronically	1%

\*The “other” category consisted mostly of mentions of specific websites.

Once the language most commonly used by respondents to describe Internet newspapers was identified, the questions for the quantitative test were then slightly modified to reflect the data. These revised questions were used in the quantitative test.

The 3 versions of the modified Scarborough newspaper readership question used in the quantitative test are shown below. The bold text indicates the additional language used to modify the current question wording.

**1. Version 1:**

Question 2a: First, I’d like to ask some questions about your readership of weekday newspapers; that is, newspapers published every Monday through Friday. **I’M NOT TALKING ABOUT COPIES OF THE NEWSPAPER WHICH CAN BE RECEIVED VIA COMPUTER USING THE INTERNET. I’LL ASK ABOUT POSSIBLE INTERNET USAGE LATER IN THE INTERVIEW.** During the past seven days, which of the following Monday through Friday newspapers, if any, have you read or looked into either at home or away from home?

**2. Version 2:**

Question 2a: First, I’d like to ask some questions about your readership of weekday newspapers; that is, newspapers published every Monday through Friday. During the past seven days, which of the following Monday through Friday newspapers, if any, have you read or looked into either at home or away from home? **PLEASE DO NOT INCLUDE USE OF ANY NEWSPAPERS ON THE INTERNET OR ONLINE . I AM ONLY INTERESTED IN THE PRINT VERSION OF NEWSPAPERS AT THIS TIME.**”

**3. Version 3:**

Question 2a: First, I'd like to ask some questions about your readership of weekday newspapers; that is, newspapers published every Monday through Friday. During the past seven days, which of the following Monday through Friday newspapers, if any, have you read or looked into either at home or away from home? AS I READ THE LIST, PLEASE INCLUDE ONLY READING OF THE PRINTED PAPER; DO NOT INCLUDE READING THE PAPER ON THE INTERNET OR ONLINE.

For purposes of the test, three "virtual markets" were created, each consisting of sample from nine different and diverse Scarborough markets. Each of these "markets" corresponded to a test cell. Setting up the test cells in this fashion ensured that a wide variety of markets were represented in each test cell. For the test, each virtual market had one of the three alternate versions of the newspaper measurement question assigned to it, and every respondent in the "market" was read that version of the question. The markets selected for inclusion were: Albany, Chicago, Dallas, Denver, Des Moines, Hartford, Minneapolis, Norfolk and Seattle.

**Results**

The following table summarizes the newspaper readership levels reported using the three different versions of the readership question:

	Version 1 (n=217)	Version 2 (n=216)	Version 3 (n=221)
Read any daily newspaper "yesterday"	55.3%	53.2%	54.8%
Read any Sunday paper "last Sunday"	71.0%	68.1%	68.3%

None of the differences among the versions are statistically significant, whether comparing any one cell to the others combined, or else comparing any combination of two cells.

**Comparison to syndicated data**

Scarborough conducted a comparison to syndicated data as a control group, to further examine the effects of the modified question wording on Internet readership. The analysis compared the test data to the syndicated interviews conducted in these nine markets during the same field period as the test: 10/22/02 to 12/1/02. This comparison is shown below.

Comparison to Syndicated Data	Combination, 3 test versions of Internet Readership questions (n=654)	Scarborough Syndicated Interviews, 10/22-12/1, 2002* (n=2,816)	Difference
Read any daily newspaper "yesterday"	54.4%	53.6%	0.8
Read any Sunday paper "last Sunday"	69.1%	64.6	4.5#

\*Interviews during same field period in same nine markets

# Statistically significant at 95% confidence level

The syndicated readership numbers from the *full* Scarborough data collection period for the same nine markets were very close to, and not significantly different from, the numbers for the test period (53.6 daily, 64.6 Sunday). This suggests there was nothing unusual about the field period during which the test was conducted.

The average readership number from the tested wordings is not significantly different for Daily readership, ("read yesterday), but does show a significant difference for Sunday readership ("read last Sunday"). So it does seem that the wording modifications mentioning Internet readership which were used in this test are associated with about a 4.5 percentage-point rise in last Sunday readership.

This finding is somewhat counter-intuitive, since one would assume that telling the respondent *not* to include Internet reading in their reported newspaper readership would result in a *decrease* in overall readership, rather than an increase.

This finding prompted some additional analysis of the data. When comparing the three test cells individually to the syndicated data for the same time frame and markets, the data show that only version 1 generates Sunday readership numbers that are

significantly different from Scarborough syndicated. However, Versions 2 and 3, taken individually, do not. The following table summarizes Daily and Last Sunday newspaper readership for each of the tested question wordings, and the same data for syndicated interviews conducted during the test period:

	Version 1 Test (n=217)	Version 2 Test (n=216)	Version 3 Test (n=221)	Any Test Wording (n=654)	Scarborough Syndicated*
Any daily newspaper “yesterday”	55.3%	53.2%	54.8%	54.4%	53.6%
Any Sunday paper “last Sunday”	71.0%#	68.1%	68.3%	69.1%#	64.6%

# Difference from Syndicated significant at 95% confidence

\*Test period only (10/22-12/1), for 9 test markets

**Discussion**

The theory that neglecting to tell respondents to exclude their Internet readership activity when reporting newspaper readership artificially inflates the readership numbers is not supported by the data. Were this the case, we would expect the clarifying language to be associated with a significant decrease in readership, not a significant increase. We would also expect this decrease to appear in both Daily and Sunday readership. Instead the findings show that the use of the modified language is associated with a statistically significant increase in Sunday readership. While counterintuitive, this increase appears to be related to the modification of the question, since all other factors were held constant.

Why might this significant increase in Sunday readership have occurred? One hypothesis is that, by mentioning Internet readership, respondents may have been reminded to include their Internet readership experience in their reported reading, despite being specifically told not to include it. Respondents may have chosen to ignore the instruction and reported all their reading, including Internet reading. Why might this change have been seen only in Sunday readership? Perhaps due to the higher incidence of Sunday readers, a larger pool of respondents was affected by the question and chose to include their Internet readership, thus resulting in a significant difference for the Sunday audience numbers. More work is needed to explore the reasons for the observed result.

The findings of the data continue to show the importance of testing small changes in wording and that even the smallest of wording changes can have a substantial impact on audience estimates. Given the expectations that Internet-published newspapers will continue to grow in a variety of manners, this finding indicates that attention must be given to even apparently simple, “common-sense” changes in question wording.

As a leading supplier of ratings data to the newspaper industry, Scarborough views this work as a continuing priority. Testing changes in wording that affect the newspaper ratings is one of our ongoing commitments. As media evolve, we must remain open to making changes in measurement when necessary. Simultaneously, we must test wording changes to avoid introducing unanticipated bias. We expect to continue the work in this area with additional testing.

~~~~~

<sup>i</sup> Source: NAA Facts about Newspapers 2003, A Statistical Summary of the Newspaper Industry, p.20  
<sup>ii</sup> Sources: Hansen, J., “Print Media Internet Portals: A Boon or Bust for Print Media Consumption?,” Worldwide Readership Symposium 2001, p. 385, and Lindner, G & Mulligan-Traub, J. “The Internet as a Medium”, Worldwide Readership Symposium 2001, page 407  
<sup>iii</sup> Source: 2002 NAA Digital Edge Report: Power Users: a Profile of Online Newspaper Consumers, Volume 1, #3, May 2002