

THE INTERVIEWER EFFECT ON READERSHIP LEVELS

**Gregg Lindner, Martin Frankel Ph.D., Simmons Market Research Bureau
Dan Mallett, Daniel Mallett Associates**

Executive Summary

The personal interview in traditional large scale print audience studies takes about an hour. Only about 20 minutes is actually devoted to print readership. Key demographics add a few more minutes, and the rest is used to cover other areas, principally broadcast media.

The print audience user community has been concerned about the effect this length interview may have on respondents and interviewers. Beyond the obvious potential for reducing overall research quality, a long and intrusive interview, in and of itself, may affect the key readership measures that are the primary study purpose.

Before launching a new print audience measurement service based on recent reading, Simmons decided to test a "pure" recent reading interview, covering print and limited demographics in about half an hour, against the "traditional" recent reading interview that lasts about an hour.

The test design included matched national samples of about 1,100 interviews each for the "pure" and "traditional" approaches, using the best interviewers and training available. The test cost more than \$250,000. To our knowledge, it is the largest and most rigorous test of its kind ever conducted in the United States.

Main test results are:

- Readership levels are about 13% higher for the shorter, "pure" interview than for the "traditional" interview.
- Audience gains in the shorter, "pure" interview are greatest in the demographic sub-groups (younger, employed, better educated) that are traditionally hardest to interview.
- Audience turnover is about 20% higher in the shorter, "pure" interview, resulting in higher multi-issue cume audiences.

In short, interview length does matter. In addition to the general quality-related benefits of reducing interviewer and respondent burden, a shorter, print-focussed interview produces different, more realistic audience measures, because it is better able to capture readership of those demographic groups that are both hardest to interview and the heaviest readers.

Based on these results, Simmons has introduced the Survey of American Readership, a survey dedicated to print audience measurement.

Design

Test Objective

Large scale syndicated print audience surveys in the United States have traditionally covered much more than print. In addition to newspapers and 200+ consumer magazines, the traditional recent reading interview also covers trade publications, radio, broadcast and cable TV, and a host of detailed and often very personal demographic areas. The interview takes about an hour to complete. Less than half this time is devoted to print.

After the personal interview, respondents are asked to complete a self-administered "product booklet", covering purchase, ownership, and use of literally hundreds of products and services. Completing the product booklet takes several hours.

For obvious reasons, the magazine research community has long been concerned about the effect this burden may have on overall research quality and on reported levels of print readership.

As Simmons began to seriously consider changing to the recent reading method for magazine audience measurement, a key issue was whether it might be possible to introduce improvements over current practice. Reducing interview length was a prime candidate.

Shorter interviews are easier to administer and offer immediate quality benefits in a number of areas.

Shorter interviews produce higher response rates, especially among busy, hard-to-interview population sub-groups. Shorter interviews reduce staffing requirements, providing the opportunity to use fewer, better interviewers and to focus on more intensive training and field supervision. And, of course, shorter interviews reduce incentives for interviewer short-cuts.

Beyond these well-known benefits, a shorter interview, in and of itself, may lead to different respondent answers.

A growing body of literature has shown that respondent answers to questions at the beginning of an interview can be affected by the number and content of questions that follow. While the respondent does not know what questioning will come, the interviewer does.

Irrespective of an interviewer's best intentions and training, this knowledge can subtly affect how the interview is conducted. In short, answers to the print audience questions asked at the beginning of the interview might be affected by the number and intrusiveness of the questions that follow.

The Simmons test was specifically designed to measure this effect. Under the best circumstances -- i.e., conscientious, motivated and well-trained interviewers -- does sheer length of interview matter?

"Pure" and "Traditional" Interview Protocols

Two interview protocols were used -- "pure" and "traditional".

Print questions were asked first in both, in the same order, with the same materials and question wording.

Daily newspapers were first, followed by Sunday newspapers, and then consumer magazines. Within each of these sections, question order was a) six-month screen, b) frequency of reading, c) publishing interval-specific reading.

The magazine section covered 215 publications. Screening and readership questions used black & white logo card decks and three position sort boards. Logo cards were shuffled before being presented to the respondent for screening. All readership questions were asked in a single weekly-to-bimonthly order (i.e., one questionnaire rotation, rather than the four used in actual surveys).

After newspapers and magazines, the "traditional" interview continued with the usual questioning on other media and demographics, and placement of the leave-behind product booklet.

The "pure" interview followed the same order, although with questions only about print media and demographics, and with no product booklet at the end.

The main "traditional" areas eliminated from the "pure" questionnaire dealt with other than consumer print media:

- Readership of 24 trade publications
- Radio listening by day-part, station, weekday and weekend
- Broadcast television viewing by day-part, station, weekday and weekend
- Cable television viewing by programming service and day-part

Other non-essential and unnecessarily intrusive questioning eliminated from the "pure" interview included:

- Elimination of 4 magazine reading qualitative questions (keeping 5)
- Elimination of most demographic questions related to the household head (as opposed to the respondent)
- Elimination of intrusive questioning (e.g., first and last names of all adults and children in the household)

Copies of both questionnaires are available on request.

Test Design

Because the test objective was to isolate the effect of interview length, design effort focussed on otherwise equal and best effort treatment in both cells.

Only the best 36 Simmons interviewers were used for the study. They were divided into two groups of 18 each, with one group assigned to the "traditional" test cell and the other to the "pure" test cell.

The groups were kept separate, and neither group was informed of the other's existence. Interviewer training was conducted jointly by Simmons' Technical Director, Gregg Lindner, and Simmons' Statistical Consultant, Dr. Martin Frankel. Material distribution and on-going interviewer contact were handled directly from Simmons' offices (without use of regional field managers).

It should be noted that many interviewers work for more than one print audience measurement service. These interviewers were equally divided between the two test cells, and were given the same training as other interviewers.

A total of 120 field locations was selected from Simmons' standard national sample, with each divided into two sub-sections.

Interviewing was conducted in two waves.

In Wave I, interviewers in the "traditional" cell were assigned the first sub-section of 60 field locations, while interviewers in the "pure" cell were assigned the first sub-section of the other 60 field locations. Wave II field assignments were reversed. Across the two waves, the "traditional" and "pure" protocols were both assigned to all 120 field locations.

Random selection rules were used in all selections and assignments described above. Standard Simmons field rules were used in selection of specific households and respondents.

Work on test design and field preparation began in earnest in early 1994. Interviewing was conducted over a four week period in July and August of 1994. Wave I ran from July 9 to July 23, Wave II from July 25 to August 6.

Interviewers were assigned 10 completed interviews in each field location for each wave. After final editing and validation, each test cell contained about 1100 completed interviews (1095 in the "traditional" cell, 1060 in the "pure" cell).

Costs for field work, material preparation, interviewer training, and travel were approximately \$250,000. To our knowledge, this is the largest and most rigorous test of its kind ever conducted in the United States.

Weighting/Projection

The "pure" and "traditional" samples were separately weighted and projected to U.S. total adult population, in two steps, using the same procedures for each.

The first step compensated for differential selection rates in the sample design (allocation by geographic strata, and selection of individual respondents within households). The second step balanced/projected each sample to U.S. totals on sex, age, and education.

Note that the test design also provides meaningful and useful comparisons on an unweighted basis -- the "pure" and "traditional" cells shared the same sample allocation, the same response rate, and so on.

For this reason, detailed test results shown in the Appendices are provided on both unweighted and weighted/projected bases. Many of the tables in the body of this report also provide unweighted and weighted/projected results. Where this is the case, discussion of results refers to weighted/projected data.

Results

Audience Levels

Table 1 below shows the average number of magazine titles per respondent screened and read in the "pure" and "traditional" interviews.

Table 1: Overall Screens and Reads

	Pure	Traditional	Pure vs. Traditional Diff	Traditional Index
Unweighted				
Average Titles Screened	15.93	13.15	+2.78	121
Average Titles Read	7.21	6.31	+ .90	114
Weighted/Projected				
Average Titles Screened	15.72	13.57	+2.15	116
Average Titles Read	7.30	6.45	+ .85	113

Across the 215 magazine titles measured, the average respondent in the "pure" interview sample screened in on about 16 titles, and read about 7. Respondents in the "traditional" interview sample averaged about 2 fewer screen-ins and 1 fewer read. These differences are statistically significant at the .05 level.

This translates into 16% higher screen and 13% higher readership levels for the "pure" interview.

Tables in the Appendix provide similar indices for various demographic sub-groups and publication types. Readership indices for some demographic groups are shown in Table 2 on the following page, on a weighted/projected basis.

Table 2: "Pure" vs. "Traditional" Readership Indices

	Pure vs. Traditional Readership Index
Total Adults	113**
Men	110
Women	116**
Age 18-34	122**
35-54	121**
55+	87
Att/Grad College	116*
Grad High School	119**
Less Than H.S. Grad	96
IEI \$50K+	127*
\$20K-\$50K	120**
\$20K-	122*
Not Employed	100

*Different from 100 at .10 significance level

**Different from 100 at .05 significance level

While "pure" vs. "traditional" readership differences were about the same for men and women, there are sharp patterns for other demographic groups. The greatest differences occur among the groups of most interest to many print media -- younger, better educated, and more affluent and more mobile sub-groups.

These are the sub-groups that are hardest to contact and find at home, who have the least time to be interviewed, and, as it happens, who often are heavy magazine readers.

It is these kinds of respondents whose readership answers are most positively affected by a shorter, easier to conduct interview.

Audience Turnover

Two-issue cumulative audiences in single interview surveys are calculated by tabulating average readership probabilities within each frequency group (reads/screens), using the binomial to extend to two issues within each frequency group, and adding resultant two-issue audiences across the groups.

While almost any result is theoretically possible, two-issue audiences generally depend on 1) the overall read-to-screen ratio, with lower read-to-screen ratios leading to higher turnover, and 2) the screen-in distribution across the frequency groups, with higher proportions of respondents in the lower frequency groups leading to higher turnover.

Table 1 in the previous section showed that the "pure" interview produced larger gains in screens than reads, and, hence, a lower read-to-screen ratio.

As shown in Table 3 below, the "pure" interview also has a higher proportion of screen-ins falling in the lower frequency groups.

Table 3: Frequency of Reading

	Unweighted		Weighted	
	Pure	Traditional	Pure	Traditional
Total Screen-ins	100.0%	100.0%	100.0%	100.0%
Frequency:				
Less than 1 of 4	18.8%	13.6%	18.1%	14.1%
1 of 4	36.8%	33.4%	36.8%	34.1%
2 of 4	17.4%	21.1%	18.0%	21.3%
3 of 4	5.4%	7.1%	5.7%	6.9%
4 of 4	21.6%	24.8%	21.4%	23.5%

For both reasons -- a lower overall read-to-screen ratio and a greater proportion of screen-ins in lower frequency groups -- the "pure" interview would be expected to lead to higher turnover. As shown in Table 4 below, this is indeed the case.

Table 4: Audience Turnover

	Pure	Traditional	Pure vs. Traditional	
			Diff	Index
Unweighted	38.4%	30.6%	7.8%	125
Weighted/Projected	38.1%	31.7%	6.4%	120

The "pure" interview produces turnover rates about 20% higher than the "traditional" interview. As a matter of interest, the "pure" interview turnover rates are at about the same level usually found in two interview, through-the-book studies.

Interviewer Effect

The test shows very clearly that a shorter, back-to-basics interview focussed on print produces different magazine audience measures than the traditional longer, multi-purpose recent reading interview. An obvious question is, "Why?"

Personal interviews are predominantly conducted in respondents' homes, and must be fit in with respondents' more important real world concerns (e.g., dinner's on the stove, someone has to be taken to a dentist's appointment, the kids are fighting in the background). Interviewers can't change these conditions; they simply work around them as best they can.

Interviewers are instructed to conduct the interview at an even, deliberate pace, with the respondent's full attention. If the respondent's attention wanders, the questions are supposed to be repeated. If the respondent is going too quickly (e.g., when sorting logo cards at the screening stage), the interviewer is supposed to stop him or her, and start again.

These are judgment calls and, despite best training and intentions, are influenced by how long the interview takes. Interviewers must strike a balance between maintaining respondent cooperation and conducting the interview as ideally designed. For respondents with limited time these goals are in conflict, and it is inevitable that a longer interview (with the added burden of asking the respondent to complete the product booklet) will be conducted at a faster pace.

The audience level and turnover data shown above support this point.

- The audience gains in the shorter interview are among the demographic sub-groups that generally are busiest and have least time to be interviewed. For demographic sub-groups with the most time to be interviewed -- older, unemployed -- the "pure" and "traditional" interviews produce about the same audience measures.
- The shorter interview picks up/captures more casual reading -- exactly the kind most likely to get passed over in a time-rushed interview.

Summary and Conclusions

The test shows very clearly that interview length affects basic readership measures.

In addition to the general quality-related benefits of reducing interviewer and respondent burden, a shorter, print-focussed interview produces different, more realistic audience measures because it is better able to capture readership of demographic groups that are both the most difficult to interview and the heaviest readers.

Based on these results, Simmons has introduced the Survey of American Readership, a stand-alone survey dedicated to print audience measurement.

Appendices

In the interest of allowing readers a closer look at test results, the tables on the following pages show "pure" vs. "traditional" comparisons in some detail.

Indices are shown for screens and reads, on both unweighted and weighted/projected bases. Unweighted values are based on comparisons of raw respondent answers; weighted results are based on comparisons after separately weighting and projecting each sample to the U.S. adult population, as described in the body of this report.

Table I shows standard demographic sub-groups, for all publications combined. Table II shows various publication sub-groups, for total adults.

The statistical significance of any result shown in Table I or II depends on several factors, including the magnitude of the reported index, the number of respondents in the "pure" and "traditional" cells, and the number and aggregate readership of the magazines combined.

To keep from reporting the most unstable results:

- Table I shows only demographic groups with at least 100 "pure" and "traditional" respondents.
- Table II shows only publication groups with at least 5 titles and 100 "pure" and "traditional" readers.

As described in the body of this report, "pure" vs. "traditional" differences are significant for all magazines for total adults and for several of the larger demographic sub-groups.

As a rule of thumb, approximate index values required for statistical significance at the .05 level for other demographic sub-groups shown in Table I are:

Size of Demographic Group	Index Required for Statistical Significance at .05 Level
100% (Adults)	110 or more
50% (e.g., Women)	115 or more
25% (e.g., IEI \$20-50K)	120 or more
10% (e.g., Age 18-24)	130 or more

Index values required for significance for the magazine sub-groups shown in Table II are more complicated, because of varying levels of readership duplication. As a very rough approximation, the index values shown above may be used, replacing "Size of Demographic Group" with "Percent of All Magazine Reading".

Appendix 1

Table I

**Indices of "Pure" vs. "Traditional" Screens and Reads
All Magazines Combined
("Traditional" = 100)**

	Unweighted		Weighted	
	Screens	Reads	Screens	Reads
Total Adults	121	114	116	113
Men	122	115	111	110
Women	122	115	120	116
Age 18-24	110	110	118	113
25-34	128	117	133	131
35-44	133	137	118	130
45-54	125	115	118	107
55-64	108	99	101	89
65+	95	88	85	84
Att/Grad College	127	117	120	116
Graduated High School	123	120	122	119
Did Not Graduate H.S.	96	99	93	96
IEI \$50,000 or more	158	135	143	127
\$20,000 to \$49,999	124	118	121	120
Less than \$20,000	115	117	117	122
Not Employed	106	105	107	100
HHI \$60,000 or more	145	130	134	127
\$40,000 to \$59,999	127	116	120	119
\$25,000 to \$39,999	104	90	105	89
Less than \$25,000	116	127	108	119
Prof/Man/Admin	132	118	130	114
Tech/Clerical/Sales	134	127	131	128
Other Employed	111	113	109	117
Not Employed	106	105	107	100
Single	101	101	110	107
Married	127	119	118	115
Divorced/Separated	124	131	119	131
Widowed	103	93	102	94
Black	105	110	107	120
Northeast	150	150	134	141
Midwest	111	110	106	111
South	99	94	98	96
West	136	114	142	124

A-2

Table II

**Indices of "Pure" vs. "Traditional" Screens and Reads
Total Adults
("Traditional" = 100)**

	Unweighted		Weighted	
	Screens	Reads	Screens	Reads
All Magazines	121	114	116	113
Weeklies	118	108	115	106
Bi-weeklies	127	111	136	126
Monthlies	123	116	116	115
Bi-monthlies	121	124	115	123
Circulation (000)				
2,000 or more	108	102	106	102
1,000 to 1,999	126	126	121	122
500 to 999	131	125	122	122
300 to 499	135	124	126	125
299 or less	137	133	122	131
Genre				
Automotive	127	117	104	106
Business & Finance	125	105	133	112
Children & Parenting	121	114	118	112
Epicurean	114	114	118	121
Fishing & Hunting	132	142	117	131
General Appeal	109	103	109	103
Health	132	140	132	147
Home & Home Services	104	106	101	104
Men's	121	131	105	118
News	116	103	113	101
Science/Technology	149	131	150	140
Special Appeal	123	108	117	114
Sports	152	153	135	143
Women's	112	108	114	113
Women's Fash/Beaut	124	125	133	145
Other	113	126	127	124
All Daily Newspapers	121	99	123	104
All Sunday Newspapers	107	99	108	103

