

Measurement of Magazine Readership via the Internet

Martin Frankel, Julian Baim, Michal Galin and Michelle Leonard, MRI

ABSTRACT

In the spring of 2002 MRI began a series of experiments that culminated with a large scale test of AIR measurement of consumer magazines via the internet. There were two basic objectives of this test. The first objective was to determine the level of readership and audience composition obtained via an internet based sample and data collection procedure. The second objective was to determine the impact of sample source, number of titles and stimuli composition within this sample and data collection context. An overall sample of 8800 respondents was distributed over a number of survey treatment combinations based on: sample source, number of titles, and type of stimuli presentation for screening. Two basic analyses of results are reported in this paper. First, screen-in and AIR levels are compared with results that are obtained with the MRI syndicated readership survey. Second, the impact of sample source, number of titles and mode of screen presentation are examined with respect to overall levels screen-in and reading.

INTRODUCTION AND DEVELOPMENT OF RESEARCH DESIGN

Over the years MRI has undertaken a number of research initiatives designed to better understand the process of readership measurement and estimation. Given the recent development of internet based surveys, MRI wished to explore the possibility of conducting measures of consumer magazine readership over the internet.

This research effort began with the development of a number of “screen designs” that might be used for web based data collection. Taking a zero based approach, we started from scratch and considered the possibility of a daily diary-based, data collection, a data collection that was “issue” specific and a data collection that did not use “screening” but went directly to a question that would produce an AIR estimate. In the latter case we considered the use of a “Frequency” scale, a “Probability Meter” and a direct aided or unaided recall question about a specific time interval.

Recognizing that departure from the generally accepted recent reading method, including the use of a “screen-in” question would produce results that would confound issues associated with the use of the internet with more fundamental measurement issues, we decided that the first large scale attempt should preserve most of the basic design that is currently used in the US to produce AIR currency levels.

Even within this overall plan, however, the number of “possible variations” was too large to contemplate a fully balanced and randomized design. For this reason we proceeded by dividing our possible research variations into three groups: those that would be examined and decided in the pre-pilot stage, those that would be examined and decided in the pilot stage, and those that would form the basic study design.

PRE-PILOT DECISIONS

In the pre-pilot stage we examined certain issues related to the design of the “web page” used for the screen-in and subsequent readership questions. Design issues focused on the layout of the individual magazine title stimuli. After examining a number of mockups we settled on the presentation of 24 magazines on each “web page.” We also decided to array the stimuli in 4 columns and 6 rows. Even though we knew that one of our variations would involve reduced size magazine covers in color, we decided to present magazine logos in black and white. We decided to equalize the screen area associated with each logo, rather than attempting to equalize logo type size.

We examined a number of possible sources of samples including companies that “conducted” internet studies using their own internet panels and companies that had mail panels with e-mail addresses. We also considered the use of RDD screening to collect email addresses and the use of email addresses from our syndicated study a year or two after our in-person interview. After considering issues of cost and sample control we decided to purchase two different samples from Survey Sampling International. SSI is well known in the US as a source of RDD phone numbers and entered the business of providing internet enabled samples several years ago. The SSI sample model is somewhat different than the model used to provide RDD samples. For RDD samples, SSI turns over a sample of telephone numbers to the research organization carrying out the research. For internet samples, SSI sends an email invitation and asks potential respondents to visit an SSI web site where they are “redirected” to the web site of the research

organization. Two different cross sectional samples are offered: SSI LITE and SSI Spot. Neither is offered as a probability sample, but the methods used to obtain email addresses are sufficiently different that we decided to make sample type an experimental treatment condition.^[1]

PILOT TESTED DECISIONS

We recognized that neither the SSI LITE nor the SSI Spot samples were probability samples, but we wished to obtain response rates from these samples that were as high as reasonably possible. For this reason we decided to offer respondent incentives that were considerably higher than customary for internet surveys requiring from 5 to 20 minutes of a respondents time. General population surveys conducted over the internet typically use incentives in the \$2-\$5 dollar range. Some may go as high as \$10. Incentives in the \$2-\$5 range typically produce response rates (i.e. the percent of completed interviews relative to the total number of email invitations) in the 10% range. We wanted to see if it was possible to increase response rates above these levels. For this reason we carried out a pilot study involving 600 invitations, randomized over two groups. Three hundred invitations offered \$10 and 300 invitations offered \$20. We used the SSI LITE for the pilot sample.

We were somewhat surprised that these two incentive levels produced only slight differences in response rates. The offer of \$20 produced a response rate of 34.5%, while the offer of \$10 produced a response rate of 29.5. Given the highly differential cost associated with the two incentives distributed over nearly 9,000 targeted interviews, we decided that there would be very little quality benefit in offering the higher incentive. In the final study all potential respondents were offered a \$10 incentive.

A second decision based on the results of the pilot test involved the nature of the response that would be required to the "screen in" question. At issue was whether or not to make use of a "yes-no" required response for each magazine or a "yes only" response to each of the 24 titles presented on a specific screen in page. In the syndicated readership study conducted by Mediamark the average number of "screen-ins" over approximately 240 magazine titles is 14-15. In the pilot test of 192 titles, the yes only screen-in produced an average of 14 screen-ins, while the yes-no required response produced 22 screen-ins on average. Because our goal was to attempt to examine the degree to which internet data collection would parallel the results of the national study, we made use of the positive only response.

THE BASIC STUDY DESIGN

Our final design involved a full factorial implementation of three treatment types. Specifically, the treatments we examined consisted of:

- | | | |
|------|------|---|
| I. | I. | Three stimuli sets: Logos, Names, Logos and Cover |
| II. | II. | Three "number of title" sets: 48, 96 and 192 |
| III. | III. | Two sample sources: SSI LITE and SSI Spot |

The test was implemented as a full-factorial design. That is, each of the $3 \times 3 \times 2 = 18$ scenarios was presented. Sufficient email invitations were sent in order to obtain approximately 450 completed web interviews for each scenario.

Examples of the web-pages shown to respondents for the three different stimuli sets are shown in Appendix A. One stimulus set showed Logos only, a second stimulus set showed magazine Names only (all in the same type face unrelated to the logo), and a third stimulus set showed Logos and Covers. The first two stimuli sets were shown in black and white. For the third stimulus set, the Logos were in black and white while the Covers were in color.

In order to maximize the sample size for the 48 largest magazines, while maintaining our ability to examine impacts of screen positioning and potential confusion, the following assignment procedure was used for the 3 "number of title"

^[1] According to Survey Sampling Inc, "Survey Spot records come from many sources, including banner ads, online recruitment methods, and RDD telephone recruitment. Survey Spot members are recruited exclusively using permission-based techniques. SSI does not use unsolicited e-mail or "spam" in building its Survey Spot panel." In contrast the SSI LITE sample is a "The SSI-LITE e-mail files are based on self-reported data gathered principally through a variety of permission-based marketing sources. These databases are comprised of e-mail addresses of people who have signed up and agreed to receive e-mail messages on a particular topic, usually by checking a box."

sets. A group of the 48 largest magazines (using MRI's audience ratings) was assigned to group A. A second group of 48 magazines (the next largest in terms of audience) was assigned to group B. The remaining 96 magazines were assigned to group C. Group A magazines appeared in all title sets. That is, these titles were presented to all respondents. Group B magazines appeared for the 96 and 192 "number of title" set scenarios. That is, these magazines were shown to approximately two-thirds of the respondents. Group C magazines were used for the 192 "number of title" set scenario only. Thus, they were shown to one-third of the respondents.

Full randomization of titles was carried out within groups A, B and C as follows: For the 48 "number of titles" scenario, the 48 magazines were randomly assigned over two web pages A1, and A2, 24 titles on each page. For the 96 "number of titles" scenarios, a random permutation of 4 pages called A1, A2, B1 and B2 was carried out. For example, a specific permutation or order might be A2, B1, A1, and B2. Group A titles were randomly assigned, 24 each to A1 and A2 and Group B titles were randomly assigned, 24 each, to B1 and B2. The same procedure was used for the 192 "number of titles" set scenario. Eight pages A1, A2, B1, B2, C1, C2, C3 and C4 were randomly permuted and 24 titles were randomly assigned with the appropriate group page type.

Appendix B, shows the specific 192 magazines by grouping A, B and C.

All titles that were "screened-in" were followed with a question about frequency of reading (x out of 4 issues) within issue period. Then, the recent reading question within issue period was asked for all screened-in titles. The ordering of titles these questions was randomized by length of issue period. That is, for a random half of the respondents issue periods were ordered from weeklies to bi-monthlies. For the other random half the ordering was bi-monthlies to weeklies. Title order within issue period were randomly assigned either alpha or reverse alpha.

The frequency and reading within last issue period questions made use of the same stimuli that was used in the screen-in process. Appendix C shows examples of these questions for the Logo stimuli set scenario.

SAMPLE RELEASE AND CONTROL

We expected that the two sample sources would have different response rate, velocity and demographic composition.^{2[2]} For this reason we divided two samples of email addresses (potential invitations) into 135 replicates of 500 addresses each. We began the study on Monday, February 24, 2003 and released 2 replicates for each sample. The release schedule originally planned for 9 replicates to be released per week per sample (2 replicates both on Monday and Tuesday and 1 replicate for each remaining day of the week). Within each replicate, the release was to be 50% male and 50% female. Both the number of replicates released and the male-female composition of each were modified as time progressed in order to account for lagging male response for both Spot and LITe and lagging LITe response overall. By March 25, 2003, the Spot sample had reached its target size quota and was completed (a total of 40 replicates accounting for 20,000 unique invitations were utilized for this sample). On March 17, the number of replicates released to the LITe sample was doubled because overall response lagged behind that of Spot. Invitations to the LITe sample continued until April 9, 2003. All 135 of the replicates generated were utilized for the LITe sample (accounting for 67,500 unique invitations). Tables 1 and 2 depict the response for the Spot and LITe samples, respectively. The "terminated" response shown in the tables describes two types of terminations: respondents being terminated based on sex quotas (specifically, females) and respondents being terminated because the treatment scenario they were randomly placed into had met quota and was closed. Of the two latter termination scenarios, the first pertains to the LITe sample while the second can be found only in cases of Survey Spot.

Table 1: Survey Spot Response

	Completes		Terminated	Total			Response			Percent including Terminated
Date	Male	Female		Male	Female	Total	Male	Female	Total	
2/24-3/9	971	1245	13	4500	4500	9000	21.6%	27.7%	24.6%	24.8%

^{2[2]} Survey Sampling indicated that Survey Spot performs about 10 times better than the LITe sample. Survey Sampling told us to expect about a 30% click through rate for Survey Spot, whereas for LITe to expect about a 2 or 3% click through rate. The click through rate is the rate at which potential respondents access the survey website through the URL on the invitation sent to them.

3/10-3/23	1226	829	117	6000	3000	9000	20.4%	27.6%	22.8%	24.1%
3/24-25	112	126	207	1000	1000	2000	11.2%	12.6%	11.9%	22.3%
Total	2309	2200	337	11500	8500	20000	20.1%	25.9%	22.5%	24.2%

* Between 2/24 and 3/9 the replicates sent out had a 50:50 male:female composition

** Between 3/10 and 3/23 the replicates sent out had a 67:33 male:female composition

*** Between 3/24 and 3/25 the replicates sent out had a 50:50 male:female composition

Table 2: LITE Response

Date	Completes		Terminated	Total			Response			Percent including Terminated
	Male	Female		Male	Female	Total	Male	Female	Total	
2/24-3/9	358	1050	5	4500	4500	9000	7.9%	23.3%	15.7%	15.7%
3/10-3/23	399	920	19	9000	4500	13500	4.4%	20.4%	9.8%	9.9%
3/24-3/27	321	264	47	7200	1800	9000	4.5%	14.7%	6.5%	7%
3/28-4/9	934	0	890	36000	0	36000	2.6%	0	2.6%	5.1%
Total	2012	2234	961	56700	10800	67500	3.5%	20.7%	6.3%	7.7%

* Between 2/24 and 3/9 the replicates sent out had a 50-50 male-female composition

** Between 3/10 and 3/23 the replicates sent out had a 67-33 male-female composition

*** Between 3/24 and 3/27 the replicates sent out had an 80-20 male-female composition

**** Between 3/28 and 4/9 the replicates sent out were 100% male

For each sample, an algorithm was used to randomly assign persons who visited the introductory URL to one of 9 treatment scenarios (3 stimuli sets x 3 number of titles sets). Thus, we were able to achieve randomized assignment to treatment combinations within sample. We did not intend to explicitly monitor the demographic composition of the sample except for gender. As described above, as the sample fulfillment proceeded over time it was clear that a the Spot sample was producing close to a 50:50 split of males and females, while the LITE sample was producing a disproportionately higher percent of females. We therefore adjusted the sample so that the Spot sample invitations were distributed 80:20 male and later 100% male.

RESULTS

Our results section is divided into three parts. First, we describe the demographic compositions of the two samples and compare them with the total US and the Internet Enabled US. Next, we examine overall differences by the experimental conditions. Finally we compare weighted results for screen levels and AIR with the Mediamark national syndicated study.

DEMOGRAPHIC COMPOSITION OF THE TWO SAMPLES

After removing cases with incomplete demographic data (these cases had started but not completed the questionnaire) we obtained final samples of 4,464 respondents from the Spot Sample and 4,215 respondents from the LITE sample.

Table 3 shows the distribution of the two samples as well as the corresponding demographic characteristics for all US adults and Internet Enabled US adults. The US estimates come from the US Census and the Internet Enabled US adults

come from the weighted (to US census) Mediamark syndicated survey. This table shows the demographic characteristics of Gender, Age, Marital Status, Employment Status, Education, Household Income, Race and Hispanic Ethnicity. The demographic compositions of the two samples are quite similar to each other, but the Spot sample, as expected, was somewhat closer to the US Census. As also anticipated there are major sample shortfalls for the two ends of the age distribution and the lower end of the Education and Income distribution. Most notable is the almost complete lack of respondents who claim less than a high school degree.

Table 3: Sample Demographics and US Census

DEMOGRAPHIC		SAMPLE SOURCE			
		Spot	LITe	Census	Internet Enabled (MRI, Spring 2003)
Gender	<i>Male</i>	51.1%	47.3%	48.0%	48.5%
	<i>Female</i>	48.9%	52.7%	52.0%	51.5%
Age	<i>18-24</i>	10.3%	5.2%	13.1%	14.6%
	<i>25-34</i>	20.5%	15.7%	18.5%	19.7%
	<i>35-44</i>	25.6%	28.3%	21.1%	23.3%
	<i>45-54</i>	24.4%	30.9%	18.9%	20.1%
	<i>55-64</i>	14.3%	14.4%	12.4%	11.7%
	<i>65+</i>	4.9%	5.5%	16.1%	10.6%
Marital Status	<i>Married</i>	57.4%	62.4%	56.7%	59.5%
	<i>Divorced</i>	14.9%	14.0%	12.5%	9.4%
	<i>Separated</i>	2.1%	2.2%	2.1%	1.9%
	<i>Widowed</i>	2.2%	2.4%	4.3%	4.1%
	<i>Single, Never Married</i>	23.5%	19.0%	24.4%	25.1%
Employment	<i>Working Full Time</i>	55.0%	48.2%	48.8%	59.3%
	<i>Working Part Time</i>	15.8%	15.8%	15.0%	11.7%
	<i>Not Employed</i>	29.2%	36.0%	36.2%	29.0%
Education	<i>Did Not Graduate High School</i>	2.3%	3.2%	16.9%	10.1%
	<i>High School Graduate</i>	19.5%	24.5%	31.8%	29.8%
	<i>Some College, No Degree</i>	36.5%	35.2%	21.5%	22.0%
	<i>Associate Degree</i>	10.5%	10.0%	5.4%	9.0%
	<i>Graduated 4 Year College+</i>	24.4%	19.9%	16.7%	19.5%
Household Income	<i>Graduate Degree</i>	8.8%	7.3%	7.6%	9.6%
	<i>Less Than 15K</i>	9.1%	9.6%	11.1%	6.9%
	<i>15-25 K</i>	13.0%	14.0%	11.2%	7.8%
	<i>25-50 K</i>	33.0%	36.4%	26.9%	25.4%
	<i>50-75 K</i>	22.7%	23.1%	20.3%	23.4%
	<i>75-100 K</i>	12.5%	10.3%	13.0%	15.3%
Race	<i>100-150 K</i>	7.0%	5.1%	13.1%	13.2%
	<i>150K +</i>	2.7%	1.6%	4.5%	8.0%
	<i>White</i>	89.4%	90.3%	77.3%	84.9%
	<i>Black</i>	5.1%	4.5%	11.1%	10.4%
Race	<i>Asian</i>	2.0%	2.3%	3.1%	3.0%
	<i>Other</i>	3.5%	2.9%	8.4%	1.7%
	Hispanic	<i>Yes</i>	3.9%	3.8%	11.9%

	No	96.1%	96.2%	88.1%	91.1%
--	----	-------	-------	-------	-------

SCREEN-IN AND READ (AIR) LEVELS BY TREATMENT

Table 4 shows the distribution of the experimental conditions among the two samples. As this table shows, the allocation algorithm within sample type functioned as planned and the distribution of respondents among the 9 treatment combinations and across the two sets of margins was sufficiently balanced to allow multi-way analysis of variance.

Table 4: Sample Distribution Among Samples and Treatments

Sample	Visual Representation	Number of Titles			Total
		48 Titles	96 Titles	192 Titles	
Spot	Logo	494	499	497	1490
	Names	495	493	493	1481
	Logo & Cover	498	498	497	1493
	Total	1487	1490	1487	4464
LITe	Logo	468	471	465	1404
	Names	468	469	470	1407
	Logo & Cover	465	469	470	1404
	Total	1401	1409	1405	4215

In order to examine differences by sample source, stimuli and number of titles we summarized the total number of screen-ins and total number reads (i.e. titles read in the last issue period) for the 48 titles that appeared in all treatment scenarios.

Table 5 shows the mean number of screen-ins, as well as standard deviations and sample sizes, in total and for the various treatment cells for the SSI Spot sample. Table 6 the corresponding summary information (mean screen-ins, standard deviation and sample sizes) for the SSI LITe sample.

Table 5: Survey Spot Screens (48 titles) by Treatment and Sample Source
(Mean, Standard Deviation, Sample Size)

Stimuli	48 Titles	96 Titles	192 Titles	Total
<i>Logo</i>	8.061	7.599	7.783	7.813
	6.212	6.126	6.254	6.196
	494	499	497	1490
<i>Names</i>	7.754	7.927	8.316	7.999
	5.747	5.746	7.069	6.219
	495	493	493	1481
<i>Logo & Cover</i>	7.924	7.902	7.620	7.815
	6.103	6.117	5.648	5.958

	498	498	497	1493
Total	7.913	7.809	7.905	7.875
	6.021	5.998	6.351	6.124
	1487	1490	1487	4464

Table 6: LITe Screens (48 titles) by Treatment and Sample Source
(Mean, Standard Deviation, Sample Size)

Stimuli	48 Titles	96 Titles	192 Titles	Total
Logo	8.517	9.363	8.910	8.931
	6.986	7.203	7.238	7.147
	468	471	465	1404
Names	9.248	8.808	9.426	9.161
	7.322	7.133	7.744	7.404
	468	469	470	1407
Logo & Cover	9.533	9.115	8.045	8.895
	7.433	6.985	6.404	6.975
	465	469	470	1404
Total	9.099	9.096	8.793	8.996
	7.257	7.106	7.167	7.177
	1401	1409	1405	4215

The overall mean number of screens for the 48 common titles differs by more than one (7.88 vs. 9.00), between the Spot and the LITe samples. This difference is statistically significant at the 1% level. Tables 7 and 8 show average number of reads (AIR levels) for the same 48 common titles. The mean number of reads also differs substantially (4.21 vs. 5.36) and significantly between the two different sample sources as well.

Table 7: Survey Spot Reads (48 titles) by Treatment and Sample Source
(Mean, Standard Deviation, Sample Size)

Stimuli	48 Titles	96 Titles	192 Titles	Total
Logo	4.322	4.210	4.020	4.184
	4.224	4.394	3.961	4.196
	494	499	497	1490
Names	4.180	4.105	4.485	4.257
	4.338	4.208	4.461	4.337
	495	493	493	1481
Logo & Cover	4.261	4.157	4.107	4.175
	4.767	4.309	4.004	4.369
	498	498	497	1493
Total	4.254	4.158	4.203	4.205

	4.447	4.302	4.150	4.300
	1487	1490	1487	4464

Table 8: LITe Reads (48 titles) by Treatment and Sample Source
(Mean, Standard Deviation, Sample Size)

Stimuli	48 Titles	96 Titles	192 Titles	Total
<i>Logo</i>	5.209	5.709	5.194	5.372
	5.656	5.634	5.328	5.543
	468	471	465	1404
<i>Names</i>	5.526	4.951	5.860	5.446
	5.647	5.416	6.258	5.793
	468	469	470	1407
<i>Logo & Cover</i>	5.794	5.333	4.634	5.251
	6.117	4.991	4.738	5.331
	465	469	470	1404
<i>Total</i>	5.509	5.331	5.229	5.356
	5.811	5.359	5.497	5.558
	1401	1409	1405	4215

Given the significant differences for both screens and reads between sample sources, we chose to perform separate analyses of variance. We specified both main treatment effects (Stimuli and Number of Titles) as well as interactions. For both sample sources, no main effects were statistically significant. For the LITe sample source the interaction between Stimuli and Number of Titles was significant at the 5% level. This significant interaction is traceable to the substantially lower number of screens and reads for the Logo and Cover stimulus presented in conjunction with 192 titles. We note, however, that overall neither the stimuli for screen-in presentation nor the number of titles were significantly different from one another at the overall level.

COMPARISON OF READ (AIR) LEVELS WITH MRI SYNDICATED STUDY.

Because the differences we observed by sample source were both substantively and statistically significant, we decided to keep the two samples separate for the comparison of results with the MRI syndicated study. Given the lack of statistical significance associated with the other main effect treatment conditions (stimuli sets and number of titles) we felt that aggregation over these conditions was appropriate for title by title comparisons. Table 9 shows the average issue audience ratings (percent of total adults) that read any issue of the title within the issue period. Four columns show separately the results from the SSI Spot sample, the SSI LITe Sample, the most recent MRI syndicated study and the “Internet enabled” persons in the most recent MRI syndicated study. In order to remove the impact of the differences in demographic composition between the MRI syndicated study and the Spot and LITe samples, both of these samples were weighted (using multidimensional raking) to the same US Census demographics used for weighting the MRI syndicated study. These demographics include Gender, Age, Income, Education, Race and Ethnicity. All weighting was carried out within gender, as is the practice for the MRI syndicated study.

TABLE 9: AIR LEVELS (As Ratings) 48 Titles

Title	Spot	LITe	MRI	MRI-Internet	Spot Index	LITe Index
<i>Better Homes &</i>	17.8%	21.4%	18.2%	19.8%	97.8%	117.1%

<i>Gardens</i>						
<i>Car And Driver</i>	8.0%	6.0%	4.8%	5.6%	167.3%	126.3%
<i>Cooking Light</i>	7.0%	8.6%	4.6%	5.3%	151.8%	186.0%
<i>Cosmopolitan</i>	8.6%	12.0%	8.5%	9.4%	101.9%	141.8%
<i>Country Home</i>	3.5%	5.4%	3.5%	3.8%	100.3%	153.0%
<i>Country Living</i>	5.6%	8.1%	5.1%	5.7%	110.8%	160.0%
<i>Ebony</i>	5.8%	5.6%	5.5%	5.4%	104.9%	100.7%
<i>Entertainment Weekly</i>	7.2%	9.7%	4.6%	5.1%	157.1%	213.5%
<i>ESPN The Magazine</i>	4.5%	6.8%	4.8%	5.5%	93.6%	142.9%
<i>Essence</i>	4.7%	4.3%	3.8%	4.0%	124.5%	113.5%
<i>Family Circle</i>	13.8%	23.9%	10.4%	11.1%	131.8%	228.7%
<i>Field & Stream</i>	5.7%	6.8%	5.1%	5.4%	112.7%	134.3%
<i>Glamour</i>	6.5%	9.0%	5.7%	6.6%	115.1%	158.0%
<i>Good Housekeeping</i>	15.0%	18.9%	11.5%	12.3%	130.1%	164.0%
<i>House & Garden</i>	8.8%	10.2%	6.3%	6.9%	139.3%	161.5%
<i>Jet</i>	4.3%	4.8%	4.0%	3.8%	106.1%	119.5%
<i>Ladies' Home Journal</i>	8.3%	10.6%	6.2%	6.7%	132.7%	170.2%
<i>Martha Stewart Living</i>	5.2%	5.2%	6.3%	7.4%	83.2%	83.5%
<i>Maxim</i>	11.4%	12.5%	6.0%	7.3%	191.7%	209.5%
<i>Men's Health</i>	6.7%	8.2%	4.5%	5.3%	149.4%	182.7%
<i>Money</i>	5.6%	7.2%	3.6%	4.2%	153.5%	197.4%
<i>Motor Trend</i>	5.4%	5.8%	3.5%	3.9%	155.5%	167.3%
<i>National Enquirer</i>	10.1%	13.6%	5.8%	5.8%	173.4%	234.7%
<i>National Geographic</i>	17.0%	14.7%	15.2%	16.8%	111.5%	96.6%
<i>Newsweek</i>	12.5%	14.2%	9.4%	10.8%	132.7%	151.0%
<i>"O, The Oprah Magazine"</i>	5.5%	7.1%	5.9%	6.8%	93.4%	120.4%
<i>Parenting</i>	7.2%	7.3%	5.2%	5.8%	138.3%	141.5%
<i>Parent's Magazine</i>	7.3%	9.2%	6.7%	7.5%	109.5%	137.5%
<i>People</i>	18.9%	22.0%	17.2%	19.0%	109.6%	127.7%
<i>Playboy</i>	11.7%	12.0%	4.7%	5.0%	247.3%	255.2%
<i>Popular Mechanics</i>	6.1%	7.3%	4.7%	5.2%	130.8%	154.8%
<i>Prevention</i>	8.2%	11.5%	5.0%	5.5%	163.3%	228.1%
<i>Reader's Digest</i>	26.9%	28.2%	20.6%	21.2%	130.7%	137.1%
<i>Redbook</i>	7.0%	12.3%	4.2%	4.7%	164.3%	291.4%
<i>Rolling Stone</i>	5.4%	9.6%	5.2%	6.1%	102.7%	183.4%
<i>Seventeen</i>	4.7%	7.0%	3.9%	4.5%	121.1%	179.7%
<i>Smithsonian</i>	5.5%	5.5%	3.6%	4.1%	154.8%	154.0%
<i>Southern Living</i>	7.9%	7.2%	7.0%	7.6%	113.4%	103.0%
<i>Sports Illustrated</i>	11.5%	12.9%	10.0%	11.0%	115.9%	129.2%
<i>Star</i>	5.0%	8.9%	3.0%	3.0%	166.8%	294.2%
<i>Teen People</i>	5.6%	9.4%	3.7%	4.2%	149.6%	252.2%
<i>The Cable Guide</i>	5.0%	3.6%	3.8%	3.8%	132.2%	96.9%
<i>Time</i>	14.0%	17.4%	11.2%	12.6%	125.4%	155.1%
<i>TV Guide</i>	20.2%	26.1%	12.9%	13.1%	156.0%	201.8%
<i>U.S. News & World Report</i>	8.9%	8.4%	5.6%	6.4%	158.1%	149.7%
<i>Vogue</i>	3.9%	5.7%	4.9%	5.5%	79.1%	115.8%
<i>Woman's Day</i>	15.3%	21.9%	9.4%	10.0%	162.9%	233.5%
<i>Woman's World</i>	7.3%	8.7%	3.6%	3.9%	199.6%	239.8%
Average (of ratings)	8.9%	10.9%	6.9%	7.5%		

Several things are notable about the AIR levels. Both the Spot and LITE AIR levels are higher, on average, than the MRI syndicated currency levels and even the MRI internet enabled levels. More importantly, however, examination of the Indices for Spot and LITE (Spot Index=(Spot Rating/MRI Rating)*100%) show that the differential among magazines is highly variable. Thus, the relative ordering of audience levels between the different Internet Samples and between these samples and the MRI syndicated levels is not constant, but highly variable.

Table 10 shows the gender composition (percent Male vs. Female) for the Spot, LITE and MRI syndicated samples respectively. This table indicates that for some titles there is close agreement while for others there is substantial difference between the web results and those found in the MRI syndicated study. We also observe attenuation toward the mean as follows: For the 20 titles where MRI shows that males comprise less than 25% of the audience, the average percent males is 15.0%. For the same titles, percent male is 20% and 18.7% for the Spot and LITE Samples respectively. Among the 9 titles where MRI shows that the male audience is 75% or above, the average percent males is 84.2%. These titles show an average male percent of 82.3% and 80.1% for Spot and LITE respectively. As expected, the internet based samples tend to flatten the gender differentiation shown by gender “targeted” titles.

TABLE 10: GENDER COMPOSITON (48 Titles)

Title	Spot Sample		LITE Sample		MRI Syndicated Study	
	Male	Female	Male	Female	Male	Female
<i>Better Homes & Gardens</i>	27.0%	73.0%	28.4%	71.6%	21.9%	78.1%
<i>Car And Driver</i>	87.2%	12.8%	81.6%	18.4%	89.6%	10.4%
<i>Cooking Light</i>	23.4%	76.6%	25.6%	74.4%	15.6%	84.4%
<i>Cosmopolitan</i>	21.5%	78.5%	17.8%	82.2%	17.4%	82.6%
<i>Country Home</i>	29.1%	70.9%	23.5%	76.5%	21.6%	78.4%
<i>Country Living</i>	26.3%	73.7%	34.7%	65.3%	23.7%	76.3%
<i>Ebony</i>	51.2%	48.8%	44.4%	55.6%	38.4%	61.6%
<i>Entertainment Weekly</i>	50.0%	50.0%	57.5%	42.5%	43.8%	56.2%
<i>ESPN The Magazine</i>	78.1%	21.9%	78.8%	21.2%	81.2%	18.8%
<i>Essence</i>	29.1%	70.9%	28.7%	71.3%	25.5%	74.5%
<i>Family Circle</i>	11.1%	88.9%	13.3%	86.7%	9.3%	90.7%
<i>Field & Stream</i>	74.9%	25.1%	76.1%	23.9%	83.0%	17.0%
<i>Glamour</i>	14.1%	85.9%	7.2%	92.8%	10.6%	89.4%
<i>Good Housekeeping</i>	17.8%	82.2%	16.5%	83.5%	12.9%	87.1%
<i>House & Garden</i>	42.9%	57.1%	31.1%	68.9%	25.3%	74.7%
<i>Jet</i>	57.4%	42.6%	40.4%	59.6%	37.9%	62.1%
<i>Ladies' Home Journal</i>	11.2%	88.8%	10.0%	90.0%	6.2%	93.8%
<i>Martha Stewart Living</i>	18.6%	81.4%	21.3%	78.7%	13.7%	86.3%
<i>Maxim</i>	80.5%	19.5%	71.5%	28.5%	76.8%	23.2%
<i>Men's Health</i>	86.4%	13.6%	83.8%	16.2%	85.9%	14.1%
<i>Money</i>	66.0%	34.0%	76.5%	23.5%	66.9%	33.1%
<i>Motor Trend</i>	89.8%	10.2%	84.6%	15.4%	90.2%	9.8%
<i>National Enquirer</i>	42.3%	57.7%	34.3%	65.7%	37.1%	62.9%
<i>National Geographic</i>	52.4%	47.6%	68.0%	32.0%	55.8%	44.2%
<i>Newsweek</i>	56.5%	43.5%	64.2%	35.8%	55.5%	44.5%

<i>"O, The Oprah Magazine"</i>	19.2%	80.8%	19.9%	80.1%	12.4%	87.6%
<i>Parenting</i>	24.8%	75.2%	22.4%	77.6%	18.2%	81.8%
<i>Parent's Magazine</i>	18.5%	81.5%	21.6%	78.4%	19.0%	81.0%
<i>People</i>	40.1%	59.9%	41.4%	58.6%	35.5%	64.5%
<i>Playboy</i>	82.6%	17.4%	82.8%	17.2%	83.9%	16.1%
<i>Popular Mechanics</i>	86.1%	13.9%	87.4%	12.6%	88.7%	11.3%
<i>Prevention</i>	29.0%	71.0%	28.6%	71.4%	20.7%	79.3%
<i>Reader's Digest</i>	43.7%	56.3%	40.3%	59.7%	41.1%	58.9%
<i>Redbook</i>	12.8%	87.2%	8.1%	91.9%	10.2%	89.8%
<i>Rolling Stone</i>	65.6%	34.4%	52.2%	47.8%	60.3%	39.7%
<i>Seventeen</i>	27.2%	72.8%	14.2%	85.8%	16.4%	83.6%
<i>Smithsonian</i>	50.6%	49.4%	63.1%	36.9%	47.2%	52.8%
<i>Southern Living</i>	30.3%	69.7%	26.1%	73.9%	23.5%	76.5%
<i>Sports Illustrated</i>	75.1%	24.9%	74.6%	25.4%	78.4%	21.6%
<i>Star</i>	35.1%	64.9%	33.5%	66.5%	31.0%	69.0%
<i>Teen People</i>	41.3%	58.7%	46.1%	53.9%	28.2%	71.8%
<i>The Cable Guide</i>	65.2%	34.8%	55.8%	44.2%	45.3%	54.7%
<i>Time</i>	56.0%	44.0%	61.3%	38.7%	54.8%	45.2%
<i>TV Guide</i>	52.7%	47.3%	49.3%	50.7%	42.3%	57.7%
<i>U.S. News & World Report</i>	68.0%	32.0%	70.7%	29.3%	62.4%	37.6%
<i>Vogue</i>	19.1%	80.9%	19.5%	80.5%	13.4%	86.6%
<i>Woman's Day</i>	7.8%	92.2%	9.2%	90.8%	5.7%	94.3%
<i>Woman's World</i>	11.9%	88.1%	6.2%	93.8%	6.9%	93.1%

CONCLUSIONS

The analysis reported in this paper only scratches the surface of the further analyses that we intend to carry out with this data set. However, based on these initial results we can conclude:

1. 1. Audience levels (AIR) obtained from an economically feasible internet based data collection procedure do not mirror those obtained by the current MRI procedure based on face-to-face interviewing in a probability sample of households.
2. 2. Differences between the audience estimates and those reported by the current MRI study are not "constant" across title. This conclusion holds for each of two internet sample sources tested.
3. 3. Differences between internet measurement and the MRI syndicated study are not restricted to total AIR. There are substantial differences in gender composition of the total audience between the difference internet based samples and with the current MRI syndicated study.