

THE MPA SYNDICATED READERSHIP RESEARCH TASK FORCE PAPER

**Charlene Tretham, Business Week,
Kathi Love, Magazine Publishers of America
Susan Smollens, Consultant**

Overview

This research paper is the result of a process of cooperative work among a diverse group of American research professionals. It is the final product of a complicated, collaborative effort among thirty-odd people from different areas of the advertising world. The project has been a success.

The issue was not methodological, nor a "beauty contest" between research studies. This self-initiated group never intended an endorsement of the "right" service. The mission was to validate the integrity of existing surveys. The approach is from the user's perspective, not a statistician's.

The complete work, a compilation of four separate papers, is now finished and available upon request from the United States' Magazine Publishers of America. There is no consensus of opinion among the members of the Task Force. Controversy still exists regarding some of the recommendations for testing. But this is acceptable: total agreement is probably impossible. However, the group accomplished its mission by working together for the industry and achieving what we have.

The major findings of the group are:

- Although there are some flaws, the existing total adult universe random sampling surveys in the United States are valid within the parameters set forth by the suppliers, which are contained in the technical guides.
- Both major studies are reliable over time, and valid within logical constraints of statistical tolerance.
- All users must accept accountability for the proper use of the published data, and the industry must initiate better education procedures. Users beware!
- Testing variations of survey implementation is recommended in order to improve the quality of the data.

This group overcame the classic premise for controversy: a triangle. Three parties with distinct perspectives and needs (Agencies and Advertisers, Publishers, Suppliers) came together and disproved sensationalism in the trade press. This is a major first step for the Publishing, Advertising and Research industries in America.

Executive Summary

Background

Complaints about syndicated audience research are nothing new. Uneasiness over some aspects of how audiences are measured has been a fact of life for all media.

What is unique is the intensity of the criticisms that members of the publishing community directed toward the industry's print measurement research services, as reported by the trade press in early 1996.

Due to the unusually contentious nature of the debate that has ensued regarding the quality and accuracy of these studies, magazine research directors collectively agreed that it was time to take an active role in addressing many of the issues raised.

In response to the concerns of its membership as well as the industry-at-large, the MPA decided to create a Task Force to serve as a forum for investigating - in a more structured and formal manner - many of the allegations and frustrations expressed by members of the publishing community.

Since the quality of magazine syndicated research is of critical concern to all industry constituencies, membership in the Task Force was expanded to include non-member publishing companies, advertising agencies and advertisers.

Task force mission

The Task Force mission encompassed two major objectives:

1. Improving the quality and reliability of the data produced by the existing services.
2. The exploration of alternative and/or enhanced audience measurement systems.

The participants decided that the Task Force's first objective should be to *take whatever steps were necessary to restore some basic credibility to the current services' data.*

Inherent in this mission statement is an underlying presumption that magazine syndicated research studies provide many benefits to publishers and agencies. Audience research serves as the industry's primary evaluative tool for determining the advertising investment in the print medium; in essence, these studies are the *currency* of the business.

The Task Force's initial goal is to make recommendations to the industry and the research firms that will strengthen that recognized currency.

Traditionally, the research community's instinctive reaction to criticisms leveled at media research suppliers has been to focus on and pursue methodological changes and improvements.

And perhaps in the past, this approach was sufficient to alleviate the crisis of the moment.

However, to attribute the current dispute regarding the magazine audience measurement services solely to methodological shortcomings ignores the broader realities in which today's controversy is playing itself out.

We can no longer afford to ignore the data usage dimension...undeniably the misuse and misinterpretation of the numbers produced by the syndicated services is as much of a contributing factor to the frustration being publicly expressed by members of the publishing community as are perceived methodological shortcomings.

Downsized agency media research departments can no longer monitor the planner's usage of the data, as they have been able to in the past. In addition, the proliferation of the PC, coupled with vast improvements in data access and retrieval, has essentially put the data into the hands of the novice user.

And from their perspective, sample bounce and response rates are arcane and cryptic terms that have little to do with the daily demands of their jobs.

Yes, undoubtedly, increased sample sizes would alleviate some of the variability in the results but it will not eliminate it. Moreover, there are limits to the number of dollars the industry is willing to invest in order to significantly increase sample size.

So before we rush off to pursue the elusive methodological solution to today's magazine measurement crisis..

We need to take a step back and ask ourselves if the inclusive measurement demands we have placed on the syndicated services are contributing to the perceived decline in the quality of the data.

Can we continue to demand that the services measure more and more magazines, products and services and expect no erosion in the quality of the results?

Is a sample of 20,000 sufficient to produce stable and reliable estimates for a host of targeted niche publications and relatively low incidence products and brands? And what happens when these data are analyzed against a highly selected target population?

Can we continue to ignore the fact that the consumer has become increasingly uncooperative and in certain instances almost hostile to the intrusive nature of survey work, but still expect that they will fill out a 96 page product booklet that most of us would describe - at best - as "burdensome".....

And as cooperation declines, the need for ascription increases. We need to understand the impact that increased levels of ascription are having on the data.

Perhaps we could all live with a little less - whether its less measured or less reported - if the tradeoff is - more reliable data.

With these issues in mind, the Task Force established three subcommittees:

Review of Trends and Critical issues such as response rates and ascription

Data Utility to investigate and document the degree of audience variability reported for various targets

AND

Product Information Booklet to critically examine the questionnaire and develop recommendations for testing alternative versions of the questionnaire.

Review of trends and critical issues

Synopsis

Purpose/Goal

- A. To provide a historical background for better understanding of the issues that may be impacting the quality of syndicated research
- B. To determine the overall validity and integrity of the major syndicated readership research studies (MRI and SMRB)
- C. To verify or refute the sensational accusations in the press which claimed that response rates are eroding to critically low levels

The sub-committee's evaluation is based on the following criteria:

- Trend of overall response rate
- Interview response rates among demographic groups
- Product Information Booklet completions by demographic segment
- Ascription levels and impact

Note: Because Ascription is such an important component of this analysis and required in-depth and thorough study, the sub-committee created an ancillary group for that purpose. The Ascription results are contained in a separate paper.

Methods/Process

- A. The primary focus of the subcommittee's analyses were the 1987, 1990, 1993 and 1996 Spring MRI releases.
- B. Respondent counts and rates were analyzed by demographic segments for both the personal interview and the Product Information Booklet.

Note: Both MRI and SMRB data were evaluated. Because of the methodological change implemented by SMRB with the 1995 Study, only 2 years of data were available for analysis and lend little insight in regard to trends.

Findings/Recommendations

- The sub-committee concluded that contrary to press reports and hearsay, the major research studies have very adequate response levels overall, are delivering the data as promised and as described in their respective technical appendices. Additionally, both services have been practicing full disclosure.
- While the personal interview response is satisfactory, there appears to be reason for concern with eroding PIB response in the population at large, and particularly among hard-to-reach upscale segments.

The response rate for personal interviews in the MRI studies has actually increased by almost 3% since 1987. Spring '96 response was 70.2%, up from 68.3% in Spring '87. (SEE EXHIBIT 1)

Personal interview upper income strata respondents still respond at rates below the population at large. However, response within this segment has actually improved relative to 1987. (*Upper Income Spring '87 response rate to the personal interview was 64.8% vs. Spring '96 response rate of 67.3% - an increase of approximately 4%*)
(SEE EXHIBIT 1A)

Response rates have declined 9.4% for the Product Information Booklet over the same period. PIB response was 59.6% in '96, down from 65.8% in '87.
(SEE EXHIBIT 1)

PIB response among upper strata populations has always been and continues to be below that of the balance of the population. However, certain demographic segments - Age 18-24, Employed, and Professionals - have outpaced the general population response declines.

The impact of ascription on MRI marketing data

Synopsis

Purpose/Goal

Response to the Product Information Booklet has declined over time, as clarified by the Review of Trends and Critical Issues sub-committee. Consequently, levels of Ascription have risen accordingly. (*Ascription is the long accepted procedure of estimating response for non-responders by method of demographic matching.*)

- A. To evaluate the impact of increased ascription upon product data, demographic profiles and readership over time.
- B. To assess ascription impact on the media planning and magazine selection process.

Methods/Process

- A. Due to the availability of historical data, MRI's '87, '90, '93 and '96 Spring releases were the basis for this analysis.

Note: With the institution of SMRB's new research technique, ascription is limited and only applied at the brand data level. Therefore, this analysis was not deemed relevant to SMRB. MRI, however, does ascribe projections for non-respondents to the PIB booklet.

- B. Ten different target/product usage definitions were analyzed - ranging from robust segments representing almost half of the population to relatively low incidence targets of less than 15% of the adult universe.
- C. Data provided were respondent counts and audience projections for *Total* respondents, *Pure PIB* respondents and ascribed respondents. Additionally, magazine data were provided for the 10 targets by the three respondent classifications.
- D. Multiple statistical analyses were executed and are provided in detail in the full paper from this sub-committee.

Findings/Recommendations

- Overall, similarities between the *Total* and *Pure PIB* segments are logical given that *Pure PIB* respondents still represent a majority of the *Total* segment. Differences between the *Pure PIB* and Ascribed segments are, in effect, averaged out to arrive at the *Total* segment profile.
- Ascription, in some circumstances, dilutes the magnitude of demographic skews and may narrow the differences between individual demo cells. However, it does not seem to seriously distort directional skews and/or differences.
(SEE EXHIBIT 2)
- Should the incidence of ascribed respondents continue to increase in the future, any differences in readership profiles between the *Pure PIB* and Ascribed segments will more seriously affect the final *Total* profiles.
- While ascription of PIB responses does produce changes in magazine audience profiles for product targets, it has negligible impact on magazine selection decisions. Magazine rankings -- whether based on composition, coverage or CPM measures -- are very similar whether using *Pure PIB* (pre-ascription) or *Total* (post-ascription) respondent data. The top quintile list changes by only 1-3 titles (out of 15-20) when moving between the *Pure PIB* and *Total* segments.
(SEE EXHIBIT 2A)

A data user seeking to identify the key demographic characteristics of a product user target will come to the same conclusions whether looking at *Total* or *Pure PIB* data.(SEE EXHIBIT 2)

Product information booklet analysis

Synopsis

Purpose/Goal

Response to the Product Information Booklet has declined over time, as clarified by the Review of Trends and Critical Issues sub-committee.

The Ascription sub-committee demonstrated that at this point in time, Ascription is not having a negative impact on audience projections nor is it significantly effecting the Media Planning decision-making process. However, any further erosion in response and increase of ascription could have a deleterious impact on the results produced in future studies.

In consideration of the integrity of future studies, as response rates to all forms of research studies decline precipitously, this sub-committee evaluated the length and content of the Product Information Booklet.

The evaluation was guided by the following premise:

“Would reducing the length and depth of the product survey as well as streamlining questionnaire design reduce respondent burden, increase response and consequently reduce ascription?”

- A. To create a measurement tool that will produce more reliable data for all users.
- B. To conduct a thorough and complete study of the two existing PIB surveys from both SMRB and MRI.
- C. To determine levels of detail at which product data are reliable and useful to the majority of data users.
- D. To develop recommendations for various tests which MRI/SMRB may conduct in the interest of reducing booklet length.

Methods/Process

- A. Both the MRI and SMRB national booklets were studied. MRI's booklet is 96 pages; SMRB's is 132 pages. For the respondent completing the booklet can be a time-consuming, burdensome task.
- B. All product categories were analyzed by advertising expenditures to determine magazine sales' potential, as well as degree of relevance to all users.
- C. Product data were studied, as well as the reliability and vitality of brand, forms/types, volume and frequency of use question structure.

Findings/Recommendations

- Product data are often unstable and should be used with caution. However, users are often unaware of limitations.
- There are wide arrays of product categories in each booklet, many of which are only relevant to a very narrow user base. In addition, many of the product categories and brands are outdated i.e., they do not represent the current marketplace.
- There appears to be a lack of consistency in each service's question structure (gateway, volume and brand questions).
- For even commonly used and stable product categories, brand data are very often unusable because of low response rates and instability. (*SEE EXHIBIT 3*)
- The sub-committee recommends multiple tests be initiated as soon as possible and that test results be shared with the user community. Recommended test areas are:
 - Current form
 - Current form, with all brand measures eliminated
 - Reduced brand data, media data detail excluded
 - Variations in questionnaire design

Data Utility

Synopsis

Purpose/Goal

The majority of titles measured by the syndicated services today are publications with circulation's under 1 million. Small circulation titles tend to be targeted to appeal to narrow segments of the population, yet they are incorporated into a broad general population survey along with large circulation, mass appeal titles. (SEE EXHIBIT 4)

Many members of the publishing community have raised concerns regarding the reliability of the estimates produced for these niche or special interest publications. Due to the very nature of these publications, the sample sizes tend to be small and the resultant audience projections subject to - what is commonly referred to - as "bounce" from study to study.

- A. To document the reliability of the audience information reported for selective targets over time.
- B. To determine if the swings in reported audience data would impact Media Planning decisions.

Methods/Process

- A. Six MRI data releases were analyzed in order to determine the stability and tolerance of the reported audience data.

Note: Both MRI and SMRB data were evaluated. Because of the methodological change implemented by SMRB with the 1995 Study, only two releases - Spring and Fall 1996 were available for analysis. However, the data showed a similar pattern to MRI leading to the same overall conclusions regarding utility.

- B. Audience changes and tolerance ranges were summarized by circulation size and percent change for different time spans:

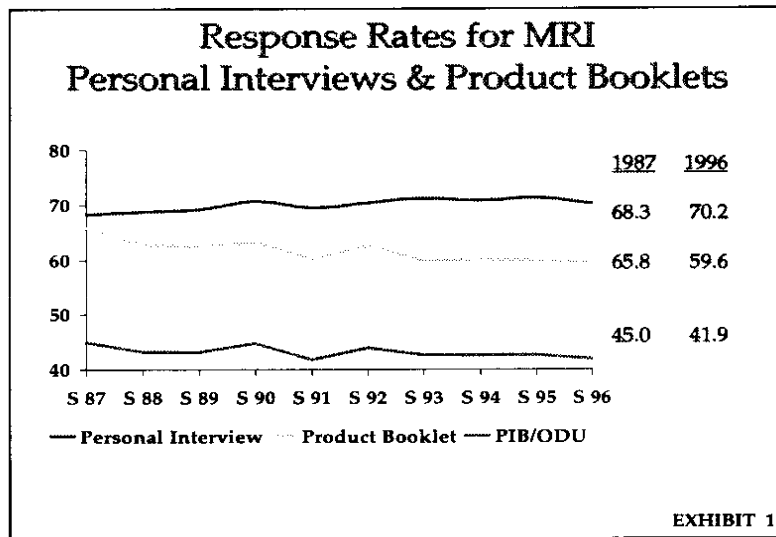
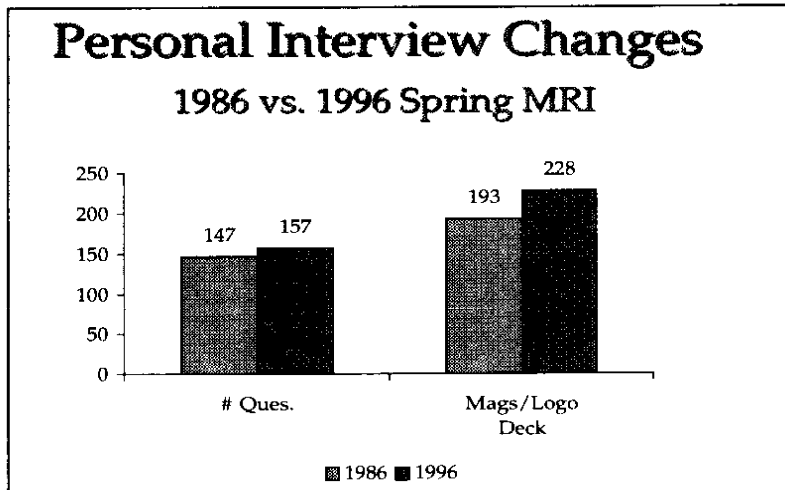
➤ Release to Release	(Spring '96 vs. Fall '95)
➤ Year to Year	(Spring '96 vs. Spring '95)
➤ Release to Doublebase	(Spring '95 vs. Dbase'95)

- C. Magazines were ranked on coverage, composition and cost efficiency for all selected targets - both demographic and product usage.

Findings/Recommendations

- Audience changes or the degree of volatility is greatest on a year-to-year basis.
- Year-to-year comparisons indicate that volatility increases significantly when analyzing low-incidence behaviors or highly selective targets. (SEE EXHIBIT 4A)
- Titles with circulation below 500,000 are subject to high tolerance ranges and dramatic audience swings on both a year-to-year and release-to-release basis. Usage of Doublebase will reduce the degree of "bounce". (SEE EXHIBIT 4B)
- For selective targets, Doublebase will reduce the number of titles with tolerance ranges of more than 20%, especially for magazines with circulation levels in excess of half a million. (SEE EXHIBIT 4C)
- Coverage and CPM rankings were highly consistent across the three time span scenarios. However, composition rankings were much less consistent and - in many instances - would have resulted in different buying/planning decisions, depending on which study was used. (SEE EXHIBIT 4D)
- Printing tolerance ranges on composition ranking reports may help Media Planners make more reasonable judgments of a magazine's performance.
- Establishing minimum reporting standards for all data would reduce the number of titles and targets that are subject to a high degree of volatility.
- Limitations of the data need to be better explained to the general user.

Exhibits



Review of Trends and Critical Issues Response Rates by Demographics

Demographics	Spring 1987	Spring 1996	% Change 96 vs. 86
Adults	68.3	70.2	2.8
Males	62.9	65.4	4.0
Females	74.5	75.5	1.3
Upper Income	64.8	67.3	3.9
Middle Income	69.1	72.3	4.6
Lower Income	75.0	76.3	1.7
PIB, Adults	65.8	59.6	-9.4

EXHIBIT 1A

Review of Trends and Critical Issues

Impact of Ascription: Demographic Profile

Base:	Men			
Target	Car Rental - Business Use Spring '96 MRI			
Incidence:	8.0%			
COMPOSITION INDICES				
	PIB	Ascribe	Total	Difference
	Sample	Sample	Sample	Asc vs PIB
Education:				
Grad College+	269	180	224	-89
Grad H.S.	40	71	55	31
Occupation:				
Professional	270	183	227	-87
Exec/Adm/Mgrl	305	222	263	-83
Not Working	32	27	29	-5
HH Income:				
\$75,000+	269	207	239	-62
\$40-49,999	83	64	74	-19
<\$10,000	16	15	16	-1

EXHIBIT 2

Review of Trends and Critical Issues

Impact of Ascription

Magazine CPM Rankings
Composite Average of 35 product targets

# of Magazines	CPM - Pure PIB					
	Q1	Q2	Q3	Q4	Q5	Total
CPM - Total						
Quintile 1	15	2	0	0	0	17
Quintile 2	2	13	2	0	0	17
Quintile 3	1	3	12	2	0	18
Quintile 4	0	0	3	13	2	18
Quintile 5	0	0	0	2	15	18
Total	18	18	17	17	17	68

EXHIBIT 2A

Breakfast Cereals

108 brands of cold cereal were run on an adult base
for the top 25 magazines/total adults

# Breakfast Cereals	...For Top 25 Magazines
17	All top 25 magazines stable
44	1-12 unstable magazines
47	> 12 magazines unstable
108	Total

Only 28 brands (26%) had an adult base of 5% or more

Source: Spring 1996 MRI

EXHIBIT 3

Data Utility Overview

Circulation Size	SMM 1966 #Titles	Spring 1996 MRI #Titles
< 500,000	8	53
500,000 -999,000	12	57
< 1,000,000	20	110
1,000,000 - 1,999,000	16	48
2,000,000 +	22	29
1,000,000 +	38	77
# Titles Measured	58	187

EXHIBIT 4

Data Utility Volatility: Release to Release / Year to Year

Target Audience	Studies	# of Magazine Titles		
		0-10% Chg.	10-20% Chg.	20% + Chg.
Adults, 100%	Spg 96 vs Fl 95	156	25	6
	Spg 96 vs Spg 95	125	45	17
Women, 52%	Spg 96 vs Fl 95	119	44	24
	Spg 96 vs Spg 95	93	52	42
Prof/Mgr	Spg 96 vs Fl 95	103	38	46
HHI \$50K+, 13%	Spg 96 vs Spg 95	80	47	60

Source: Magazine Audience Changes, MRI Studies
Spg 96, Fall 95, Spg 95

EXHIBIT 4A

Data Utility Tolerance

Circulation Size	All Titles	# Titles > 20% Tolerance	
< 500,000	53	50	
500,000 -999,000	57	47	
< 1,000,000	110	97	88% of Titles
1,000,000 - 1,999,000	48	25	
2,000,000 +	29	2	
All Publications	187	124	66% of Titles

Professional/Managers, HHI\$50K+
Source: Spring 1995 MRI; Target Audience = 12.6% US POP.

EXHIBIT 4B

Data Utility Tolerance

Circulation Size	All Titles	# Titles > 20% Tolerance
< 500,000	53	35
500,000 - 999,000	57	17
< 1,000,000	110	52 47% of Titles
1,000,000 - 1,999,000	48	8
2,000,000 +	29	1
All Publications	187	61 33% of Titles

Professional/Managers, HHI\$50K+
Source: Doublebase 1995 MRI; Target Audience = 12.0% US POP.

EXHIBIT 4C

Data Utility Media Planning Implications

Composition Rankings	R/R S'96/F'95	Yr./Yr. S'96/S'95	DB/Rel. DB'95/S'95
Men (47.9%)			
Top 10	6	5	8
Top 20	15	13	18
W25-54/HHI\$30K+ (20.5%)			
Top 10	9	8	8
Top 20	16	16	17
PM/HHI\$50K+ (13.3%)			
Top 10	6	7	9
Top 20	16	16	18

EXHIBIT 4D

Data Utility Media Planning Implications

Composition Rankings	R/R S'96/F'95	Yr./Yr. S'96/S'95	DB/Rel. DB'95/S'95
Hot Bkfst. Cereal/Past 6 Mnth. Fem. Hmks.(49.6%)			
Top 10	4	1	5
Top 20	10	7	13
Drank Rum/Past 6 Mnth. (8.5%)			
Top 10	2	0	6
Top 20	6	1	13

EXHIBIT 4D