ANATOMY OF A MAGAZINE AUDIENCE ESTIMATE:

- THE ARF COMPARABILITY STUDY REVISITED -

by Valentine Appel Ph.D.
Senior Vice President, Research Director
Backer Spielvogel Bates, Inc.

Draft: 12/28/90

In the United States there are two principal methods for estimating the size and composition of magazine audiences -- one called the Through the Book (TTB) method and employed by the Simmons Market Research Bureau and the other called Recent Reading (RR) and employed by Mediamark Inc. Both methods involve a two step procedure in which probability samples of adults are personally interviewed.

The first step screens a large number of magazine titles by showing each of their logos and asking the respondents to indicate whether or not each magazine might have been read or looked into in the past six months.

In the next step, the few titles which "screen-in" are subjected to a second line of questioning. Although the first step is basically the same for the two methods, the second step is quite different. The TTB method requires that each interviewer carry a kit of suitably aged stripped down issues of each title to be measured. For each title screened-in, the respondents are taken through the stripped issue and asked whether they had read or looked into that particular issue before. Those saying that they were sure that they had seen it before are classified as issue readers. The RR method simply asks whether the respondents had read or looked into any copy of the magazine in the last publishing interval -- past week in the case of weeklies, past month in the case of monthlies, etc.

1

The two methods do not produce the same results. RR estimates are generally higher than TTB estimates, the more so as the publishing interval increases. That was the basic conclusion of the ARF Comparability Study (1980) which was presented by Paul Chook at the first International Symposium in New Orleans, and which has been confirmed by comparison of the separate audience estimates produced by Simmons Market Research Bureau and by Mediamark Research Inc.

The Comparability Study, with an out of pocket budget of 350 thousand 1979 dollars ($350,000) and sponsored by 83 media, agencies, and advertisers, was launched as a result of the industry outcry which resulted following Simmons' announcement that it would increase the number of magazines measured in a single interview by using what later came to be called the mixed method: the smaller audience monthlies were to be measured using the RR method while the larger audience monthlies and the weeklies would continue to be measured using the TTB method.

The Basic Design

The ARF Study was conducted using three probability samples of respondents as shown in Exhibit #1: Sample I, consisting of approximately 1000 respondents, was used to measure the audiences of 68 titles using only the TTB method. Samples II and III, consisting of approximately 1800 respondents each, were employed for purposes of measuring 124

Exhibit #1

EXPERIMENTAL DESIGN

(Sample Size)	Three Independent Samples I (1084) %	II (1858) %	III (1796) %
All Titles Measured	68	124	124
TTB Measured Titles	68	34	34
Monthlies	54	27	27
Weeklies	12	6	6
Tri Weeklies	2	1	1
RR Measured Monthlies	0	90	90
TTB in other sample	0	34	34
Not measured TTB	0	56	56

2-a

titles: 34 using the TTB method followed by the measurement of an additional 90 using the RR method; 34 of the 90 titles which were measured using the RR method were the same as had been measured by the other of the two samples using the TTB method. The remaining 56 RR measured titles were the same in Samples II and III and were used as filler to approximate the length of the Simmons mixed method interview. The measurement of the 68 titles in Sample I was referred to as the TTB-T method (T for traditional), while the method employed in Samples II and II was referred to as TTB-M (M for mixed).

The visual stimuli used for the TTB screening were a series of life-sized, four color logos arranged both in alphabetical and reverse alphabetical order, placed in a spiral bound binder and presented three to a page, 22 pages in all. Order of presentation of the pages was appropriately rotated, with two additional pages, one containing the logo for the New York Times Magazine and the other containing The Star's logo, always being presented last.

All 24 pages were presented to the TTB-T sample. The odd numbered pages were presented to Sample II and the even numbered pages were presented to Sample III. The net result was that the TTB-T sample was required to screen twice as many titles as were each of the two TTB-M samples.

Following the TTB-M portion of the interview, the respondents in Samples II and III were presented with a deck of 90 miniature-sized, appropriately rotated, four color logo

cards, one logo per card, taken thorough the screening process, and then subjected to the RR line of questioning.

The ARF Study represents a significant achievement, not only because of the understanding it provided regarding the performance of the TTB and RR methods, but also by virtue of the rich source of data it supplied and which is still available in the form of volumes of tabulations and a computer tape waiting further analysis.

The purpose of this paper is to report the results of one such analysis recently completed by the author based upon data which had already been tabulated.

Excluded from the analysis are Family Circle and Woman's Day. These two titles were excluded because of their unusual publication frequency and the fact that having only two titles precluded reliable statistical generalizations. A third title, Travel and Leisure, was excluded following the discovery of a field administrative error involving the improper use of local editions in the TTB interview. The analysis to be presented, therefore, is based on a total of 53 monthly and 12 weekly titles.

Two Principal Components

The appearance of the ARF Study was the first time that data had been published in the U.S. which permitted the decomposition of the TTB and RR audience estimates into their

two principal components: (1) the screen-in level and (2) the read/screen ratio, i.e. the proportion of those claiming to have read or looked into a given magazine in the past six months who went on to qualify as average issue readers.

The audience estimate is mathematically equivalent to the product of these two numbers, although in actual practice the multiplication is not literally performed. Rather, respondents are classified as readers of a given magazine if they "screen-in" and go on to answer the RR/TTB question with affirmative certainty.

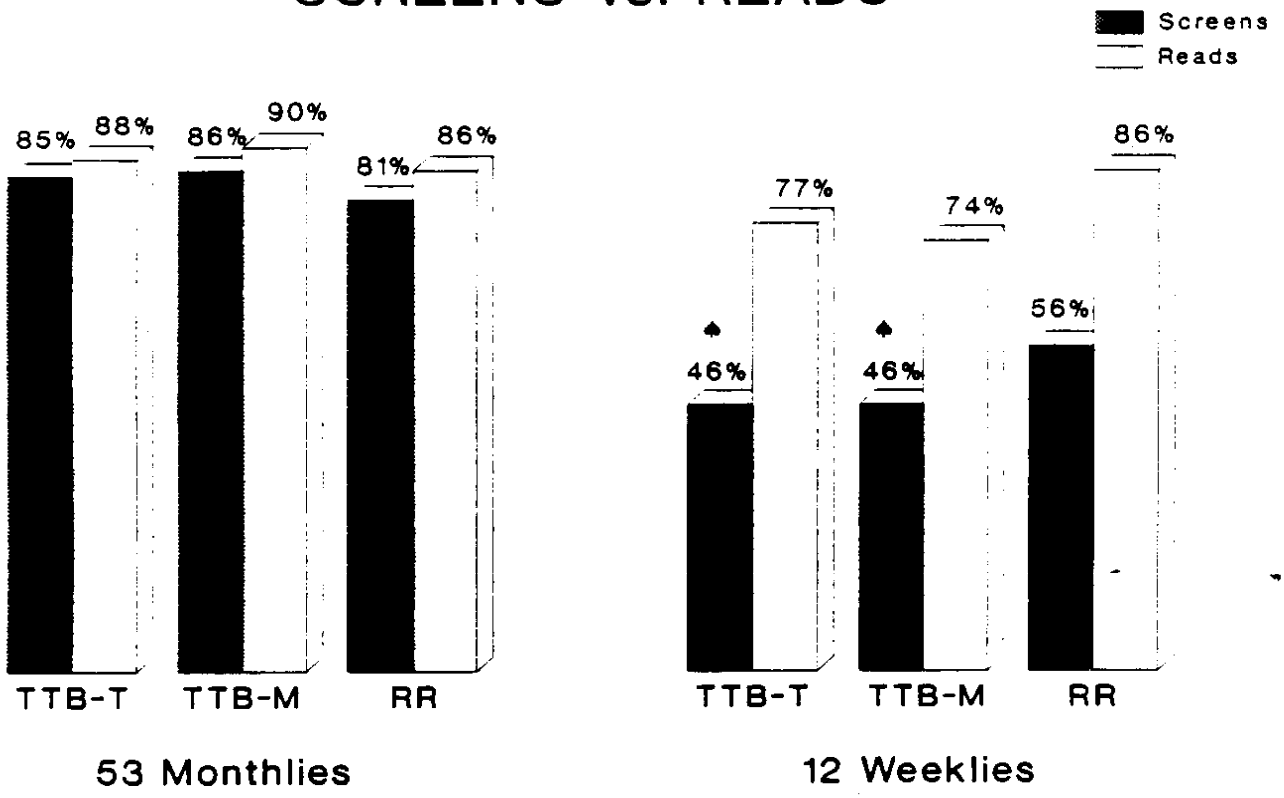
Circulation Effects

Exhibit #2 summarizes the percent of variance in the screen-in levels and in the estimates of average issue audience (reads) which is accounted for by the circulation of the magazines. Note that throughout the presentation the data for the 53 monthlies will be shown on the left for each of the three measurement methods, and the data for the 12 weeklies will be shown on the right.

The percentages at the top of each of the bars now on the screen are the squares of the correlations with circulation multiplied by 100. The slide confirms that the read levels are strongly accounted for by circulation with upwards of 74% of the variance being so explained. There are no significant differences between the heights of any of six bars showing the relationship between circulation and the read levels.

Exhibit #2

VARIANCE ACCOUNTED FOR BY CIRCULATION SCREENS vs. READS



♠ Significantly different from the monthlies

5-a

Looking at the screen-ins, however, we find a decidedly different picture: for the monthlies circulation accounts for about as much variance in the screens as it does the reads. In the case of the weeklies, however, circulation is much less important. For the two TTB methods, the differences between the monthlies and weeklies are statistically significant.[1] Circulation is clearly more important in explaining the variation in the screen levels for monthlies than it is for weeklies.

Because of these relationships with circulation, before proceeding further it was thought best to remove it as a source of unwanted variance. To that end, the coverage percentages reported in the ARF's report were first multiplied by the size of the U.S. adult population to provide a projected average issue audience and a projected number of screen-ins. Each of these audience projections was then divided by the appropriate circulation for each magazine to produce (1) a readers per copy estimate (RPC) and (2) an estimate of screeners per copy (SPC). This procedure was followed for each magazine for each of the three measurement methods.

---

[1]Throughout the paper the correlation coefficients have been squared and expressed in terms of percent of variance explained. Tests of statistical significance were calculated using Fisher's Z transformation. Without exception, for the means as well as the correlation coefficients, all significant differences meet the .05 level of confidence or better.

Readers Per Copy (RPC)

Let's first compare the mean readers per copy estimate produced by the three methods. The pattern of readers per copy estimates, which are shown in Exhibit #3, are essentially similar to the pattern of audience estimates some of you saw in New Orleans 11 years ago. The reason they are not identical is because what was shown in New Orleans were mean coverage percentages uncorrected for differences in circulation.

Note that in the case of the monthlies, the RR method produced readers per copy estimates about twice the size of that produced by the TTB methods. And in the case of the weeklies, the RR method produced audience estimates which were about 25% larger than those produced TTB. Both of these differences are statistically significant.[2]

Note also that, although the difference is not statistically significant, the TTB-M RPC estimate for the monthlies is somewhat higher than that achieved using the TTB-T method.

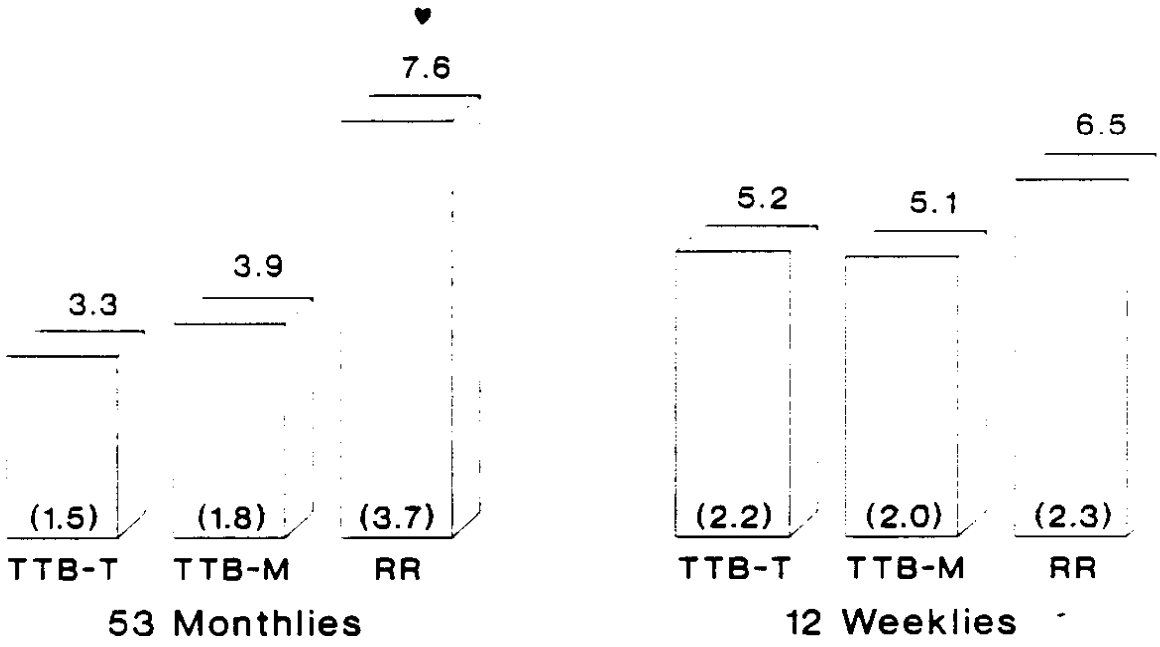
Now, let's explore how these differences came about.

---

[2]Throughout the paper the significance of the difference between means was calculated using six replicated sub samples, and entering the t-table with five degrees of freedom. The information necessary to do this was taken from tabulations made available by the ARF.

Exhibit #3

MEAN READERS PER COPY*



♥ Significantly different from both TTB methods

* Standard deviations shown in parentheses ( )

7-a

Screeners Per Copy (SPC)

Turning first to the screeners per copy data, the means and standard deviations of which are shown in Exhibit #4, note that for both the monthly and the weekly magazines the TTB-M samples produced higher screen-in levels than did the TTB-T sample. Note also that for the monthly magazines the RR sample produced the highest screen-in level of the three. All of these differences are statistically significant.

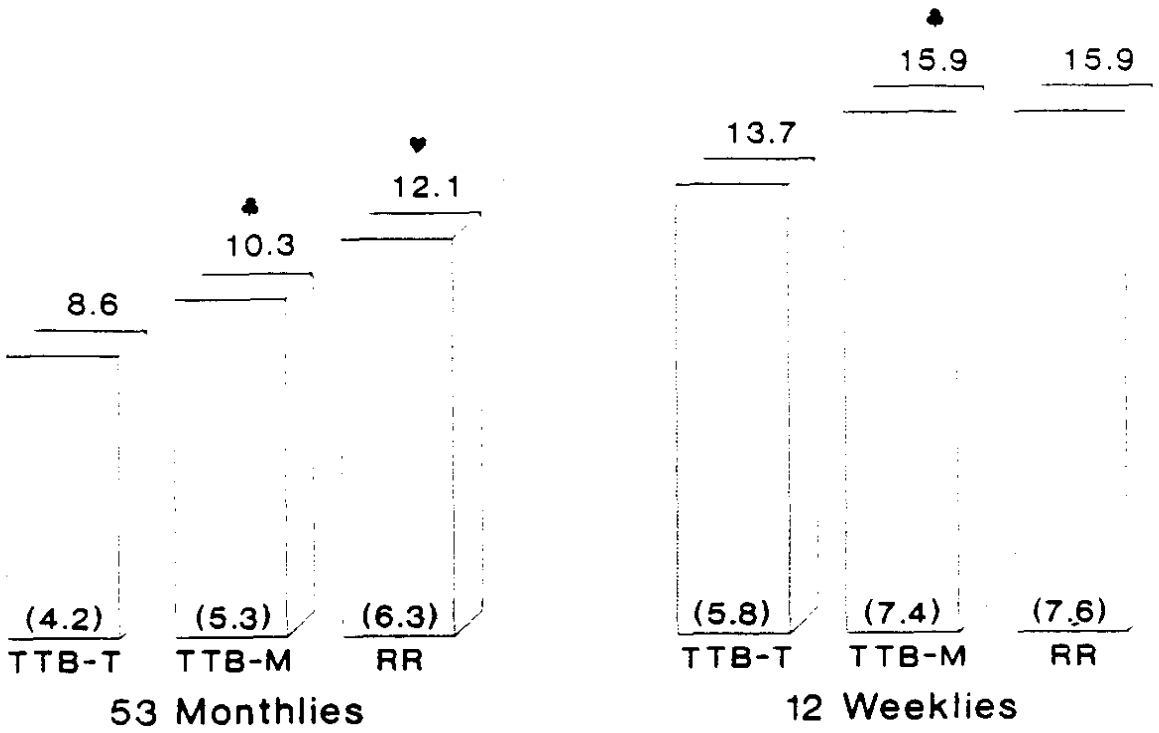
But whether these differences were caused by differences in numbers and/or mix of titles screened, differences in format and/or presentation of logo cards, placement in the interview or some combination of these is unclear. What is clear, however, is that seemingly minor changes in screening procedure can and does produce materially different results.

Read/Screen Ratios

The means and standard deviations of the distribution of read/screen ratios are shown in Exhibit #5, from which it is clear that the read/screen ratios generated by the RR procedure was much higher than those produced by the two TTB methods, the more so in the case of the monthlies where both differences between the RR and TTB estimates were statistically significant. For the weeklies the difference between the RR and TTB-M method and the difference between the TTB-T and RR method were also statistically significant.

Exhibit #4

MEAN SCREENERS PER COPY*



♥ Significantly different from both TTB methods

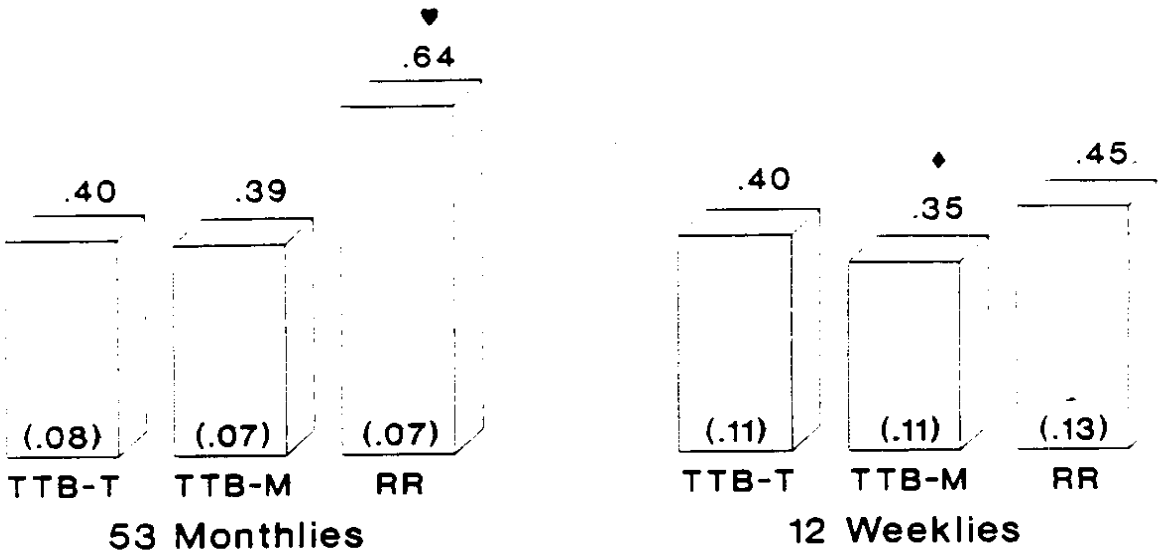
♣ Significantly different from TTB-T

* Standard deviations shown in parentheses ( )

8-a

Exhibit #5

MEAN READ/SCREEN RATIOS*



♥ Significantly different from both TTB methods

♦ Significantly different from both TTB-T and RR

* Standard deviations shown in parentheses ( )

8-b

R/S vs. RPC Correlations

In order completely to understand the interaction of the R/S ratio with the screeners per copy in the production of a readers per copy estimate, it is necessary to know the correlation between these two variables.

These data are shown in Exhibit #6 again expressed as the squares of the coefficients multiplied by 100. Unlike the other exhibits, the heights of these bars are shown upside down to signify the fact that all of the correlations are negative, significantly so except for the TTB-T method.

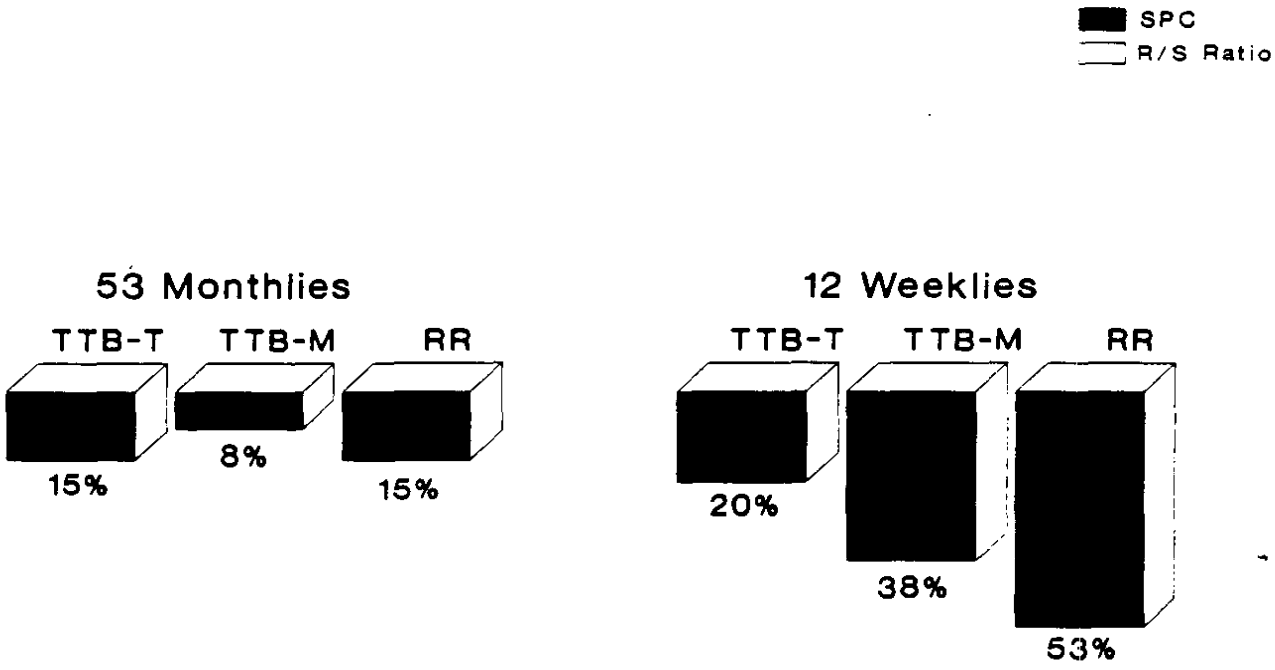
Note also that for each of the three measures the negative correlation is stronger for the weeklies than it is for the monthlies. Although none of these three differences is statistically significant in and of itself, the average of the differences is significant. It would appear, therefore, that the size of the downward adjustment to the SPC which the R/S ratio provides becomes smaller as the SPC becomes smaller, and that this is particularly true for the weekly magazines.

Compensating Effects

Now, given the fact presented earlier that the TTB-M method produced significantly higher mean screen-in levels than did the TTB-T method, for the TTB procedure to provide the same estimate of the size of the average issue audience, there has to be a reversal in the read/screen ratios.

Exhibit #6

VARIANCE EXPLAINED
SPC vs. R/S RATIOS



9-a

Such was the case for the weeklies, as can be seen in Exhibit #7, where the difference in SPC was almost completely compensated for by a statistically significant reversal in the R/S ratios and virtually identical RPC levels.

For the monthlies, however (shown in Exhibit #8), the significant difference in screen levels was not compensated for by a reversal in the R/S ratio resulting in a somewhat higher through not significantly different RPC estimate having been generated by the TTB-M method.

Coefficients of Variation

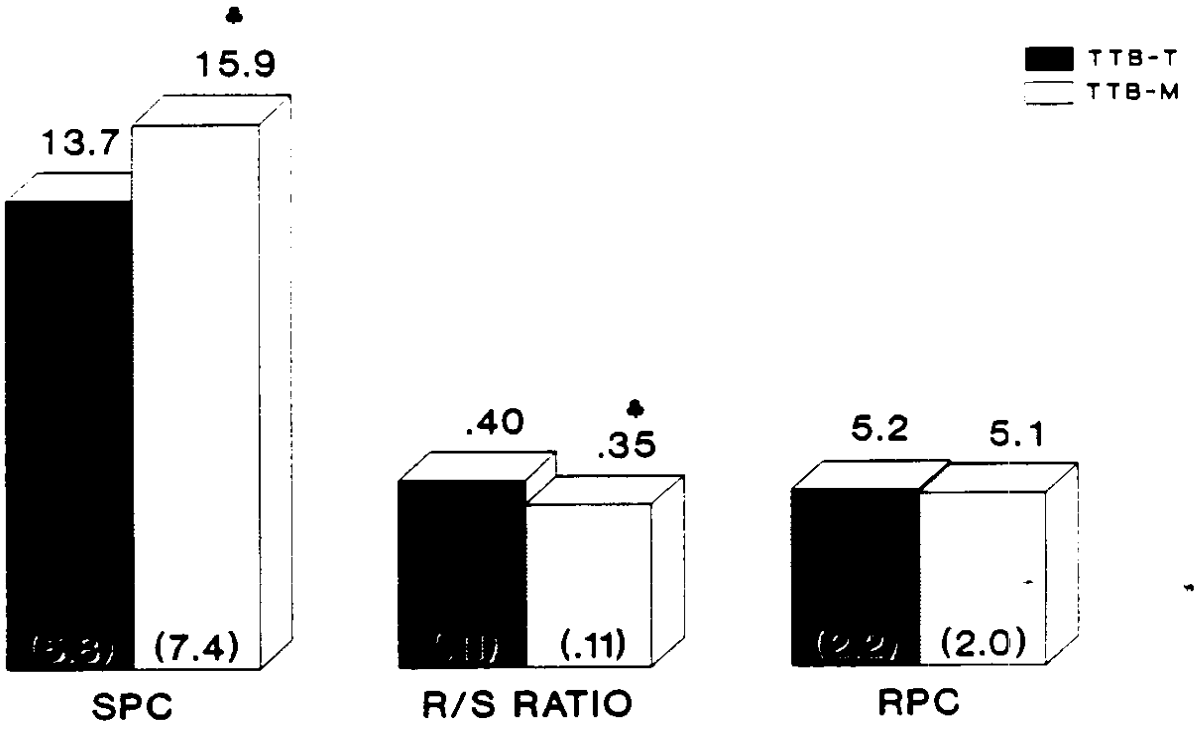
To this point we have been concentrating on the mean screeners per copy and read/screen ratios produced by the three methods. The standard deviations of the distributions of the monthlies and weeklies were also shown but not commented upon.

Exhibit #9 presents the coefficients of variation of these measures calculated by dividing the standard deviation of each distribution by its mean. The screeners per copy are shown in the left hand member of each pair; the data for the read/screen rations are shown on the right.

Looking now to the relative sizes of the coefficients of variation, note that while the screeners per copy uniformly show more variation between magazines than do the read/screen ratios, these differences are much more pronounced for the monthlies than for the weeklies.

Exhibit #7

WEEKLY MAGAZINE MEANS*



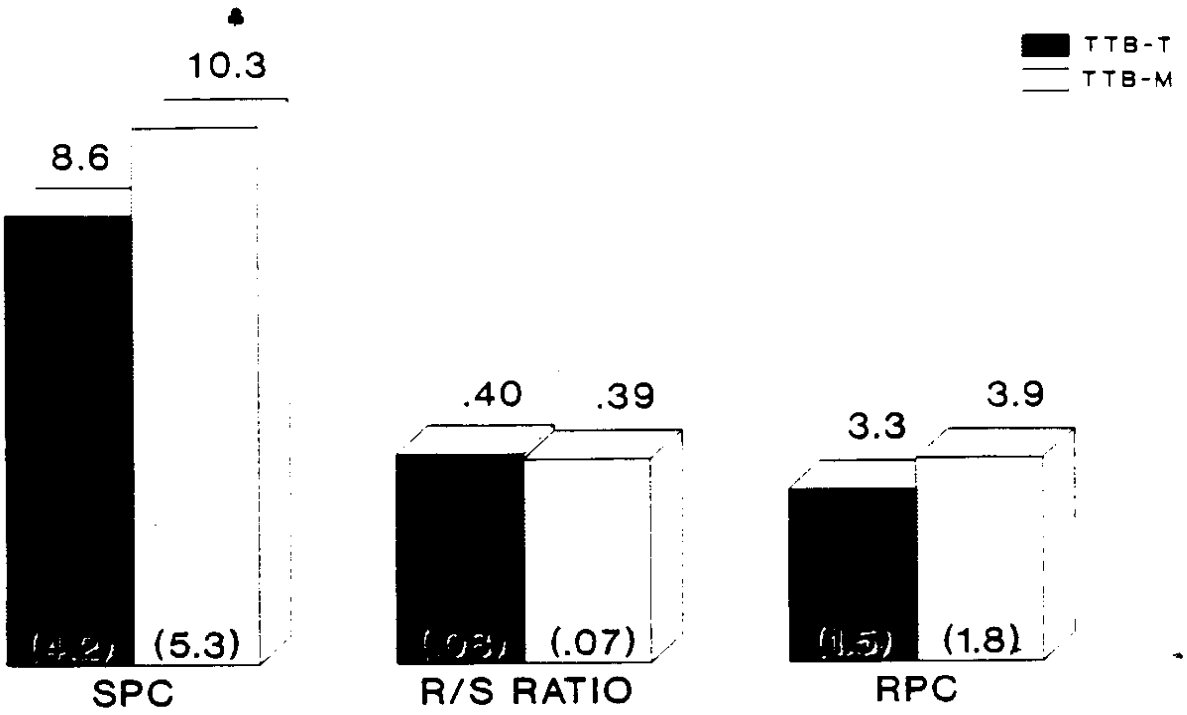
♣ Significantly different from TTB-T

* Standard deviations shown in parentheses ( )

10-a

Exhibit #8

MONTHLY MAGAZINE MEANS*



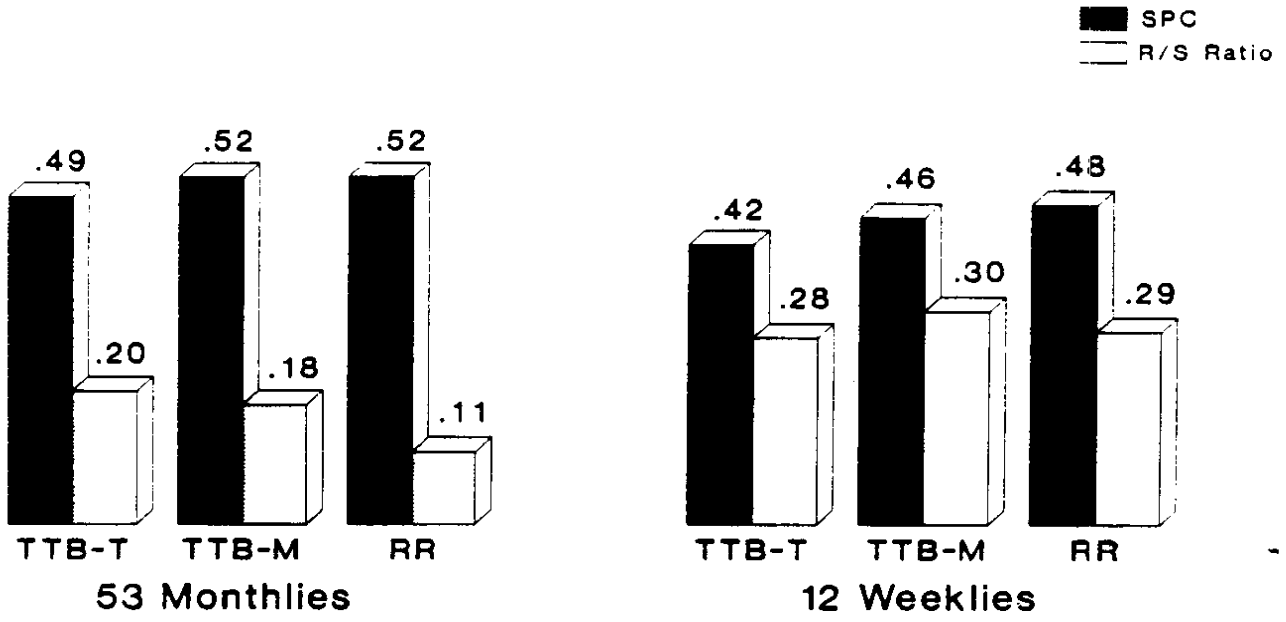
♣ Significantly different from TTB-T

* Standard deviations shown in parentheses ( )

10-b

Exhibit #9

COEFFICIENTS OF VARIATION SPC AND R/S RATIOS



10-c

These differences in coefficients of variation and the differences in correlation shown earlier are extremely important because they entirely determine the extent to which the variation in readers per copy is determined by the two components of that statistic: the screeners per copy and the read/screen ratio.[3]

Based on these data, using statistical theory (Snedecor, 1946), we would predict: (1) that the screeners per copy would explain more of the readers per copy variance than would the read/screen ratios, and (2) that this relationship would be stronger for the monthlies than for the weeklies.

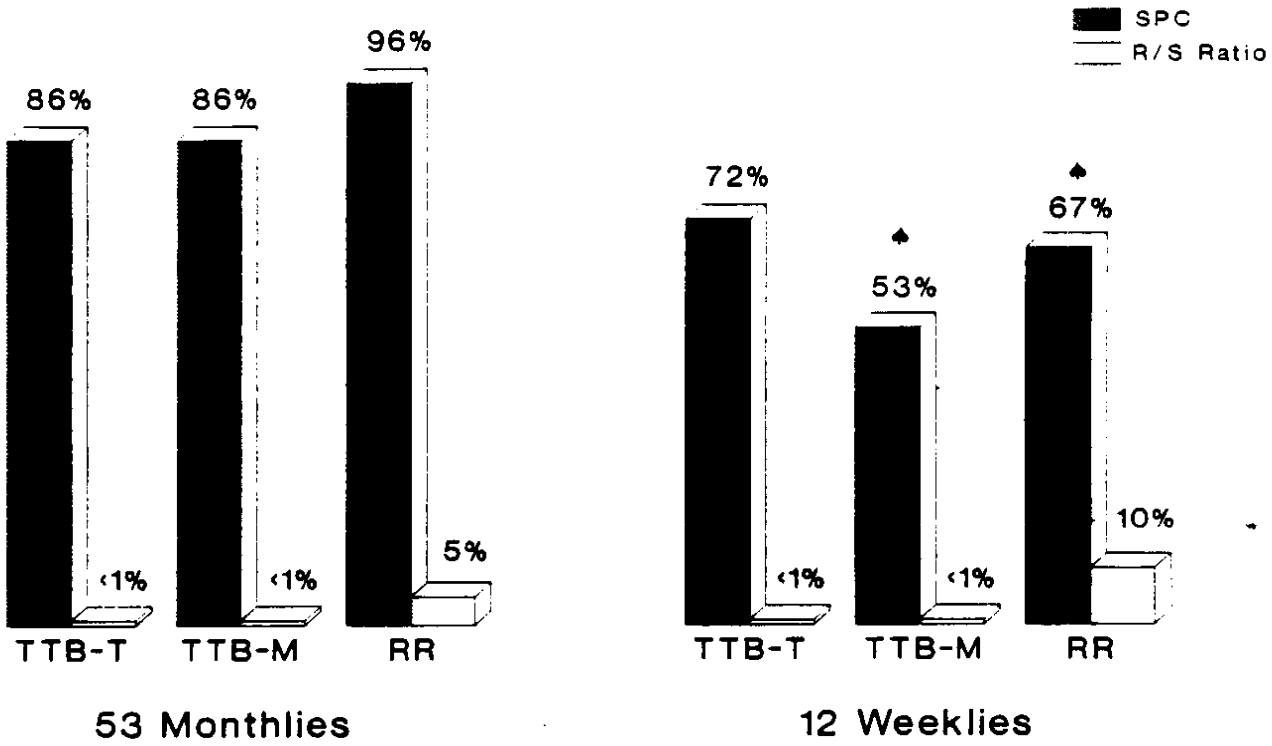
Explained RPC Variance

The accuracy of this prediction can be seen from Exhibit #10 which presents the squares of the correlation coefficients between the readers per copy estimates vs. the screeners per copy and the read/screen ratios. In the case of the monthlies, the overwhelming proportion of the variance is accounted for by the screeners per copy, and in the case of the weeklies a lesser but still substantial majority of the

[3]When one variable is defined as the product of two others, i.e., A x B = C, the correlations of A and B each with C are entirely determined by the coefficients of variation (standard deviation ÷ mean) of A and B and the correlation between them. The larger the difference in the coefficients of variation, and the more the correlation between the two variables departs from +1.00, the larger will be the difference in the correlation between A vs. C and B vs. C.

Exhibit #10

SOURCES OF RPC VARIANCE
SPC vs. R/S RATIOS



♠ Significantly different from the monthlies

11-a

variance is so explained. In two of the three instances -- TTB-M and RR -- the differences between the weeklies and the monthlies are statistically significant. Virtually none of the variance is explained by the read/screen ratios alone. And for those who may have noticed that the two sources of variance do not sum to 100%, the reason is that the strong negative correlation between them has not been taken into account.

Discussion

After controlling for circulation and publishing interval, most of the variance in audience size is determined by the screen-in levels, with the variation in the R/S ratios having a decidedly lesser effect. For monthly publications, variation in the screen-in levels all but overwhelms the variation in the R/S ratios.

The read/screen ratio appears to be a more important determiner of audience size for weeklies than it is for monthlies. Very likely this difference is attributable to the fact someone saying that they might have read a given monthly in the last six months has a higher likelihood (1/6) of having read or looked into a particular recent issue than does someone giving the same response for a weekly (1/26).

One must wonder, therefore, why so much attention has been paid to the different methods for establishing the R/S ratio -- witness the never ending TTB vs. RR controversy -- and why so little attention has been paid to the screening process.

The effect upon audience size of number of titles screened, the physical proximity of the logos of titles likely to be confused, order effects, etc. have been largely neglected topics in the U.S. The effect of number of titles screened is particularly important given the industry's ravenous appetite for having more and more measures for more and more magazines.

Instead, the industry continues to concern itself with such matters as the relative validity of the TTB and RR methods, the increased weight of the TTB interview magazine kit if the number of screened titles increases, and the consequences of reducing the number of articles exposed in the TTB interview in an effort to reduce the kits' bulk. All of this to improve the validity of that portion of the interview which contributes least to the determination of audience size.

REFERENCES

ARF Comparability Study: A Controlled Experiment Comparing Three Methods of Estimating Magazine Audience. Advertising Research Foundation, New York. The first of two reports, January 1980.

ARF Comparability Study on Methods of Measuring Magazine Audiences. Advertising Research Foundation, New York. Report Number Two, December 1980.

Chook, P.H. "The ARF Comparability Study," Readership Research: Theory and Practice. Proceedings of the First International Symposium, New Orleans, 1981, pp. 122-127.

Snedecor, G.W. Statistical Methods. The Iowa State College Press, Ames, Iowa. Fourth Edition, 1946, pp. 162-164.