Timothy Joyce
Mediamark Research Inc
New York, USA

The two syndicated magazine audience research services in the United States, run by MRI (Mediamark Research Inc.) and by SMRB (Simmons Market Research Bureau), report audience levels for magazines which differ considerably.

The first stage in audience measurement with MRI is a six-month screen, conducted by the sorting of logo cards by the respondents. The screened-in cards are separated into weeklies, monthlies, etc. They are then re-sorted by the respondent on sort boards appropriate to each magazine frequency. Those who place the cards in the 'yes-sure have' position for having read in the last seven days for a weekly, 30 days for a monthly etc are counted as readers. The MRI technique is therefore Recent Reading (RR), though with procedures developed in the United States and not exactly the same as those used in any previous survey.

The first stage with SMRB is essentially the same as MRI: a six-month screen conducted by logo card sorting. However, the audience measurement is based on specific test issue recognition. Issues of each screened-in title are shown which average in age, from the on-sale date, five-six weeks for weeklies and 10-12 weeks for monthlies. The test issues are skeletonised, nine editorial items being retained per issue. Respondents are asked which items they would find especially interesting - and then, whether they had read the issue before. Those saying 'sure have' are counted as readers.

The SMRB technique is therefore Through-the-Book (TTB), though full magazine issues are not used. Obviously they could not be, since SMRB studies 100 magazines in total, and the kit would be unmanageable for personal interviews and the interview impossibly long if they were not severely stripped down.

MRI reports audience levels which are approximately 20% higher for weeklies, and 40% higher for monthlies, than SMRB levels.

Apart from the weekly/monthly difference, in *relative* terms the audience levels are not too different. A number of independent evaluations of the services have reached this conclusion, based on comparisons for all adults, men and women, and key demographic target groups. Magazine selections based on the services therefore tend to be very similar.

However, the substantial overall differences remain. If one or the other service is distorting audiences for some reason, it could be introducing relative (even if generally slight) distortions between specific magazines. Further, in inter-media comparisons, the overall levels do matter.

THE DIFFERENCES

There should be nothing alarming or even surprising about these differences to those who attended the previous Symposia in New Orleans or Montreal, or read the papers delivered at them. It is now well understood that different measurement techniques yield different results. Further, RR and TTB themselves are not monolithic, since variations in execution even with one system lead to variations in levels which may be extreme. The screening procedure employed, the precise wording of the readership question, the number of response categories etc are all of importance.

MRI and SMRB, however, both use essentially the same screening question. Both acknowledge the same definition of a reader, and use the same question language (ie "did you *read or look into*"). Both effectively use three response categories for readership qualification - sure did, not sure, sure did not - and only count

as members of the audience those who are sure they did read or look into.

One might conclude, therefore, that the differences between MRI and SMRB audience levels are primarily due to the basic difference between the RR and TTB approaches - provided it is understood that TTB as executed by SMRB involves skeletonised, not full, test issues. What would cause these differences in results between RR and TTB? Obviously this question receives two sets of answers, depending on one's viewpoint.

CRITICISMS OF TTB

The criticisms of TTB are twofold. First, in consequence of the need to measure a large number of titles, it is generally carried out with skeletonised test issues: as with SMRB, which uses nine-item issues. But these retain only a fraction of the full editorial matter in the actual issues, and they do not *look* like magazines. Perhaps some readers, especially casual readers, fail to recognise them: the interview may, after all, take place several weeks after the reading event.

There is indeed evidence that *full* Through-the-Book yields higher audience levels than skeletonised issue TTB. In 1975, the last published study to employ full TTB in the United States was carried out. (A 15-magazine study carried out by Simmons in the following year omitted regular editorial features.) This 1975 study was carried out by Audits & Surveys for Time Inc, and surveyed ten magazines. It obtained audience levels substantially higher than the contemporary Simmons study, and than the current SMRB study. The obvious reason was the use of full rather than skeletonised issues. The additional readers found, in net terms, were all out-of-home (ie on average more casual) readers.

The second criticism of TTB applies to *any* form of TTB and, indeed, to *any* 'issue-specific' recognition

measurement. It has to do with the age of the test issues.

It is reasonable that given their topical nature, the audiences of weeklies may build fully over five-six weeks at most, and that test issues of this age should not encounter appreciable memory loss. But monthlies are not as topical, and it is a common experience that copies of them can stay in public places, or be passed on from home to home, over many months. The issue age of 10-12 weeks is a compromise - probably losing some readers due to forgetting, and others simply because they have not read it yet. The supposition is that total losses are minimised at that age - a supposition not checked empirically for many years. In any case, the issue age problem could account for understatement of monthlies relative to weeklies. TTB was, after all, originally developed for weeklies and bi-weeklies - *Life* and its competitors.

CRITICISMS OF RR

From the standpoint of someone favouring TTB as conducted by SMRB, RR must of course be supposed to overstate audiences. The main reasons why it could do so which have been put forward have again been two.

First, it is well known that replicated reading can cause RR to overstate, although on the other hand parallel reading can cause it to understate. Perhaps in net terms, errors due to replicated reading outweigh errors due to parallel reading.

Some years ago, when RR was first introduced to syndicated research in the United States, this criticism was common. It is heard less now, perhaps due to a realisation that errors due to replication must be quite limited. They can only arise with irregular readers (since regular readers will be correctly counted anyway) - and irregular readers tend to read issues on only one day, ie tend not to replicate.

A more popular criticism now is the second, which is that RR audiences are inflated by `telescoping', ie a tendency on the part of respondents to bring remembered events forward in time. On this model, the audiences of weeklies would be overstated to some extent and of monthlies to a greater extent. This case was advanced by Dr Valentine Appel of SMRB in papers at both previous Symposia. I shall only comment that no *direct* evidence has so far been presented that RR as conducted by MRI leads to net error due to telescoping.

THE SCREEN

These are the reasons suggested why, on the one hand, TTB may understate or on the other hand, RR may overstate. There is however a difference between MRI and SMRB which does not have inherently anything to do with RR versus TTB. The *screening* levels are not the same.

Both surveys use essentially the same screen, ie a six-month screen by logo card sorting. The SMRB-type screen levels have been made public twice. The 1979 ARF comparability study (reported on by Paul Chook at New Orleans) used this type of screen following the then SMRB practice, before RR questions. These screen levels were virtually identical with the current MRI levels (actually 1% higher).

In 1983, the Media Research Committee of the American Association of Advertising Agencies (AAAA) asked both services to supply their screen levels. Both did so, and it emerged that the SMRB levels were in the aggregate 20% *below* MRI's.

Given the earlier comparison where the levels were the same, it seems unlikely that the current difference is due to the slight differences between the screen procedures (eg SMRB uses a two-position sort board, MRI a three-position board). It seems more

likely that it is due to the fact that TTB questions now follow the SMRB screen, while RR questions follow the MRI screen.

It has been conjectured that this could be a result of the mental set of the SMRB interviewers leading unconsciously to their putting the question in a way which would discourage heavy screening-in, since with the TTB double-interview procedure this must lead to *two* lengthy interviews.

In any case, lower screening levels may be part of the explanation for lower reading levels. Someone who has not screened in cannot subsequently qualify as a reader.

THE COMPARATIVE STUDY

With this as background, it seemed that it would be very interesting to carry out a methodological study with the following features:

(1) RR and TTB would be compared starting with the *same* screening procedure and levels.

(2) The test issues would be *full* TTB issues, not skeletonised.

(3) TTB test issues would be studied at different ages.

This study would remove possible differences due to screening, and would compare RR with full TTB. It would also assist a judgement as to how far any differences which did emerge could be due to the issue age problem.

501 adults were interviewed in May 1985, in five interviewing facilities in shopping malls across the United States. In these facilities, complete kits of full TTB test issues could be on hand, which would be impossible with in-home interviews.

The interviews began with a six-month logo card screen with the MRI procedure, with 101 titles out of the

110 currently measured by SMRB (bi-monthlies, cable TV guides, and titles not also measured by MRI were dropped). Following the screening of all 101 titles on the same lines, reader qualification was achieved by TTB for some of the titles, and by RR for the remainder of the titles in each interview.

In half the interviews, the monthly magazines were measured by TTB and the other publications by RR; in the other half of the interviews, the monthly magazines were measured by RR and the other publications by TTB. In half of the interviews, the TTB measurement was first; in the other half, the RR measurement was first. In half of the questionnaires, magazines were printed in alphabetical order, and in the other half, in reverse alphabetical order. There were thus eight versions of the questionnaire, used in rotation.

Questions other than magazine questions were confined to major demographics. Interviews took an average of 40-45 minutes. Respondents were paid for their participation.

The RR measurement was exactly the same as MRI, ie within each publication frequency group, logo cards for the screened-in titles were re-sorted by the respondent using sort boards specific to each group (eg last seven days for weeklies, last 30 days for monthlies).

Publications were prepared for full TTB measurement on conventional lines. Facing pages of ads and editorial continuation material were stapled together. The opening page or pages of each editorial item was, however, displayed to the respondent. Editorial items were numbered. Items averaged 27 across all magazines, with a maximum of 65. After conventional editorial interest questioning, the SMRB reader qualification question was asked.

Test issues were used at five different ages: almost two, four, six, eight, and ten weeks old for weeklies except

TV Guide; almost one, two, three, four and five months old for monthlies; and correspondingly for bi-weeklies and tri-weeklies. Thus, the monthly test issues were the January, February, March, April and May issues in each case, with the fieldwork in early May. The middle age issues were equal in age to the average of SMRB test issues; two issues being younger than this, and two being older.

Rotation of age of test issue was achieved by sending different ages of each publication to each interviewing location, in rotation. Note therefore that *each* respondent was subject to both RR and TTB measurement; *and* (if they screened in) would be subject to TTB measurement with issues of all ages. Those who said they were *sure* they had read or looked into, within the publication-interval in the case of RR, and the specific issue in the case of TTB, were counted as readers.

Three points may be noted about the presentation of the results of this study.

First, the study was not designed to report individual titles: it was designed to report aggregates. However, for obvious reasons, weeklies and monthlies were to be reported separately.

Second, a decision had to be made about the reporting of the four bi-weeklies and the two tri-weeklies (the latter being the 'store books', *Family Circle* and *Woman's Day*). Since the former compete primarily with weeklies and the latter with monthlies, it was decided to pool their data with the 14 weeklies and the 81 monthlies, respectively. In what follows, therefore 'weeklies' embrace bi-weeklies; and 'monthlies' embrace tri-weeklies (as one might say, 'men' embrace women). The great majority of cases, of course, *are* weeklies or monthlies.

Third, the common starting point for both RR and TTB measurement was the screen. This was deliberate, given the

desire to compare the RR and full TTB methodologies per se. The best way of examining the data therefore is in terms of 'read/screen ratios'. These are quite simply the proportions of screeners-in who read. Thus if half those who screen-in qualify as readers, the ratio is 0.50. The use of read/screen ratios removes the sampling error specific to the screen, and is equivalent to the ratio estimation of directly comparable RR and TTB levels.

RESULTS - WEEKLIES

The aggregate read/screen ratios for weeklies, with RR and TTB across all issue ages, were as follows:

RR	.46
TTB	.47

Recent reading and full Through-the-Book, then, yield the same levels for weeklies when the screen starting point is the same.

TTB levels by issue age were as follows:

2 weeks	.47
4 weeks	.47
6 weeks	.49
8 weeks	.45
10 weeks	.47

(These are approximate - rounded up - test issue ages for all weeklies except *TV Guide.*)

It appears that weeklies accumulate essentially all their eventual audiences in a week or two, and issue age is therefore not a factor in distorting audience levels.

The fact that full TTB and RR levels for weeklies agree implies that the lower levels obtained by SMRB are a result of the latter's use of skeletonised issues rather than full issues. This is borne out by the comparison between the 1975 Audits & Surveys ten-magazine study and the contemporary Simmons levels.

RESULTS - MONTHLIES

The case is different for monthlies. The aggregate RR and TTB read/screen ratios are as follows:

RR	.55
TTB	.41

However, the variation of TTB levels by issue age is most interesting:

1 month	.38
2 months	.38
3 months	.42
4 months	.45
5 months	.43

(These are approximate test issue ages - rounded up- for all monthlies. Bases for the ratios average about 750 unweighted counts.)

It appears that TTB levels increase with increasing issue age up to about four months, and then decline.

In actual fact, of course, the levels cannot decline because a magazine issue cannot lose part of its audience once it has obtained it. But the declining audience phenomenon has been found in previous studies (Joyce, 1982). In the past it has been attributed to *forgetting.*

The data just reported suggest that TTB - or more generally, any issue-specific method - is incapable of measuring the audiences of monthlies without substantial understatement. Their audiences build over a number of months; but to use test issues which are effectively at the end of their 'lives' will incur substantial memory loss.

The peak with full TTB is at four months - older than the usual TTB practice - with a ratio of .45. At .55, the RR ratio is still 22% higher. It seems quite possible that this difference is accounted for by the combination of (a) still later readers not being counted, (b) memory loss at this test issue age. But with the data

available, we cannot quantify either of these.

An interesting tabulation is of the proportions who answer "may have seen" with increasing age (who are uncertain, but who are not counted as readers). These were as follows (as proportions of the screens):

1 month	.044
2 months	.046
3 months	.042
4 months	.059
5 months	.067

With increasing issue age, reasonably enough, increasing numbers of respondents are unsure whether or not they read or looked into the test issue before.

THE WEEKLY/MONTHLY COMPARISON

It is worth dwelling for a minute on the comparison between weekly and monthly read/screen ratios obtained by the different techniques. For TTB, the maximum across all test issue ages is shown below:

	Weeklies	Monthlies
RR	.46	.55
TTB (maximum)	.49	.45

TTB finds a *lower* ratio for monthlies than for weeklies; RR finds a *higher* ratio for monthlies than for weeklies. Which is the more reasonable?

Keep in mind that both sets of ratios are based on six-month screens. In literal terms, a six-month screen is a 26-issue cume for a weekly, and a six-issue cume for a monthly.

Obviously the screens cannot be taken literally in this fashion, but there should be *some* relationship. And if turnover levels for weeklies and monthlies are on average about the same (which both RMI and SMRB levels suggest), we would *logically* expect the read/screen ratios for monthlies to be

higher than those for weeklies, since fewer issues of them are published in a six-month period, and the total turnover involved for them over this period will therefore be less.

This is consistent with the RR relationship between monthlies and weeklies, which therefore appears to be more logical.

ROTATION EFFECTS

Since with half the interviews, RR questions were first and with the other half, TTB questions were first, it was possible to determine the effects of rotation on the RR and TTB levels. It was possible that the levels could have been affected by position in the interview and/or whether or not they followed questions for the other methodology.

The results tabulated by the rotation were as follows:

		Weeklies	Monthlies
RR:	Total	.46	.55
	RR first	.47	.56
	RR second	.45	.54
TTB:	Total	.47	.41
	TTB first	.44	.40
	TTB second	.50	.43

With the possible exception of TTB for weeklies, all the levels are very similar across this rotation. It seems that in these terms each of the two methodologies - RR and TTB - is 'robust'.

CONCLUSIONS

This study was carried out to determine what differences, if any, there are between aggregate audience levels for weekly and for monthly magazines obtained by recent reading and full Through-the-Book, starting with common screening levels, and to seek to explain whatever differences did emerge.

For weeklies, RR and TTB yield the same audience levels. Issue aging is not a problem for TTB, since weekly audiences are essentially fully accumulated in a couple of weeks.

For monthlies, it is a very different story. Their audiences accumulate over a long period. By the time they are essentially fully accumulated, TTB measurement will lead to considerable understatement due to forgetting. Indeed, in this study as in previous ones, reported audience levels rose with increasing issue age up to four months and then declined again.

At four months, the TTB levels are at a maximum. The RR levels are still however 22% higher than this maximum level. It is quite possible that this remaining difference is accounted for by a combination of (a) eventual audience members who still have not read, (b) forgetting on the part of some of the early readers.

This study cannot be itself confirm this. However, the relationship between monthly and weekly read/screen ratios is more reasonable for RR than for TTB, which is circumstantial corroboration.

The TTB methodology used in this study employed full test issues with all editorial items exposed to the respondent, and the same screening procedure as that used for RR. The RR procedure was identical with MRI's. The still lower audience levels reported by SMRB, therefore, are clearly due to a combination of the use of skeletonised test issues and lower screening levels.

Finally, this was a comparability study, not a validity study. The determination of what is 'truth' must await the development and employment of tested validation tools.

REFERENCE

Joyce, T (1982) *Recent reading* New York: Mediamark Research Inc. (Reference 15).