

A BRIEF ANALYSIS OF
MEASURING THE AUDIENCES OF MAGAZINES

Produced By

THE AMERICAN MARKET RESEARCH BUREAU

The introduction to this document indicates that it is attempting to establish "which technique or techniques are best capable of eliciting accurate recollection of reading events, in the context of a survey which must establish this for a large number of magazines?"

The primary conclusion (page 31) reached as a result of the investigation is "that the balance of advantage stay with the technique which has been specifically designed for handling large numbers of publications. This is the recent reading technique. It has stood the test of time and has produced authoritative results in many countries. It has the theoretical advantage of only requiring memory of recent events, and the practical advantage that the respondent's task is relatively easy, and that it can be administered by any data collection technique with essentially the same results."

Between the pages of the introduction and the conclusion drawn on page #31, little valid evidence can be found to substantiate the conclusion ultimately reached and indeed except for the obvious "practical advantage that the respondent's task is relatively easy", the investigation would appear to lead to just the reverse of this conclusion.

While this analysis will not undertake to cover every little point made in the document, many of which are not directly related to the conclusion drawn, it will cover the crucial points that bear directly upon the validity of the measurement technique.

On page #7, it is pointed out in reference to the through-the-book method that...

"if an issue is shown which is old enough to have accumulated all readers, some early readers may have forgotten it".

In support of this point, the results of a Study are cited on page #8 showing that three magazines obtained larger audiences with the use of older issues than with the use of younger issues in using the through-the-book method whereas for four other magazines just the reverse happened. How inconclusive can the argument get? The results were about equally divided leaving only to the conclusion that the evidence is self-contradictory.

Quite apart from this extremely flimsy evidence, however, it would be very difficult not to agree that if one used very old issues, it is bound to result in some forgetting of the reading that actually took place. If this were not true, it would certainly upset a lot of learning theory. In the test case cited, the use of six months old issues may very well have contributed somewhat to this kind of result.

On page #9, the suggestion is made that "there is evidence that the technique may under-report casual or irregular readership". Here the results of the Marder Study were cited in which reading in waiting rooms had been observed by planted interviewers and the readers subsequently interviewed.

When this Study was released a few years ago, the results surprised hardly anyone who had any real research knowledge of audience measurement. While the Study could be faulted seriously on many technical grounds, one does not really need a Study to indicate that the extremely casual and cursory reading of a magazine in a doctor's or dentist's office that the respondent might rarely have a chance to read could easily be forgotten in the space of several weeks. Thus, even with the substantial through-the-book memory aids such reading may not always be reconstructed at the time of the interview. The real point to bear in mind, however, is how much greater the likelihood of forgetting would be in such cases-or any other cases of reading - where the respondent is not even shown any memory aids to help him reconstruct the event of reading.

The further point on page #9 relating to over-claiming of reading of issues not yet published was refuted in a Study conducted by Politz which found no such over-claiming and also in a Study the Simmons Company conducted for

...cont...

Progressive Farmer under ARF's supervision in which only two such false claimers were found out of over five-hundred interviewed on a prepublished issue.

The document apparently attributes fallacies to the through-the-book method arising out of the use of the screening question and cites the Marder Study as evidence of this. In this connection, one should bear in mind that a failure to pass the initial screening question not only indicates that the respondent does not remember reading the particular issue of the publication that might have been observed by a planted interviewer in a dentist's office but it also indicates that the respondent has no recollection that he even may have read any issue whatever of the magazine within the last six months, for that is what the screening question asks him. In the case of the highly artificial Marder Study, the further evidence seems to bear out the fact that in these unusual situations respondents were indeed looking at magazines that they almost never have a chance to see otherwise. It would indeed appear remarkable if an appreciable number of respondents who state flatly that they do not think they even may have read any issue of the magazine within the past six months would, in fact, prove to be a reader of a relatively recent issue contained in the interviewer's kit.

One objective test of the screening question was conducted by the Simmons Company in a survey conducted for Mac Fadden-Bartell with ARF consultation.

... cont...

In this Study, as a methodological test, provision was made to take respondents through-the-book for two down-scale publications that the committee felt respondents might be reluctant to admit having read when asked the screening question. Only thirteen respondents out of a sample of four thousand indicated reading either of the two magazines after having first failed to pass this screening question. The committee attributed this phenomenon to a tendency of some women - somewhat naturally perhaps - to be reluctant to admit for reasons of prestige to having read the rather sensational romance magazines.

There is indeed an implicit assumption in the use of the screening question followed by the through-the-book procedure for those that screen in. It is that a respondent who has no recollection that he even may have read any issue of a magazine in six months will not, in fact, be a reader of a recent issue that the interviewer contains in the kit. The evidence of the test conducted in conjunction with the Mac Fadden-Bartell Study fully supports the position that this assumption is substantially true in the vast majority of real reading situations that occur. This is true despite the fact that in highly artificial circumstances of extremely casual reading under, perhaps, circumstances of tension imposed by an impending visit to the doctor or dentist, such reading may occasionally be forgotten.

In considering the use of stripped issues, one needs to bear in mind it is highly exceptionable for any respondent to qualify as a reader because he remembered

... cont. ...

seeing only one or two of the ten items included in the skeletonized issue. Quite the contrary, on the average, respondents will claim having seen seven or eight of the articles and features shown and very rarely does the number fall below three. Since the actual qualification as a reader is because of the memory of several of the articles and features shown the respondent, it can hardly be sustained that many actual readers of some other article or feature not shown in the kit could be missed in this way. The articles and features used in the kit are not a sample of just any kind of editorial matter but are selected specifically to include all of the prominently featured articles in the magazine including those listed on the cover and those that carry the large illustrations and so on. It must indeed be a rare reader who would miss seeing any of ten such major leading articles and features in the issue.

The final point the document makes in attacking the through-the-book method is a question of fatigue arising out of the large number of magazines to be covered (page#14). In point of fact, out of over sixty magazines covered by the Simmons Company in 1971, the average respondent proved to be an actual reader of only 4.3 such issues shown. This corresponded to the average respondent screening in on about seven such magazines. Since it takes only a couple of minutes to go through each of the ten articles or features with respondents using the reader-interest questions, this does not make for a burdensome interview. On the contrary, the entire readership questions on the average requires no more than twenty minutes and very rarely will they exceed thirty minutes although to be sure, occasionally, respondents will be found who read

... cont. ...

a great many different magazines so that it leads to a long interview.

In the infrequent cases where this happens, the respondent has an obvious interest in magazines and consequently is far less bored or fatigued by the interview than perhaps would be the case with others who are uninterested and who rarely read magazines at all.

Regarding the question of the feasibility of the through-the-book method using skeletonized issues for a large number of magazines, it is obvious that here the emphasis is being placed upon the wrong factor. It is not the number of magazines covered that makes for the long interview, it is the amount of magazine reading that in rare instances brings about this result. For instance, the total combined reader base the Simmons Company found in 1971 out of twenty-seven thousand interviews for the smallest twenty magazines collectively included in the Study was less than ten thousand readers. This so happens to be smaller than the base of readers found for Reader's Digest all by itself. What this means clearly is that the amount of time the interviewer spent in taking respondents through all twenty of these small magazines was less than the amount of time the interviewer spent in taking readers of Reader's Digest through the stripped down issue. In addition, about half of the total readership base for all magazines was accounted for by the largest ten magazines included in the Study.

Regarding the filter-recall questioning method, the document states on page #17 "whether people can accurately report the proportion of issues they read is a moot question". A Starch Study is cited as confirming that the two techniques produce roughly comparable results while the ARF - Politz Study clearly indicates substantial over-statement in audience figures produced in this way.

The conclusion that this is a "moot" question apparently is arrived at by overlooking a great deal of other evidence existing on the point. The Study by Corlett and Osborne, the Study reported by Don McGlathery while at SRDS, as well as several comparative tests which were produced by the Simmons Company in three separate years all indicate a very substantial degree of over-claiming of frequency of reading in response to such questions. This point was covered rather fully in the debate between Ed Barz and W. R. Simmons at the ARF conference in 1969 and the evidence need not be repeated here. It is contained in the proceedings of the Fifteenth Annual Conference of the Advertising Research Foundation October 14th 1969. The evidence strongly indicates that respondents will rarely admit reading only one of two issues of a magazine that they alledge to read at all in response to an unaided recall question (filter-recall method). On the contrary, they tend to report a great many more instances of reading four out of four issues or three out of four issues than can be reconciled with the average issue audiences produced by the through-the-book method (see the McGlathery Study reported in the ARF Journal of Advertising as well as several tests reported by the Simmons Company).

To suppose that respondents can accurately report the number of issues of say, Reader's Digest, that they have read in the last four months is to credit people with prodigious memories. Because they do not really know how many different issues of Reader's Digest they may have read in the last four months, it leads to the kind of respondent speculation that is seriously colored by considerations of prestige that in turn, leads to claiming of more frequent reading of books that they deem to be prestigious. Yet it appears that the recall of the number of issues read over a period of time is a method which the document in question promotes as the means of getting the information for reach and frequency projections.

The claim for validity of the "recent reading method" of estimating magazine audiences rests upon a rather extraordinary assumption which is stated rather clearly in the last two paragraphs of page #28. However, the enormity of the assumption is minimized by the rather simple statement that "these two effects (parallel reading and replicated reading)" in practice tend to balance out. The reason why we must consider this assumption rather enormous can be illustrated very simply by reference to table #2 on page #20 where it is shown that in that particular, well contrived example it does turn out, in fact, that the number of issues read in total exactly agrees with the number of weeks in which at least one issue were read. In other words, there are nine columns representing the weeks in which some reading took place and there are also exactly nine rows

... cont...

representing the issues that were read at some time.

This fine balance between the rows and columns would not exist at all if any one of a great many other reading patterns happened to occur. For example, if the second issue had happened to have been read in the second week instead of the third week it would have produced ten weeks of reading for only nine issues. If, in addition, the repeat reading of the fifth issue in the eighth week had taken place in the ninth week instead, we would now have eleven weeks of reading of the nine issues. Finally, if, in addition, the tenth issue had been read in the tenth week instead of the eleventh week, by now we would have had reading in all twelve weeks although, there are still only nine issues read. What special reason is there to suppose that the patterns of reading so conveniently group themselves in such a way that they balance the number of weeks in which the issues are read precisely with the number of issues ultimately read. Conversely, it is quite possible to have more issues read in fewer weeks as might occur for any given magazine - although we would deem this less likely to occur than the other way around.

Probably the most serious aspect of this assumption of an off-setting between "parallel" and "replicated" reading - as they are defined in the document, is a large extent to which the results would tend to favor some types of magazines as against others. We know, for example, that magazines vary widely in the lengths of their useful lives and the extent to which they engender

...cont...

repeat reading or a great many reading days. There may well be a substantial number of people who do, in fact, read say, Reader's Digest, over a period of say, three months. If they happen to be interviewed in any one of three different months they would thus be counted as a reader of the same issue magazine. (They have three chances to be counted as a reader compared to only one chance for a person who reads only in one month.) If, for example, at the end of the three months such persons acquire another pass-along copy of Reader's Digest which they read for another three months, they would be counted for six months of reading of only two issues with a consequent substantial inflation in the audience estimate it engenders. Unfortunately, one might expect magazines having a shorter useful life span to enjoy much less favoritism with respect to the inflation of their estimates of audience.

The somewhat more technical explanation of how this balancing is supposed to work out theoretically is shown in table #3 on page #22. The explicit assumption underlying this theoretical "proof" of the balancing proposition is stated in the last sentence on page #21, as follows:

"All we assume is that these **are** independent probabilities which are the same for all issues."

Now the hooker lies in the use of the word "independent" because it means, for instance, that the probability of reading a magazine in the second week

... cont...

by any individual is in no way affected by his having read it the first week. The statement also assumes that the probability of a person reading issue number two in no way affects the likelihood that the same person will also read issue number three. Looking at the problem in this light, it seems obvious that the probabilities can in no sense be deemed to be "independent". How can one possibly assume, for example, that the probability of a person reading Time magazine say, in the third week of its life, is unaffected by the fact that he had already read that issue of Time in the first week of its life. The probabilities of reading will vary from magazine to magazine and of course, from person to person and also from week to week but they do not vary in a random and independent way but according to many different reading patterns that occur among various population groups as they differentially apply to different kinds of magazines and different editorial formats and periods of issue. A person may well read an issue of Life the second week because he had seen something interesting in it the first week that he didn't have a chance to finish reading. In this case, the second week's reading is a direct result of the first week's reading and is not independent at all.

Having, therefore, made this assumption which we must consider contrary to known facts, the document proceeds with the "proof" that hangs upon this assumption by demonstrating that the probability is exactly the same for each row and the corresponding column (foot note at the bottom of page #23).

This statement of probability of reading, however, is completely fallacious

...cont...

where the probabilities of reading in successive weeks do, in fact, depend upon one another, according to the particular pattern of reading behavior any given individual may happen to exhibit.

For the technicians, where the probabilities of reading in one week do depend on whether or not the issue was read in previous weeks, the correct formula for the expression in the foot note on page #23 must be written in terms of conditional probabilities, that is, the probability of reading in the second week knowing that the issue was read in the first week. Conversely, the conditional probability will sometimes be the probability of reading in the second week knowing that the issue was not read in the first week. For the third week in any given case, the conditional probability might be the probability of reading knowing that the issue was read in the first week and was not read in the second week. Obviously, many variations in the conditional probabilities of reading in any week will occur and the correct expression for them becomes an increasing complex as the age of an issue increases.

One of the advantages that the document points out as being in favor of the "recent reading" method is the fact that it only asks people to remember their reading behavior for a relatively short period of time, that is, a week or a month according to the period of publication. While recent memory is obviously more reliable than memory for a long period of time, this

... cont...

proposition completely ignores the very big difference in what the respondent is asked to remember in this method as opposed to the through-the-book method. It is one thing to recognize an object or a person or an article or feature in a magazine as something that one remembers seeing before, but it is quite another thing to remember just when the object in question was seen. Few people, for instance, could be expected to remember that they had fish for dinner even three nights ago although, almost everyone can recognize fish as a dish that they have indeed eaten.

People do not read by the calendar anymore than they read by the clock and the significance of the fact of whether it was last Tuesday or Wednesday or maybe the previous Friday that they last looked into an issue of say, Sports Illustrated, eludes them completely. No one ordinarily pays any attention whatever to the particular date or day of the week that they may happen to pick up a magazine issue and thus, they form no associative memory whatever between the act of reading and the day on which it occurred. It just doesn't make any difference to anyone. Yet the validity of the "recent reading" method hangs mightily upon a correct memory of the time the reading took place. If, for instance, respondents mistake a reading event that happened, in fact, ten days ago as having occurred more recently say, within the last seven days, it leads to a 43% bias in the response. (The three extra days are 43% of the seven days asked about.

A great deal of work has been done that shows respondents tend to fore-shorten time when asked questions about the last time they did something. This is, in fact, seldom balanced by other respondents who mistakenly lengthen the time interval, possibly because the memory of more recent actual events is more accurately recalled. If a questioning method really seriously undertook to discover whether a reading event took place eight days ago or nine days ago or only seven days ago, it would seem to require a rather painstaking reconstruction of the time of occurrence of many other events that the respondent might possibly associate with the reading activity.

Possibly, the most serious consequence of this kind of question lies in the fact that respondents simply will not make the mental effort necessary to reconstruct the events surrounding the reading so that the time can be fixed in his mind, not only because it is time consuming but also because he simply attaches no significance to the day the reading occurred. Thus, the question becomes a loose question in which the respondent speculates off the top of his head regarding whether the reading was in the last week or not. The **greatest** trouble with loose questions in media research is they would almost invariably lead to serious inflationary biases for audiences of publications that are prestigious. Since the respondent does not, in fact, know the answer he tends more often to **supply the missing information in a manner**

...cont...

that satisfies his ego most satisfactorily. Obviously, the extent of the errors engendered by this form of question may very well be considerably larger for monthlies than they are for weeklies, since the time span over which the specific date of last reading must be recalled is much longer.

In considering the possibilities of the "recent reading" method, one must bear in mind an interesting experience of the Simmons Organization with such a method. In 1952, Simmons conducted a Study for Crowell-Collier based on thirty-three thousand interviews in which they asked respondents not about their reading over the last week or month - but about their reading "yeaterday" and the "day before yesterday". This is really what one might call a "recent reading" method. Even so, it turned out, in fact, and was so reported to the ARF technical committee, that a great deal more reading was reported "yesterday" than was reported for the "day before yesterday" whereas in point of fact the results of the two days should have been virtually identical, since over a period of time they reflect substantially the same days. The number of readers found "yesterday" was from 12% to 20% higher than those found for the "day before yesterday", where the reading of both days were reported by the same respondents. The **only explanation** that could be found for this curious state of affairs in such a large probability sample lies in the fact that respondents frequently reported when asked about their reading "yesterday" a reading event that actually took place "the day before yesterday". Thus, when they were asked a second

...cont...

question as to what reading they did on the "day before yesterday", there was nothing further to report.

The question arises, of course, as to how respondents can be expected to remember accurately their reading for a much longer period of time if they cannot even remember correctly their reading for the last two days. It is worth noting that this substantial 12% - 20% difference in the reading of the magazines reported for "yesterday" and the "day before yesterday" cannot possibly be a function of sampling errors, since the responses for both days were obtained from the same thirty-three thousand respondents. Thus, the differences for the two days must indeed reflect invalid responses which may be logically attributed to faulty memory about the time the reading took place. Incidentally, the size of the difference between the results for the two days was obviously a source of embarrassment to the sponsor and the Simmons Company as well and thus, one can be sure that every reasonable effort was made that could be made within the limits of accurate and forthright reporting to minimize the differences rather than to accentuate them. The conclusion nevertheless was inescapable; many respondents simply cannot report accurately their reading of six major magazines that took place within the last two days.

Again, in 1960, the Simmons Company conducted a Study for the purpose of measuring claimed ad readership scores among respondents who read issues

... cont...

of magazines "yesterday", the "day before yesterday", and the "day before that". While the average daily audience of the magazine was not a major purpose of this Study, as a by-product, again, the same phenomenon was noted; the number of respondents reporting reading "yesterday" exceeded substantially the numbers reporting their reading the "day before yesterday" which, in turn, was greater than the number reporting reading on the previous day. The results of this Study were published in Printer's Ink under the provocative title, THE SIMMONS - DODGE DEBATE (the subject under debate, incidently, had nothing to do with the reading of the issues but was concerned with the accuracy of ad page exposures).

Now, the document under review suggests the use of a not so "recent reading method" in which respondents are asked to report their reading accurately for a period of seven days for weeklies and thirty days for monthlies.