

NEW ORLEANS TO VIENNA: A RETROSPECTIVE

Valentine Appel, Consultant and Virginia Cable, The Wall Street Journal

This year will mark the 26th anniversary of the first Worldwide Readership Research Symposium in New Orleans. Since I was there and have attended all of the symposia which have been held since then, and have authored or coauthored sixteen papers – three of them prize winners – I think it appropriate today to offer a personal retrospective along with Ginny Cable of The Wall Street Journal who coauthored all but one of the most recent papers.

At the time of the first symposium there were two major print audience measurement procedures in use in the United States: (1) the through-the-book (TTB) method which was in use by the Simmons Market Research Bureau and (2) the recent reading method (RR) which was in use by Mediamark Research (MRI).

The Advertising Research Foundation (ARF) had just completed its Comparability Study and concluded that the RR method generated much larger audience estimates than did TTB, the more so as the publishing interval increased. Monthlies, for example, were more strongly advantaged by the RR method than were weeklies.

Both the RR and TTB methods employ a two step procedure. The first step presents the respondent with a large number of magazine logos and the respondent is asked to indicate whether or not each title might have been read or looked into in the last six months.

Those answering “yes” are said to have “screened-in.” Each title screened-in is then subjected to the next step which is quite different depending upon which method is employed.

Those for whom the TTB method is to be employed are presented with a suitably aged stripped down issue of each title screened-in and asked whether they had read or looked into it before. Those for whom the RR method was to be employed are simply asked whether they had read or looked into any issue in the last publishing interval – last week in the case of weeklies, last month in the case of monthlies, etc.

During the early years of the Symposium I was employed by the Simmons Market Research Bureau and I spent a great deal of time debating Timothy Joyce, the president of MRI, as to why the recent reading method produces incredibly large audience estimates.

In 1981 at the New Orleans symposium I introduced the concept of telescoping which explains that respondents cannot accurately recall when a particular media event may have occurred and imagine that it occurred more recently than it actually did. As a result, the recent reading method produces spuriously high audience estimates the more so as the publishing interval increases.

At the next symposium in Montreal in 1983 I demonstrated that the U. S. Census Bureau had been wrestling with the telescoping problem for more than 20 years and that it has been observed in such diverse areas as crime victimization, home repair and others.

Later in Salzburg, I expanded upon the telescoping notion and demonstrated that relative to TTB the RR method also produced disproportionately higher estimates for titles which are read with lesser regularity i.e. titles with higher turnover rates.

At the Hong Kong symposium in 1991 I presented a paper entitled “Anatomy of a Magazine Audience Estimate.” The paper consisted of a reanalysis of the ARF Comparability Study which had been published in 1980. The Hong Kong paper deemphasized the importance of the through-the-book vs. recent reading controversy, concluding that “regardless of whether the TTB or RR procedure is employed, the screen-in level is the principal determinant of the differences in magazine audience size after circulation levels and publishing frequency are taken into account.”

Later the MPA (Magazine Publishers of America) had commissioned Dan Mallett to verify the results by reproducing my Hong Kong analysis using the MRI and Simmons data bases.

The Mallett analysis was presented in San Francisco in 1993, and set out to accomplish two objectives: (1) to verify my Hong Kong findings, and (2) to explore the relationship between the Simmons and MRI year-to-year changes in screen-in levels and average issue readership estimates.

The Mallett paper essentially confirmed my Hong Kong findings, but the year-to-year analysis produced a perplexing finding. As expected, the MRI data showed a close relationship between year to year changes in their screen and read data, but the Simmons data did not, implying the superiority of the TTB method.

To everyone's surprise, notwithstanding these findings, shortly thereafter Simmons announced that in order to increase the number of titles measured they were abandoning the TTB method and adopting RR.

Shortly after the publication of the Hong Kong paper I entered retirement and began a consulting relationship with The Wall Street Journal. Our first undertaking was the paper I presented with the late Mike Stien at the Berlin Symposium in 1995. Then and now, the syndicated print measurement services typically use a six month screening interval for magazines but a seven day interval for daily newspapers, a procedure which we suspected disadvantaged the newspapers.

The paper presented the results of a study which compared the average issue readership levels obtained for daily newspapers using a six month screen vs. those observed using a seven day screening interval. The study clearly showed that the national newspapers – The Wall Street Journal, USA Today and The New York Times – generated much larger audience estimates when a six month screen was used.

As a result of these findings, the ARF amended its Newspaper Research Guidelines to make it clear that its seven day screening interval recommendation did not apply where both newspaper and magazine audiences are to be compared.

Based on these findings and the ARF amendment of its Newspaper Research Guidelines, MRI changed its procedure. The Wall Street Journal, The New York Times and USA Today were added to the logo deck and screened in random order along with the magazines. Also, consistent with the procedure used for magazines, the screening interval for the two national newspapers was changed from seven days to six months, and those screening-in were directly asked whether they had read the publication yesterday.

As soon as the changes were made, a dramatic and alarming increase in readership was observed, and in subsequent waves MRI reverted to its earlier procedure, which it continues to use to this day.

Changing the subject, at The Berlin symposium in 1995 MRI and I presented a paper describing the development of a computer assisted interviewing procedure for measuring magazine audiences. The paper demonstrated that although computerized interviewing produced results which were highly correlated with those obtained via personal interviews, it also generated audience estimates which were materially higher.

At the Vancouver symposium in 1997 the Journal and I presented the results of two studies, one executed by MRI and one by Simmons, which attempted to understand the reasons underlying the dramatic increases experienced by MRI as a result of their having changed their questioning procedure. Although the studies did not completely explain the reasons underlying the surprising MRI findings the importance of the change in screening interval from one week to six months was confirmed.

At the same symposium I, along with Mike Skrapits, presented an analysis of the relationship between changes in circulation and changes in readership. The analysis clearly showed that, contrary to the commonly held belief, when a circulation change is sufficiently large relative to the sampling error of the audience estimate, a change in circulation is indeed highly predictive of a change in audience.

At this point I would like to turn the podium over to Ginny Cable who coauthored with me papers presented in 1999, 2001 and 2003.

At the Florence symposium in 1999 the paper we presented took a totally different tack, correcting survey non-response bias. The paper demonstrated how a large scale database can be used to correct some of the bias of non response, which is a problem in all surveys. As a result of this bias, before correction The Wall Street Journal's audience was older than it should have been, and the correction procedure was adopted in future studies.

At the same symposium Andrew Elder, Tony Incalcetera and Val reported the results of a controlled experiment which demonstrated that although a self-administered questionnaire conducted via the internet produced readership estimates which were highly correlated with those produced using a self-administered paper and pencil questionnaire response rates were materially lower and readership estimates were materially higher on average.

At the Venice symposium in 2001 we presented a study which explored, using an on-line self-administered questionnaire, the effects of using three different questionnaire formats upon the audience estimates produced.

The study findings were: (1) Isolating specific groups of titles and questioning about them separately ala MRI, raised their audience estimates, (2) placing the grouped titles early in the questionnaire, also ala MRI, also raised their audience estimates and (3) contrary to expectation, theory and other research grouping titles by similarity of name or content raised their audience estimates as well.

Again changing the subject, recent years have seen the development of on-line editions of print media vehicles. Our Cambridge paper in 2003 demonstrated, that unless respondents are instructed specifically not to include on-line reading they will do so thereby exaggerating the extent of their print media exposure.

Most of these issues are as relevant and important today as they were 26 years ago. However, you don't hear much about these topics anymore. You hear about fusion and cross media measurement and issue specific readership. As we all know, a fusion is only as good as the data being fused, and if the underlying studies have not addressed the age old questions about title confusion, telescoping, non-response bias, the effects of screening intervals, etc., we're no further ahead than we were in New Orleans in 1981. The same is true of cross media measurement and issue specific readership. As long as print readership data is collected via surveys, we shouldn't lose sight of the underlying factors that impact the accuracy and validity of our decision-making tools.

Most people feel that audience measurement is at a crossroads and that a significant overhaul of the way we measure readership, and all media usage for that matter, is imminent. I read on the ARF website that the industry is at an "axiom-shifting perfect storm moment". If this is true, it's more important than ever that we revisit, re-evaluate and really act on what we've learned so far as we embark on the next chapter of audience measurement. With that in mind, I'd like to leave the media research community with the essence of that learning:

1. Because respondents cannot accurately report the recency of occurrence of a particular media event, they tend to overestimate it, the more so for the publications with longer publishing intervals. As a result of this, the recent reading method produces spuriously high audience estimates which are biased in favor of less frequently published titles.
2. The screen is what determines the size of audience of a particular title regardless of the level of readership claimed among those screening-in.
3. For this reason, longer screening intervals produce larger audience estimates than do shorter ones.
4. The existence of online versions of print publications results in inflated estimates of the audiences of the print vehicles themselves unless some mechanism is adopted which excludes exclusive on-line reading from their readership claims.
5. Computerized interviewing has many advantages. In addition, it produces audience estimates which are highly correlated with those obtained via conventional methods. At the same time it produces audience estimates which are materially higher.
6. The bias of non-response can be partially compensated for by merging a large-scale database with survey results and weighting them appropriately. The Journal now uses this procedure for weighting its subscriber studies.
7. Notwithstanding studies which have failed to demonstrate changes in audience corresponding to changes in circulation, circulation changes are predictive of changes in audience. The reason these studies have failed to demonstrate it is because, although the circulation changes may have been relatively large, they were often smaller than the sampling errors of the measured audiences.

BIBLIOGRAPHY

(In chronological order)*

- New Orleans, 1981. "Telescoping: The Skeleton in the Recent Reading Closet." Session 2.17.
- Montreal, 1983. "Hypnosis and Telescoping: Some Rediscovered Facts." Session 4.5.
- Salzburg, 1985. "Reading Regularity and Inflation of Magazine Audience Estimates." Session 2.10.
- Barcelona, 1988. "The Editorial Environment Controversy." Session 6.5.
- Hong Kong, 1991. "Anatomy of a Magazine Audience Estimate – The ARF Comparability Study Revisited."
- San Francisco, 1993. Mallett, Daniel T. Jr. "The Relationship of Screens And Reads and the Role of Screening In Readership Measurement." Session 4.1
- San Francisco, 1993. "How Changes in Screen-Ins Affect Reads." Session 4.2.
- Berlin 1995. Baim, Julian, David Arpin and Valentine Appel. "Using Casio-Audio (Computer Assisted Self Interview With Audio in Readership Measurements." Session 2.2
- Berlin, 1995: Appel, Valentine and Michael G. Stien. "Newspaper Screening Intervals: Six Months vs Seven Days." Session 3.1
- Berlin, 1995: Pinnell, John and Valentine Appel. "How Computerized Interviewing Eliminates the Screen-Bias of Follow-up Questions." Session 3.6.
- Vancouver, 1997. Skrapits, Mike and Valentine Appel. "The Relationship Between Changes in Circulation and Changes in Readership." Session 9.2.
- Vancouver, 1997. "Frequency, Recency and the Moderating Effects of The Screening Interval." Session 3.6.
- Florence, 1999. Cable, Virginia, Dan Jennings and Valentine Appel. "Using Database Overlaps to Correct Survey Non Response Bias" Session 4.9.
- Florence, 1999. Elder, Andrew, Tony Incalcaterra and Valentine Appel. "Measuring Print Audiences Via the Internet." Session 9.4.
- Venice, 2001. Cable, Virginia and Valentine Appel. "Title Confusion and Physical Proximity in a Self Administered Questionnaire." Session 5.3.
- Cambridge, 2003. Cable, Virginia, Gennaro Nunziato and Valentine Appel. "The Impact of Online Reading on Traditional Print Media Estimates." Session 5.3.

 *Where no author is listed, Valentine Appel is the sole author