Wayne P Eadie
Newsweek
New York, USA

2.4   WAYS OF MEASURING OVERCLAIMING

In Montreal Steve Douglas and I reported on a series of studies which involved the first-time read yesterday approach by telephone for estimated average issue audience. These studies included experimental work, a validation study relative to underclaiming and a large scale readership study. Here I report on a natural extension of this work also conducted by Audits & Surveys which uses the same method of estimating average issue audience. The new work tests some methods of evaluating the first-time read yesterday method relative to overclaiming.

As noted in Montreal, in the United States the Recent Reading and Through-the-Book techniques are competing methods for measuring average issue audiences. These two methods exhibit systematic and substantial differences in audience estimates. The Recent Reading estimates are higher and the differences are particularly pronounced for monthly magazines.

This chronic state of affairs has created substantial interest in developing a validation measure against which the audience data of the two commercial services could be compared. The beginning of such work was reported in New Orleans in 1981 under the heading, the ARF Certitude Tests. This work did not yield a validation measure: rather it uncovered some problems related to high levels of underclaiming for observed public place readership. In deciding how best to proceed in pursuit of a valid method of measuring audiences, we noted that in New Orleans there was considerable interest in, and discussion about, developing the first-time read yesterday approach as a validating measure.

The initial small scale validation studies conducted on underclaiming which we reported in Montreal were successful. There were two studies in which eventual respondents were surreptitiously observed actually

reading new issues of *Newsweek*. In one case the observations were made in-home by spouses of *Newsweek* subscribers. In the other interviewers observed normal customers or patients in barber shops, beauty shops, doctors' offices or dentists' offices. The underclaiming levels were about 10% in each case. We were particularly encouraged as this was the first time we know of where a measurement technique generated such low levels of underclaiming for public place reading.

We have pursued this technique beyond the work we reported in Montreal not only because the initial validation studies gave positive results but because our exploration of some types of overclaiming could parallel and aid the ARF validation programme. Thus, two studies we have just completed not only have value on their own, but they are also, to some degree, pre-tests for studies two and five in the ARF validation plan, as described by Dick Lysaker in Session 1.

Before proceeding, it is important to point out that a good part of our initial interest in the first-time read yesterday technique is the fact that it is the only technique for measuring Average Issue Readership that lends itself to fairly extensive and realistic measures of overclaiming. All methods can be evaluated in terms of underclaiming. It is fairly direct and simple to observe someone reading an issue of a title and subsequently ask questions about readership either by using the Through-the-Book method or the Recency method. Analysing such results for the Through-the-Book method is straightforward. Respondents known to read the issue in question do or do not claim readership of it. With the Recency method, the analysis is the same. However, there is no way of directly relating observed readership of a specific issue to the claim of having read one or more issues of the magazine during the publication interval.

With overclaiming, the situation is far different. Neither technique allows validation testing for overclaiming where the criterion is observed reading or non-reading of an issue or a title across extended time periods. Some of the numerous weaknesses of the confusion control method are discussed in the paper by Dieter Müller-Veeh.

The situation is much better with the first-time read yesterday technique. Although complete validation is not possible across all conditions, a substantial amount of validation checking is possible. As noted in the earlier description of the ARF work, overclaiming can come about in a number of ways.

(1) Readers who did not read yesterday can claim yesterday first-time reading

(2) Readers on multiple days who read yesterday, can claim yesterday was the first day of reading.

(3) Non-readers can claim first-time reading.

Furthermore, such occurrences can involve new or old issues. Or they can involve an issue which was new as of the first reading but old at the time of a subsequent reading.

The levels of overclaiming generated in connection with the first-time read yesterday technique which can be subjected to direct observation examination include the following reading conditions:

New issues read in-home or in a public place
where the issue could be
- Not read at all or
- Not read yesterday or
- Read yesterday but not for the first time

New or old issues read in-home or in a public place
where the issue was read some time earlier and was
- Not read yesterday or

- Read yesterday.

Claimed first-time reading of older issues does not lend itself to direct observation. Here, only a surrogate measure is possible. The measure we used is a determination of whether or not the issue of the magazine was in fact physically available at the place readership is claimed.

The developmental work we have done includes a number of these conditions. One study deals with reading of new issues in subscribing households. The other deals with locating public places where respondents claim to have read an issue yesterday to determine whether or not the issue still is available there or was, in fact, available there yesterday.

The general model and methodology used for our prior studies as well as the current ones have a number of key elements.:

- Interviewing is conducted by telephone across a time period equal to one or more publication intervals

- The readership question is time specific but not issue specific

- Parallel reading is directly accounted for by determining the number of different issues read yesterday

- Replicated reading is accounted for by including only first-time readers in average issue audience estimates

- The estimate of average issue audience is simply the number of first-time issue readers generated across the publication interval

- All work reported in Montreal and here is based on a constant group of 31 titles which represent a wide range of weekly and monthly magazines as well as both general and special interest books.

The first-time readership question we are using is: "Did you happen to read

or refer to that issue of (magazine) however briefly, before yesterday?"

In order to measure overclaiming of new issues in subscribing households we selected a sample of subscribers in zip codes known to regularly receive Tuesday delivery of *Newsweek* magazine. Qualified subscribing households contained a husband and a wife where the wife would be at home at all times when the husband would also be at home across the period of the study. This generally involved households where the husband is working and the wife is not. In each household the wife was recruited to surreptitiously note any magazine reading done by the husband when he was at home. Additionally the observers noted any magazines that arrived at the house through the mail during the observation period. In this way we know when the target issue arrived and could have been read but was not. Neither the wife nor the husband knew why they were selected or that we had any special interest in *Newsweek* magazine.

Participating housewives observed their husbands from Tuesday to Friday or until an interview which could detect overclaiming could be conducted. A total of 46 interviews were completed. An analysis of the relevant reading condition and readership claims shows that overclaiming occurred in six or 13% of the cases.

	Number of observations	Number of overclaims
Read yesterday not first-time	11	1
Read but not yesterday	27	4
Did not read at all	8	1
Total	46	6

As is shown above all types of overclaiming were detected, but at low levels. Our conclusions are that this first-time read yesterday technique continues to show promise. Relative to the ARF experiment 2, we have determined that it is, in fact, feasible to recruit spouses to observe and report on readership and magazine receipt at least over a four-day period.

The second study dealt with measuring the availability of copies of issues respondents claim to have read yesterday in a public place. For the purpose of this study a public place included the following types:

Barber shop	Library
Beauty shop	News-stand
Dentist office	Hospital waiting-room
Doctor's office	University lounge

Respondents were interviewed using the first-time read yesterday questionnaire described earlier. A qualified respondent was one who claimed yesterday readership of any one of the 31 target publications and who claimed that the reading occurred in one of the kinds of public places just listed.

When a respondent claimed yesterday readership of a magazine, specific questions were asked to determine as much information as possible about the location of each public place. This included name, address and general location. Specific questions were also asked about each issue regarding date, cover illustration and contents. In this way we were able to characterise all potential respondents as being able or unable to give accurate enough information about both the location of the reading and the issue read so that copy availability could be determined. One day after the readership interview, visits were made to each specified public place where reading was claimed. In each public place, the magazine inventory was physically examined. If this examination did not uncover the issue in question, the receptionist or another knowledgeable person was questioned about whether or not the issue in question was, in fact, present two days ago.

The relevant sample size for this work consisted of 96 issues of magazines for which yesterday readership was claimed. These claims came from 70 respondents. The base of establishments to which visits were made is 55.

The first question is to what extent the questions we used were able to elicit sufficient information from respondents to allow a follow-up check to specific public places for specific issues.

An issue was considered located if we found an issue of the specified magazine with a date, or cover or specific contents which corresponded to the respondent's description.

Information obtained from respondent

	Number	%
Issue and place	70	73
Place but not issue	18	19
Issue but not place	3	3
Neither	5	5
Total	96	100

Thus in about three or four instances where respondents' claimed yesterday readership in a public place, we obtained specific enough information to make the validation check. However, in about one out of five cases we obtained enough information to locate the place of reading but not enough information about the issue read. Clearly this is a problem area. As a result, after the experiment, we conducted some further interviews to see if revised, more elaborate questioning could improve our capture rate for issue identification.

When respondents could not remember anything specific, we asked, "Here are some types of stories or articles that traditionally appear in (Magazine title) - do you recall seeing or reading any stories or articles like these?" If they did we asked, "Please tell me what it was about. What did it say? What did it show?" When respondents mentioned only a subject

which would have occurred in many issues and could give no particulars, we asked, "Was this a general story or did it deal with a particular incident? Were there any photos? Where in the magazine did you find it?" The sample was very limited but in six of eight cases where the initial questioning failed to elicit information on the specific issue being read, the follow up questions succeeded. Clearly more work should be done in this area.

In total we were able to attempt validation of issue availability for 70 issues in a total of 55 establishments. For 15 of these issues, we were unable to determine whether or not they had been present on the day for which the respondent claimed readership, that is two days prior to the visit to the establishment. The receptionist or another knowledgeable person said the issue may or may not have been there.

Locating issues claimed to have been read

	Number of issues	%
Could determine if issues had been available	55	78
Could not	15	22
Total	70	100

In those instances where we made a positive determination on the presence or absence of an issue the validation level was 78%. These results were virtually the same for those issues respondents said had been read yesterday for the first time as for those respondents said they had also read previously.

Locating issues claimed to have been read

Issue available	Number of issues	%
Found issue	41	78
Issue removed	2	3
Issue not available	12	22
Total	55	100

This represents the first test of this method of validation. The results are encouraging but the actual validation level must be higher in order to say that the readership measurement technique is satisfactory. The validation level we achieved is, of course, both a function of the accuracy of the readership measurement and our determination of place of reading and issue read. We suggest that further work could well improve the information that can be obtained on both place of reading and issue read. This of course will be part of the ARF work. One possibility is to forewarn the respondent tactfully that we will be going hunting at the claimed place of readership for the issue in question. For example, did they claim the wrong place of readership or did they claim the wrong issue? This type of research is clearly in line with the plans of the ARF Gold Standard Committee.

Up to the present, we at *Newsweek* have devoted funds towards independent studies of the validity of one version of the first-time read yesterday approach. In the future we plan to continue our contribution by supporting the Gold Standard work of the ARF.