

4.4

DEVELOPING AND VALIDATING A MEASUREMENT OF 'FIRST READING'

INTRODUCTION

Throughout the majority of its life, the South African All Media and Products Survey (AMPS) has employed Recent-Reading-based estimates of average issue readership in the case of newspapers and reading frequency-based data for magazines using, in the latter case, a sixpoint, numerical scale of claimed frequency and (simplifying slightly) the probabilities nominally implied by the scale.

In this market, the debate on and criticism of the validity of the AIR estimates has been as intense as anywhere and, recently, more vociferous than in most other countries where regular, large-scale readership research takes place. High readers-per-copy figures have come repeatedly under attack. As regards the causes of alleged overestimation of readership, whilst the likelihood of sheer prestige related overclaiming (particularly amongst certain population groups) was widely mooted, it was also accepted (on the basis of strong, if indirect evidence) that replication could also be a major culprit. In 1986, within a continuing programme of developmental and experimental research, the South African Advertising Research Foundation (SAARF) decided to test the practicality of obtaining estimates of 'first' reading and the apparent, relative validity of such estimates, against a circulation criterion.

Two general points are to be underlined at this stage.

First, the objective was not to check-out a technique which would necessarily be suitable for full-scale field use (in terms of interview length and complexity, for example), but rather to maximise the chance of obtaining 'good' first

reading claims. If the results proved promising, it was then the intention to consider the 'streamlining' of the first reading measure.

Second, the study concerns first reading *in the issue period*, rather than 'yesterday'. Whilst, obviously, this adds a further limitation to likely accuracy, the adoption of a CATI/FRY approach would not be practicable in South Africa.

Under the constraints of space and presentation time, this paper concentrates just on three aspects of the research: its design, with particular reference to certain features of the questionnaire and fieldwork; the first reading estimates vis-à-vis conventional AIR estimates; and their apparent, relative validity. A post-script links this study with later SAARF work.

RESEARCH DESIGN

The universe sampled comprised adults of 16 years of age or more, in one South African Province only – the Transvaal – with the exclusion of four (numerically negligible) population subgroups.

The planned sample size comprised 900 completed interviews, with some disproportionality of geographical allocation.

Required numbers of households were obtained by dividing numbers of interviews by estimated household sizes, it being the intention to attempt interviews with *all* members of a selected household: this critical feature of the research design reflected the aim to classify households as 'primary' or not on the basis of direct evidence.

Numbers of households were, in turn, converted to numbers of clusters by dividing by five for urban areas and by four for rural ones. Clusters were allocated to urban and rural areas proportional to population, using a conventional random start, fixed interval technique.

The location of cluster starting addresses employed telephone number listings as a sampling frame for Whites in cities and rural areas only and randomly generated grid references for other groups. Selection of addresses subsequent to a starting address employed a strictly controlled random route technique.

Each selected address was randomly pre-designated as requiring an (initial) male or female interview. Within household, the selection of the man or woman for initial interview employed a Politz grid.

Subsequently, further interviews were attempted with all other adults in any household where a successful first interview had been achieved, up to a total of five in any one household. For both initial and other respondents, up to four calls-back were made.

Turning to the questionnaire, the interview commenced with some cinema-going questions, solely for the purpose of establishing initial rapport.

Respondents were then shown, in randomised order, masthead cards for the 19 daily, Sunday and weekly newspapers relevant to the survey area and for ten magazines (a materially smaller number than in AMPS itself), selected as having the largest average issue readerships in previous findings. The masthead cards showed monochrome illustrations of the publications' logos.

Four filter questions were then asked, sequentially and 'vertically' (ie. Q1, title by title; Q2, title by title; and so on).

The first of these questions related to claimed readership 'sometimes'.

The publications not read 'sometimes' having been set aside, the second question employed a sorting board with a four-point, verbal frequency of reading scale: 'every issue', 'most issues', 'some issues', and 'very few issues'.

The masthead cards were then sorted by publication group and the third filter question was asked in relation to monthlies first, then fortnightly, then weeklies and finally dailies; it comprised a straightforward six-issue-period when-last-read sort.

The fourth question in this initial set covered habitual source of copy; again employing a sorting board, it asked "... show me how you *usually* get the copies of (*Title*) you read".

For each newspaper or magazine claimed as read 'sometimes' – but irrespective of the other claims so far collected – further questions were then asked, but now 'horizontally'; that is to say, all these questions were put, one after another, in respect of the first title considered, before passing on to the next publication.

The first of this second set of questions asked:

"Think of the last time you read or paged through (*Title*). Where were you the last time you read or paged through a copy of (*Title*)?"

Next, source of the specific copy *last* read was established (as distinct from 'habitual' provenance, covered in an earlier question).

The date of the last reading event was then questioned, using a sorting board offering six mutually exclusive time periods as possible answers, so chosen that three fell within the publication interval (and thus would lead to the 'positive' classification of the respondent as a

'reader') and the other three did not. Self-evidently, it was necessary to name different sets of time periods for dailies, weeklies, fortnightlies and monthlies.

The interview then turned to the critical measurement of 'first' reading. Initially, respondents were asked on how many different days they usually read a copy of the title concerned – but note that (unlike some techniques), this question was *not* employed as a filter to the one that followed. The timing and place of the last reading occasion were then rechecked, before the interviewer said:

“Think of that occasion. Imagine yourself holding a copy of (*Title*) and reading or paging through. It was a particular copy with a particular cover and particular articles and pictures in it. Try to remember it.”

“Was that occasion the *first day* you read or paged through that *particular* copy of (*Title*) – or had you read or paged through *that* one on some earlier day? Think carefully.”

If earlier reading was claimed, its date was checked in terms of the number of days or weeks ago.

This second section of readership questions concluded with a measurement of reading frequency, in terms of “out of every six different issues how many do you usually read or page through?”.

The interviewer then established the first names of the other adults in the household – with all of whom interviews were to be attempted – before reshowing the full set of masthead cards and asking, in respect of each of the other adults and each title, whether they “ever read or page through” or whether the respondent just did not know. These ‘indirect’ data were not, however, collected at the last interview in a given household because, by then,

‘direct’ data would have been gathered from everyone.

The questionnaire just described was formally piloted on 12 respondents, six White and six Black, with probing for comprehension and with particular attention to the flow of the interview. A proportion of these pilot interviews were held under viewing-room conditions.

SELECTED RESULTS

Achieved sample

876 adult interviews were completed, against the target of 900. Of these, 573 were in ‘complete’ households – households, that is, where all adults (up to the maximum of five in any one household) were successfully contacted.

Because certain aspects of the analysis turn on individuals and others on households, the results reported below relate to the smaller, 573 base, it having been thoroughly checked, as a preliminary, that there were no major significant differences (in terms of readership behaviour) between adults in ‘complete’ households and the total sample of individuals.

Readership levels

As we shall see, the pattern of results obtained differed somewhat between the racial groups, which are therefore considered separately, here and throughout this section.

Considering gross readership, summed across titles and indexing the first reading results on the recency levels, the data for the newspapers are shown in Table 1.

(Note that six minor titles, of negligible penetration on any of the survey’s measures, have been excluded from the analysis.)

DEVELOPING AND VALIDATING A MEASUREMENT OF 'FIRST READING'

Table 1

Gross readership indexed

	Whites	Blacks
Five daily newspapers		
Recent reading (issue period)	100	100
First reading in issue period	71	35
Seven weekly/Sunday newspapers		
Recent reading (issue period)	100	100
First reading in issue period	88	65

The drop in gross penetration with the imposition of the first reading requirement was considerable and, further, observable for daily titles, in the case of which replication would not normally be regarded as a likely bias.

Turning to the magazines, it is of interest to show the indexed reading-frequency-based average issue readership levels also, as being comparable with the estimates previously published in AMPS (Table 2).

Table 2

Gross readership indexed

	Whites	Blacks
Ten magazines		
Recent reading (issue period)	100	100
Reading frequency	97	117
First reading in issue period	76	35

Thus, for Black respondents, the first reading 'drop' would have been even greater if referenced on the AMPS measure previously employed for magazines.

Since all of the above comparisons are based on data aggregated across titles, it is reasonable to ask whether the general pattern of the results holds good at the publication level. For the product-moment correlations, across titles, between the readerships based, respectively, on Recent Reading and on first reading in the publication interval, the coefficients were as shown in Table 3.

Thus, with the one exception of Black respondents' reading claims on magazines, there was very little between-title variation in the first reading/Recent Reading ratio.

Validity of readership claims

Whilst, as we have seen, the imposition of the 'first reading' qualification materially reduced the Recent-Reading-based AIR levels, to decide whether this reflects a reduction in the bias arising from replicated readership and, consequently, an improvement in validity requires further evidence.

In respect of each newspaper or magazine on the media list, each 'complete' household (where all adults had been successfully interviewed) was classified as 'primary' or 'not primary'. A household was operationally

Table 3

	Whites	Blacks
Daily newspapers (7)	0.98	0.96
Weekly/Sunday newspapers (5)	0.99	0.95
Magazines (10)	0.96	0.53

defined as 'primary' if, for the title concerned, one or more respondents in that household claimed that the copy last read had either been personally acquired by them or by another member of the household. This classification was carried out using both issue-period readership and first reading within the issue period.

(A methodological loose end, which will be readily apparent, should be pointed-up: since the readership measurement was, by its nature, not issue specific, it is quite possible for two adults in the same household to claim personal acquisition in relation to different issues of a title).

Next, the regressions, across titles, of survey area circulations on the numbers of primary households were examined, separately for newspapers and magazines. The results are shown in Table 4.

Table 4

12 daily, Sunday and weekly newspapers

$$C = 23.7 + 0.31 H_R \quad r^2 = 0.54$$

$$C = 6.8 + 0.48 H_F \quad r^2 = 0.81$$

10 magazines

$$C = 37.8 + 0.57 H_R \quad r^2 = 0.51$$

$$C = 29.6 + 0.69 H_F \quad r^2 = 0.81$$

where

C = survey area circulation ('000)

H_R = number of primary households ('000),
recent-reading-based

H_F = number of primary households ('000),
first-reading-based

These results strongly support the *relatively* greater validity of the 'first reading' estimates, since the variations in numbers of first-reading-based primary households more closely model circulations, in terms of the proportion of variance explained, of the nearness of the intercepts to zero and of the closeness of the slope coefficients to unity.

But – and it is a considerable 'but', in the overall assessment of the research methodology and of the success of the validation route adopted – the slope coefficients are *not* 1.0, by a wide margin; we have approximately 2.1 primary households (on the 'first reading' classification) per newspaper copy sold and approximately 1.4 in the magazines' regression.

There are, of course, a whole number of possible sources of error: non-private-household circulation has been ignored – but deduction of 'institutional' copies from total net sales would worsen the results just cited. Recent Reading may still have been overclaimed, for prestige, via telescoping or from other causes. Too high a proportion of issue-period reading events may have been classified as 'first' (although the implied numbers of reading days are not incredible). And, finally, we have no direct evidence at all on the validity of the source-of-copy claims.

Space does not permit full reporting here of all the very considerable volume of other analyses carried out, but one result is of interest as being indicative of problem areas.

A multiple regression analysis was carried out, now across all titles, newspapers and magazines, with circulation again as the dependent variable and, as independent variables, numbers of primary households (using the 'first reading' definition), the proportion of total audience represented by Black readers and the proportion accounted for by readers claiming a frequency of six issues out of six. The multiple

r^2 was 0.92, with the latter two variables accounting for 14% of variance.

ENVOI

Despite the limitations discussed here, the results of this study were felt, by SAARF, to provide sufficiently encouraging evidence of the relatively greater validity of a first-reading-based estimate of average issue readership to justify further work. In a later, major test (referred to in the paper by Gert Yssel at this Symposium) four versions of a 'first reading' questionnaire and two methods of interview administration were tested in a full experimental design; circulation/primary household correla-

tions were again employed, *inter alia*, in arriving at a choice between the methodologies. The 'winning' version was introduced, full-scale, from July 1988.

Acknowledgements

The author would express his thanks to Mr G J Yssel, Managing Director, South African Advertising Research Foundation for permission to quote these results and to Mr W A Hunt, Director, Research International South Africa (Pty) Ltd who was in charge of the execution of the research and who made invaluable contributions to its design, analysis and interpretation.