

A REVIEW OF AMPS READERSHIP METHODOLOGY

Gert Yssel, SAARF, South Africa

Abstract

Up to 1986, the South African AMPS survey used recency to estimate readership for newspapers and frequency for magazines. The resulting readers per copy figures were considered too high. A 1986 experiment and its results led to the introduction of a readership questionnaire for the AMPS survey from 1987 onwards, which uses 'first reading within issue period' coupled with the grouping of titles within publication issue periods. That method had potential theoretical weaknesses which are discussed, and which led to a systematic series of experiments.

The experiments in 1989, 1990 and 1992 investigated possible changes to circumvent the perceived weaknesses without introducing unacceptable changes to the data levels. An evaluation of the results of the experiments ends with a justification of the decision on the readership questionnaire to be used from 1993 onwards.

Introduction

Two papers dealing with the introduction of 'first reading within issue period' to AMPS were presented at the Barcelona Symposium in 1988. The first described the 'first reading within issue period' experiment conducted in 1986 for the South African Advertising Research Foundation.

The second paper referred briefly to the adoption of first reading as the measure for estimating AIR, and to the interim 1987 results produced by the method. It mentioned the fact that a proportion of the total variance of the readership estimates, was related to the so-called rotation effect and to the circumstances of the new readership section of the questionnaire. These and other factors resulted in a series of tests between 1989 and 1992 of amended versions of the questionnaire. A somewhat detailed description of the readership questionnaire which was introduced in 1987 is necessary to understand the reasons for these studies.

The 'first reading within issue period questionnaire'

Titles were grouped by their publication period into the five publication groups: daily newspapers, weekly newspapers, weekly magazines, fortnightly magazines and monthly magazines.

The prompts consisted of a number of show cards, each containing black and white replicas of the mastheads of up to six publications with the same publication interval. To reduce the number of prompts required for any particular interview, the mastheads of only those newspapers which were distributed in the area of an interview, were asked about. For this purpose, separate sets of regionalised newspaper prompt cards were produced.

Each interviewer was instructed to commence his interviews with daily newspapers being asked about first in the first interview. He then had to rotate the order of the publication groups in subsequent interviews according to a fixed pattern. For an interview which commenced with questions about fortnightly publications, the sequence of the publication intervals would be fortnightlies, followed by monthlies, daily newspapers, weekly newspapers and ending with weekly magazines. Theoretically, this approach would result in closely similar numbers of interviews in which each publication group would be first, second etc in the sequence. The purpose was to average any 'order bias' which might exist.

Within each publication interval group, all the readership questions about the titles in that group were asked before proceeding to the next publication interval group of titles.

The sequence of questions in the readership section is listed below:

**1987/88 readership questions and sequence,
within publication groups.**

A six months filter question,
A frequency of reading question,
A place of reading question,
An origin of copy question,
A recency (last read) question,
A first reading question,
A currency of copy question, and
A thoroughness of reading question
(Newspapers only) Supplements: recency of reading

Readers per copy

As we had expected, the introduction of the 'first reading within issue period' approach to estimating Average Issue Readership did have an effect on readership levels and the numbers of readers per copy. Table 1 shows the readers per copy figures (rpc's) for a variety of publication groups, for AMPS 83 when the furore about the reported readership levels was raging, for AMPS 88/89 which was shortly after the 'first reading' method was introduced, and for AMPS 92, the most recent survey. To be noted are the differential effect on magazines and newspapers, and on publications which are intended for the Black sector.

Table 1 Readers per copy by publication group, AMPS 83, AMPS 88/89 and AMPS 92

Publication Group	AMPS 83	AMPS 88/89	AMPS 92
English dailies	5.8	5.2	4.7
Afrikaans dailies	5.6	4.2	3.9
Eng & Afr dailies	5.8	5.0	4.6
Eng weekly newsp	6.9	3.7	3.8
Afr weekly newsp	5.0	3.5	3.8
Black weekly newsp	6.5	4.8	5.9
Eng weekly mags	7.3	3.7	4.2
Afr weekly mags	6.9	3.8	4.0
Fortn mags	7.0	4.2	4.6
E & A monthly mags	5.7	3.1	3.5
Black monthly mags	11.1	7.9	10.8
All magazines	7.5	4.2	4.8
All reading	6.9	4.6	4.6

Credibility of the AMPS readership data

One of the realities which must be faced by the providers of media data, is the vital importance that their data be perceived to be credible. The results of the best designed and most meticulously executed surveys will not be used if the data are considered to be invalid or unreliable - and it does not really matter whether such perceptions are based on sound scientific reasons, gut feel, or result from self-interest.

Despite the fact that the down-scaling which resulted from the introduction of first reading was more marked in the case of magazines than of newspapers, and despite the fact that some of the readers per copy figures for individual titles remained high, there has been fairly general acceptance of the AMPS readership figures by both media owners and advertisers. The data are perceived to be close to reality (with some technical dissension) and have become accepted as the currency for buying and selling print space.

This situation has strengthened year by year to the point where one of the considerations, when the data of further development tests were scrutinised, has been whether data resulting from amended procedures remained close to existing levels, or at least did not increase them. Such considerations were not as unscientific as might appear at first glance. A considerable body of evidence indicated that if the data were erring, they were erring on the high side.

The acceptance of the readership data did not preclude us from undertaking a series of exploratory tests on possible changes to the readership questionnaire. These tests stemmed from two main considerations. The

first related to theoretical considerations; the second to puzzling aspects of readership data derived from the Black sector of our population. The latter problem is discussed in a separate paper.

Theoretical considerations

Quite soon after the first reading readership data were released in 1988, members of our Advisory and Technical Committees (which typically meet in joint sessions) expressed concerns about aspects of the procedure which was in use.

An analysis of the AMPS 1987/88 readership data again confirmed the 1986 finding that some of the variance was due to a rotation effect. Tabulations also showed that the desired balance of equal numbers of interviews which commenced with each of the five publication groups, had not been achieved. This second problem was immediately and effectively solved by revised administrative procedures. But the rotation effect was still there.

One needs to consider the practical interview situation. The procedure does not provide for an up-front filter which forces the respondent to claim, or not claim, readership of all titles in the questionnaire within the filter period of six months, before other more detailed questions are asked. The filter question is asked separately for each publication group. Therefore, it was argued, the less masochistic respondents would realise very soon that a claim to have read a publication in the filter period, resulted in a spate of further questions about that publication. Such a respondent thus had a reason to restrict the number of his filter period reading claims when he came to the second and later publication group titles. The analysis showed that there was a reduction in the number of titles which came through the filter in the second and subsequent publication groups. If the ultimate survey results were close to the truth, as seemed to be the case, then they were derived more by accident than by good design. And the correction of the imbalance in the number of interviews starting with each of the publication groups, spread this error variance more evenly, but did nothing to remove it.

The solution seemed straightforward: introduce an up-front filter question. At the same time, the re-introduction of individual masthead cards which had been used for AMPS up to 1986, and which are shuffled before each interview, would theoretically be a better system for randomising the sequence of titles for the filter question. From previous experience, there was also some evidence that the use of individual masthead cards would better retain the interest of the respondent who sorted cards onto one or more sorting boards, while the duration of the interview, at least for the more literate respondents, could be reduced. These changes should remove the need for the rotation by publication interval groups, also remove the rotation effect, and do away with the possibility that experience with the first group could condition the responses to subsequent groups.

From 1989 to 1992, three separate attempts were made to achieve a questionnaire which would incorporate these changes without drastic changes to the readership data levels. Collectively, these experiments became known as the **AMPS Readership Fine Tuning Experiments**.

The 1989 fine tuning experiments

For the 1989 Fine Tuning Experiments, two experimental groups 89A and 89B of 400 each, were matched to a control Group of 400 respondents from the AMPS 89 fieldwork. Each of the three main groups (the control group and two experimental groups) contained equal numbers of males and females, and of Blacks and Whites. The control groups, being part of the larger normal sample, were subjected to the entire current AMPS interview. The interviews of the experimental groups commenced in the same way as the main AMPS interview, with the cinema section, and terminated after the readership section which then followed.

The main differences between the interview for experimental group 89A and the standard interview, lay in the introduction of an up-front six months filter question, and a single random shuffle of the prompt pages instead of the normal rotation pattern. The questions in the remainder of the interview were asked within publication groups, which were rotated in the same sequence as for the matching AMPS interview.

The procedure for experimental group 89B was similar to that for experimental group A, but single title masthead cards were used instead of prompt pages. The cards were shuffled before the filter question was asked; subsequent questions were asked within publication groups, where the sequence was again the same as for the matching AMPS interview.

Based on a number of pointers, the assumption was that the AMPS data were if anything, an over-estimate of true readership. Therefore, any procedure which resulted in higher AIR's would probably represent a move in the wrong direction. The results of the 1989 experiments were to some extent predictable. Both experimental procedures resulted in higher proportions which came through the six months filter, and in

higher AIR estimates and higher readers per copy figures. Experiment 89B with the individual masthead cards, resulted in slightly smaller increases and was, of the two experimental procedures, the one on which further experimentation would be based.

Experiment 89B resulted in an overall increase of 24% in AIR estimates, with the increase being slightly higher for White respondents than for Black respondents. The increases to magazine AIR's were markedly higher than for newspapers. Across all titles, the average rpc from the AMPS 89 survey was 4.7. For experiment 89B the comparable figure was 5.9.

The results of Experiment 89B were obviously not of such a nature that they could be applied to the survey itself. We decided that a further experiment would be conducted in 1990. In the hope of obtaining approximately the same proportion of respondents who came through the up-front six months filter as those who came through that filter in the AMPS interview, it was decided to introduce a further filter before the six months filter question was asked.

The 1990 fine tuning experiments

Once more, two experimental groups were required. The sample sizes and the control group from the AMPS 90 fieldwork, were as for the 1989 experiments. For both experiments, individual masthead cards, shuffled up front, were used.

For Experiment 90A, the first question, which we referred to as the status deflating question, was as follows:

"Here is a pack of little cards. Each card has the name of a South African Newspaper or Magazine on it. I would like you to go through these cards and show me which of the statements best describes your own position for each publication".

The four statements which appeared on a sorting board, were:

- i *I do not know the publication - I have never heard of it.*
- ii *I have heard of the publication, but have never read or paged through it.*
- iii *I have read or paged through this publication in the past.*
- iv *I read or page through this publication either regularly or from time to time.*

All titles which came through the status deflating question (categories iii and iv) then went on to a six months filter question. The remaining questions were then asked within publication group, with the group sequence again matched to that of the corresponding AMPS interview.

Experiment 90B used largely the same procedure except that after the six months filter question, while the other questions were still asked within publication groups, the order of the groups did not change between interviews. The fixed order started with the longest and ended with the shortest publication interval. This procedure was suggested by Belson ('Studies in Readership', IPA, 1962).

The status deflating question was pre-tested for comprehensibility in a small set of 25 interviews of Black and a further 25 interviews of White respondents.

Overall, the Experiment 90A's levels were higher than either the AMPS 90 control figures or those of experiment 90B. This situation occurred across publication groups and for both black and white respondents, with only the black daily newspaper readership deviating slightly from the pattern. The all readership figure exceeded the AMPS control figure by 3.9% in the case of Whites, and by 5.7% in the case of Blacks.

For Blacks, Experiment 90B generally delivered lower readership levels than Experiment 90A, but with daily newspaper readership slightly higher. The Experiment 90B Black levels are generally higher than those of the AMPS control data. For all readership, the Black levels are about 5% higher than the AMPS levels, but the increase is much higher for daily newspapers - by 33%, and with a 17% increase for all newspapers. The level for all magazine reading shows a small decrease of 1.7%.

For Whites the AMPS control data and the Experiment 90B data agree closely. There is a difference of only 0.2%.

The status deflating question appeared to have worked well for whites with both newspapers and magazines. The Experiment 89B increase of 23.2% in black readership of all newspapers, reduced to 17% in Experiment 90B; the Experiment 89B increase of 24.8% for magazines, changed to a slight reduction in Experiment 90B. The status deflator question worked well for black magazine readers, but not nearly as well for black newspaper readers, where the Experiment 90B levels were unacceptable.

Newspapers vs magazine rpc's

Little mention has been made hitherto, of the concern which was expressed because newspapers had on average attained higher rpc's than magazines in the AMPS surveys since the 'first reading within issue period' method was introduced. For instance, AMPS 88/89 shows that English and Afrikaans daily newspapers had an average rpc of 5.0, while English and Afrikaans weekly magazines had an average rpc of 3.7.

In view of the questions about the earlier, higher rpc's for newspapers than for magazines, the relevant AMPS Committees were reluctant to consider any change which was likely to increase rpc's generally, and those of newspapers over magazines specifically.

Any decisions on major changes to the AMPS readership section were held in abeyance during 1991, when experimental work was devoted to an attempt to gain further insight into puzzling aspects of readership claims by Black respondents. That study is reported in a separate paper.

The 1992 readership fine tuning experiment

The results of the 1990 fine tuning experiments did not put an end to the conviction held by some, that an up front filter, preferably linked to the use of shuffled individual masthead cards, would represent a worthwhile technical improvement to the AMPS readership questionnaire.

The opinion had been expressed that the status deflating question in another format might achieve the type of desired results which its 1990 format had failed to deliver.

Therefore, as part of a larger experiment which is not reported here, the following revised status deflating question was used. It had five options compared to the four options in 1990:

- i *I do not know this publication.*
- ii *I have heard of this publication, but have never read or paged through it.*
- iii *I have read or paged through this publication in the past, but I no longer do so.*
- iv *I do read or page through this publication, but only occasionally.*
- v *I regularly read or page through this publication.*

Respondents who selected options iv or v went on to the six months filter question. Individual masthead cards and the 1990 fixed order of publication groups were used. The study was imbedded in an omnibus survey with 2508 respondents. As far as achieving data levels comparable to main AMPS was concerned, the experiment failed. The experimental results for English and Afrikaans weekly newspapers were much higher than from a comparable sub-set of 8913 respondents to the AMPS 92 questionnaire. For certain publication groups, the increases were more than 50%. Therefore, measured against the criterion of achieving data levels comparable to the AMPS questionnaire, the 1992 version of the status deflating question was not as good as the 1990 version.

Other changes

About eight years ago, the relevant AMPS Committees decided that changes to the media sections of the AMPS questionnaire would not be introduced before they, and their effect on the data, had been tested. Along with many other media researchers, we had learned that media research data tend to be volatile, and that apparently insignificant changes to questions and procedures, can lead to unexpected and exaggerated consequences.

Nevertheless, since then we have introduced, on judgement, a number of changes to the questionnaire and procedures which came into use in 1987. Thus far we have been fortunate in that our assessment that the changes were minor, were not confounded by results. The following are some of these changes.

Rotation

The change to the procedure for rotating the sequence of publication groups, has already been referred to. It was purely administrative in nature, was transparent to the respondent, and did not cause any perceptible changes to data levels.

The 'when last' or recency question

The first version of the first reading questionnaire contained different recency response categories for each publication group. Each of these questions presented the respondent with six time intervals to choose from. Three of the options fell inside the publication interval, and three outside it. The application of this variety of versions of the same question proved difficult and time-consuming in practice. In order to devise three questions which fell within the publication period of daily newspapers, the somewhat inelegant solution was to use 'yesterday morning', 'yesterday afternoon' and 'yesterday evening'. We now use the same nine-point recency scale for all the publication groups.

The initial questionnaire used different recency questions for daily newspapers, depending on whether the interview took place on a Sunday, a Monday or some other day of the week. That section, too, has since been simplified and standardised and the computer sorts our 'yesterday' readers.

The 'first reading' question

Mention has already been made of the fact that the first reading question which is used in the AMPS survey, goes more directly to the determination of first reading than the more roundabout but effective sequence which was used in the 1986 test.

The 'first reading' question which was used from 1987 to 1992, reads as follows, using fortnightly magazines as an example.

"You said that you read(title) within the last fourteen days. Please think back to the FIRST day you read this particular issue of(mention title). Was it within the last 14 days, or more than 14 days ago?"

A small scale qualitative investigation, which is also dealt with in another paper, indicated that some respondents had difficulties with the question. Consequently, the idea of getting at first reading within issue period via reading days found favour. If a publication is only read on one day, there can be no replication and the need for the first reading question falls away. The majority of reading claims refer to reading a particular issue on only one day. Therefore, we have introduced the following first reading question sequence into AMPS 93 (still using fortnightly publications as an example):

i Only in those cases where the answer to the recency question places the last reading of a title inside the issue period prior to the interview, is a 'reading days question' asked.

ii The 'reading days' question is worded as follows:

"Think back to the specific issue of..... (mention title) you read last. Please tell me whether you read that issue on one day only, or on more than one day."

iii If reading on more than one day is claimed, the 'first reading' question is asked:

"When did you read this particular issue of.....(title) for the FIRST time? Was it within the last fourteen days or more than fourteen days ago?"

We hope that the new procedure will have two advantages. If multiple reading days are not claimed, the need to determine the first reading event falls away. If multiple reading days are claimed, the question on the number of reading days should make the concept of the first of those reading days, easier to understand. Time will tell whether our optimism is realistic.

The 'reading days' question was used in a small scale test of its comprehensibility and appeared to work well. There was no opportunity to test it on a substantial scale prior to its implementation. The decision to use it, was based on assurances that this approach had worked well in projects which were undertaken years ago.

Evaluation

The failure of the 1992 fine tuning experiment to simultaneously maintain previous AMPS readership levels, and to remove the perceived weaknesses of 'learning' because of the lack of an up front filter, and thereby to remove the rotation effect by using shuffled individual masthead cards, caused us to re-evaluate the AMPS readership questionnaire in its operational format.

In 1992, a South African paper was presented at a symposium in Toronto, in which the results of three different methods of estimating television audiences were compared. Some of the results were illuminating. Remarkably high correlations indicated that the three methods were measuring the same behaviour, but at different currencies and thus at different levels. That finding prompted the calculation of correlation coefficients between sets of comparable data for different types of publications and for different race groups, using figures from the AMPS 92 readership questionnaire and the 1992 experimental study. Table 2 shows the results.

Table 2 Correlations between data from AMPS 92 and the AMPS 92 experiment

	All Races	WCA	Black
Daily English newspapers	0.98	0.98	0.98
Daily Afrikaans newspapers	0.98	0.98	0.99
Weekly English newspapers	0.96	0.98	0.91
Weekly Afrikaans newspapers	0.98	0.98	0.31
Weekly Non-White newspapers	0.97	0.97	0.98
English magazines	0.99	0.98	0.98
Afrikaans magazines	0.98	0.98	0.94
Non-White magazines	0.99	0.93	0.99
Daily newspapers	0.98	0.97	0.98
Weekly newspapers	0.96	0.98	0.95
Magazines	0.97	0.98	0.99
All publications	0.95	0.97	0.95

The correlations are very high. The variance, r^2 , of all titles and the total sample was 0.905. We are confident that much of the small variance which is not common, is due to the rotation effect which has been referred to several times.

The high correlations between the experimental data and the AMPS data are reassuring. The experimental data result from a methodology which is considered to be free from some possible technical weaknesses which are present in the AMPS questionnaire, and are based on a status deflating question which was developed experimentally over a number of years. The high correlations are interpreted as indicating that although the data levels are different, the possible weaknesses of the current AMPS methodology do not harmfully discriminate between titles and publication groups.

Which of the two is closer to the truth? We have some pointers, already referred to, that the current AMPS results are closer: firstly, they agree better than the experimental results with the maximum possible number of readers per copy, calculated from circulation data and origin of copy claims. Secondly, some users argue that a part of the AMPS readership numbers are on the high side. If so, then the experimental data are even further from the truth. The AMPS data are subjectively more credible, an aspect which we have learned not to ignore. For the time being, we will continue to use the current questionnaire.

