Pym Cornish
Research Services Ltd
London, UK

7.6 SECOND THOUGHTS ON VALIDATION

INTRODUCTION

National readership surveys are undertaken to provide buyers and sellers of space in print media with the information they require for the orderly transaction of business. To meet their needs the complex relationship between a publication and its readers must be represented by a small number of simple measures. These measures should reflect the features of readership that are most relevant to advertising decisions.

The ideal readership currency would be based on standard units of potential advertising exposure. That is to say, a reading claim by any informant should represent an equal probability that the informant would be exposed to a standard advertisement, whatever the publication for which the claim is made. Unfortunately, the direct measurement of potential advertising exposure is time-consuming and uncertain; it is not a practical possibility to do it for the large numbers of titles that national readership surveys must cover. We have to make do with indirect measures.

The measure that is actually used in place of potential advertising exposure in national readership surveys is Average Issue Readership (AIR). AIR is a measure of claimed exposure to issues of publications as a whole, rather than to their contents. It is therefore two steps removed from our ideal measure. In the first place AIR estimates will not correspond precisely to the actual number of exposures to the average issue. Secondly, actual exposure to an issue does not imply a constant probability of exposure to the advertising it contains.

As we know, different techniques of AIR measurement generate different levels of claims. We therefore have to be clear in deciding between alternative measures on what criteria they are to be judged. Should we be looking for the most accurate measurement of actual exposure to issues? Or are we looking for the best issue based predictor of potential advertising exposure?

If we choose the first of these possibilities, methods will be more or less valid according to the accuracy with which they predict issue exposure, measured or observed by some independent technique. If we prefer the second, validation will be concerned with advertising exposure; that is, not only with exposure to issues but also with the intensity of the exposures that occur.

The study reported in this paper approached this problem from the starting point of the AIR method now used in the UK. It examined the actual behaviour of each informant making a readership claim for the publications concerned. It found that AIR claims came closer to representing actual levels of potential advertising exposure than actual levels of issue exposure for sub-groups of the population defined by reading frequency and source of copy. While the study reported was a small one, covering only two types of publication, the results suggest that measured AIR automatically includes an allowance for variations in reading intensity. If this is accepted, it implies changes in the interpretation and validation of AIR estimates.

JICNARS MEANING OF READING STUDY, 1984

In May 1984 the Meaning of Reading Study Group of JICNARS carried out a study of the actual behaviour associated with AIR claims for Sunday newspapers and their supplements.

All informants for the month in question completed standard NRS interviews. At the end of these interviews informants were also asked in detail about their readership of specific issues of selected newspapers

and supplements published between eight and thirteen days before the interview. These questions generated specific issue or SIR readership claims. AIR claims and SIR claims could then be compared for various categories of informant, defined by demographics, source of copy and claimed reading frequency in the light of two different measures of reading intensity.

Table 1 shows the gross numbers of AIR and SIR claims made for the test publications, broken down by sex, claimed reading frequency and source of copy. In the table frequency claims are divided into regular - made by informants claiming to read a given title 'almost always, at least three issues out of four' - and 'other' - those claiming to read titles less frequently, or not in the past year. Source of copy claims are divided into primary claims - for copies bought by household members or delivered to the home - and other claims.

It can be seen that the SIR measure generated 4% more claims in total than the AIR measure. Arguably, part of this difference could arise from overclaims for the specific issues, but SIR claims were only accepted if they were supported by details of source of copy and place of reading. Differences by frequency and source of copy cannot credibly be explained in this way. There is clearly a tendency for regular and primary readers to make more AIR claims than SIR claims. There is a greater tendency for occasional and secondary readers to make more specific issue claims than AIR claims.

We now turn to the two measures of intensity of reading used in the study. In the first place all informants making SIR claims were asked how long in total they had spent reading the test issues. Secondly, they were asked to say for each spread, in up to two test issues per informant, whether they had read or glanced at it before, giving a measure of spread traffic.

Table 2 shows the results broken down in the same way as in **Table 1**.

The combined index is obtained by giving equal weight to time spent reading claims and spread traffic claims. It may be felt that each of these variables contributes to the probability of effective advertising exposure.

The table demonstrates that men have a somewhat greater mean intensity of exposure than women to the set of Sunday newspapers and supplements. These differences are very much greater for the other two breakdowns. Regular and primary readers have considerably higher levels of claimed intensity of exposure than occasional and pass-on out of home readers.

In **Table 3** we show the effects of weighting SIR claims by the combined index of reading intensity. The results are also compared with AIR claims.

TABLE 1
AIR and SIR claims for selected Sundays and supplements

	Total claims	*Sex*		*Frequency*		*Source of copy*	
		Men	*Women*	*Regular*	*Other*	*Primary*	*Other*
Gross AIR	859(%)	51	49	83	17	80	20
Gross SIR	896(%)	50	50	71	29	72	28

Base: 1,646 informants

The profile by sex of specific issue readership, weighted by reading intensity, is more male than the AIR sex profile. In turn, the AIR profile by sex is more male than the unweighted sex profile of specific issue readership. Similarly, the profile of intensity weighted readership by source of copy is rather more skewed towards primary readers than the AIR profile, which is considerably more skewed towards primary readers than the unweighted profile. Lastly, in the case of the frequency profile, weighted specific issue readership is closer to the AIR profile than to the unweighted profile.

Overall, for this group of publications, the AIR measure is a good approximation to intensity weighted issue exposure, which is a measure of actual potential advertising exposure. It does not correspond very closely to unweighted specific issue exposure, which is a measure of actual complete issue readership.

We may also note from **Table 4** that intensity weighting has little effect on the breakdown of all readership claims for Sundays and supplements by three categories of publication.

TABLE 2
Intensity of reading measures for SIR claimants

	Total claims	*Sex*		*Frequency*		*Source of copy*	
		Men	*Women*	*Regular*	*Other*	*Primary*	*Other*
Mean claims:							
Reading time (minutes)	37	39	36	42	26	44	22
Spread traffic (percentages)	70	74	66	74	60	79	51
Combined index (total = 100)	100	105	95	109	77	114	65
Bases:	896	448	448	640	256	641	255

TABLE 3
Intensity weighted SIR claims for selected Sundays

	Total claims	*Sex*		*Frequency*		*Source of copy*	
		Men	*Women*	*Regular*	*Other*	*Primary*	*Other*
SIR unweighted (%)	100	50	50	71	29	72	28
SIR weighted (%)	100	52	48	78	22	82	18
AIR unweighted (%)	100	51	49	83	17	80	20

Base: 1,646 informants

We may also show the comparison between SIR claims and AIR claims for the Sunday supplements only. In **Table 5** the frequency scale is further broken down to show the profile of informants claiming not to have read a publication in the last year in the main NRS survey.

As in the case of all Sundays, the intensity weighted profile is very close to the profile of AIR in terms of primary and other readers.

SUMMARY

(1) The AIR measure gave quite a good approximation to the *numbers* of specific issue readers of each publication category, and their breakdown by individual publication, sex and by other demographics. But this net total included overclaims by regular and primary readers and a similar number of underclaiming by occasional and secondary readers.

(2) Regular and primary readers were found to read with much greater intensity than occasional and secondary readers on the measures of reading time and claimed spread traffic.

(3) The propensity to make a reading claim varied proportionally with intensity of reading. Hence measured AIR was a good predictor of intensity weighted specific issue readership, or potential advertising exposure. But measured AIR was a poor predictor of actual exposure to complete issues between groups of intensive readers and groups of light readers.

TABLE 4
AIR and SIR claims by publication category

	Total	*Broadsheets*	*Tabloids*	*Supplements*
Unweighted AIR	859(%)	32	32	37
Unweighted SIR	896(%)	31	32	37
Weighted SIR	100(%)	31	32	37

Base: 1,646 informants

TABLE 5
Intensity weighted SIR claims for supplements

		Frequency			*Source of copy*	
	Total	*Regular*	*Less often*	*Not past year*	*Primary*	*Other*
SIR unweighted	332(%)	66	21	13	69	31
SIR weighted	332(%)	72	18	10	78	22
AIR unweighted	314(%)	81	19	-	79	21

Base: 1,646 informants

IMPLICATIONS FOR VALIDATION STUDIES

The implication of this experiment for partial validation studies is that they should be as concerned with the quality or intensity of contacts with issues as with the numbers of actual contacts that take place. That is to say, validation studies should place as much emphasis on AIR as a predictor of intensity weighted issue exposure as on its role as a predictor of unweighted issue exposure.

Unfortunately, the method used in the reported study cannot be applied to other publication categories, since other categories of consumer magazines undoubtedly have much longer active lives than Sunday newspapers and their colour supplements. One possibility is to concentrate on short recent time periods, as in the case of the MPX Study (1983) and the Claimed First Time Reading Study (1983) reported at Montreal.

The measures in such a study, based on all reading events in the past 24 hours, and possibly forming an independent part of a national readership study, might include:

(1) Identification of all publications, and all issues of those publications, seen at all through examination of day parts;

(2) Estimated reading time in past 24 hours per publication;

(3) Estimated proportion of spreads per issue opened in past 24 hours;

(4) Number of reading days per issue to date;

(5) Source of each copy;

(6) A re-check of frequency for each publication.

It is necessary to check frequency in order to cover publications which had not passed through the read-in-past-year screen in the main interview. We suggest that the details in points (3) to (5) should be collected for two issues per publication only.

If this secondary information was collected on a major national survey based on 20,000 or more informants a year, it would be possible to analyse it for the larger individual titles. Other titles would be grouped for analysis purposes.

Combinations of the data collected at the main and secondary stages of the interview would generate a number of measures of considerable interest. They would include:

(1) AIR per day, obtained by dividing AIR by the length of the publication interval in days;

(2) First issue readership, all issues not read before the past 24 hours;

(3) Total issue readership, all issues read in the past 24 hours;

(4) MPX, issues read in past 24 hours weighted by proportions of spreads seen, divided by first issue readership;

(5) Time exposure per issue, time spent reading in past 24 hours, divided by first issue readership;

(6) Total page exposure, issues read in past 24 hours weighted by proportions of pages seen;

(7) Total time exposure, time spent reading in past 24 hours.

Then the comparison between AIR and first issue readership is a guide to AIR as a predictor of actual issue contact. The comparison of AIR, broken down by frequency measured at the main interview, against first issue readership, by frequency measured at the supplementary interview, will show whether the patterns of overclaiming by regular readers and underclaiming by occasional readers also occur for

publications other than Sunday newspapers and supplements.

Lastly, comparisons of AIR, first issue readership, total page exposure and total time exposure, in each case broken down by reading frequency, will provide evidence to test the hypothesis that standard AIR is a better predictor of potential advertising exposure than of contact with the average complete issue.

REFERENCES

Douglas, Stephen and Lysaker, Richard (1983) 'Audience levels produced by the claimed first reading method' (Montreal Proceedings)

Gropp, Marvin (1983) 'MPX - A study of magazine page exposure' (Montreal Proceedings).