

6.1 Simulating readership data surveys in Great Britain – a review of current practices

INTRODUCTION

Simulate – to assume falsely the appearance or signs of anything; to feign, pretend, counterfeit, imitate; to profess or suggest anything falsely (*Oxford English Dictionary*).

This suggests that simulating data smacks of dishonesty. One is asking people to pay for data one did not collect in the first place.

So, before I begin to describe the various methods of *massaging* data currently used in Great Britain I want to make three points.

- (a) circumstances beyond the control of researchers forced them to embark on 'simulation'.
- (b) where 'simulation' has been used, every effort is made to acquaint users of the data with the facts, including a description of how the operation is performed.
- (c) I am solely concerned with 'simulation' of data collected on a continuous basis on two large-scale surveys in Great Britain – the National Readership Survey and the Target Group Index – where, during the course of the surveys, some real data exist which can be expanded to cover periods of interruption.

Having, I hope, covered myself adequately, I shall now discuss the three basic methods of *massaging* data currently used on these surveys which can be labelled – (i) adjustment; (ii) cosmetic weighting; (iii) ascription – and I shall not use the word 'simulate' again.

ADJUSTMENT

The National Readership Survey has been published in Great Britain twice a year since 1957 and is the accepted 'bible' for those interested in the print media. Its longevity is important since it antedated processing of data by computers. Net readership and all combination analyses were run laboriously on counter-sorters, but, since frequency data for all publications were not included until 1967, reach and frequency analyses, schedule building and optimisation models were unheard of.

In other words, the printed reports were the desired result. If some publication suffered industrial trouble and failed to appear for a time, then the Technical sub-Committee responsible for the control of the survey might consider basing their readership claims on a period during which no interruption had occurred. In other words, the printed reports might contain (and frequently did) data for titles based on nine rather than 12 months

fieldwork but expanded (ie multiplied by 1.333) to appear as a year's data. There was nothing underhand in this. It was widely agreed that, since the data are used for forward planning, one wants figures based on periods of continuous rather than interrupted publication. And, of course, any title treated in this way was clearly listed in the introduction to the report together with the fieldwork period on which the data were based.

But no adjustment was made to the tapes issued to the bureaux. Therefore, if the bureaux included in their analyses of a title so treated the quarter where the interruption had occurred, they would not arrive at the same totals as those published in the reports.

The occurrence of such interruptions prior to 1974 was not very serious, and certainly media planners – and the media owners themselves – were content to live with the situation. This changed drastically during 1974. During that year, of the 44 weekly periodicals covered by the survey, 23 suffered interruptions which, on average, amounted to 5.7% of their annual circulation. Of 55 monthly periodicals 12 suffered interruptions, averaging a loss of 19.0% of their annual circulation.

It was, therefore, decided that some action should be taken to adjust the data for publications which had suffered interruptions in their circulation. The method used to effect this is, theoretically, quite simple and is explained in the JICNARS reports. But the ground rules for deciding when to adjust are quite interesting:

- (1) adjustment would only be considered for a title if:
 - (a) at least one entire issue was lost; or
 - (b) partial losses of successive issues totalled to at least one entire issue.
- (2) only the titles which suffered losses would be considered for adjustment. In other words, no account was taken of the possible gains in readership which a rival title might win during the absence of the affected title.
- (3) adjustment would only be considered during the period starting on the day after the missing issue should have been published and finishing on the day on which the next issue was published.
- (4) adjustment would only be carried out if there was an apparent loss in Average Issue Readership.

Readers of this paper should bear in mind that the National Readership Survey in Great Britain uses the recent reading technique, and that a claim to have read a title may be based on an issue of any age – not just the most recent issue. Let us assume a very simple life model for a monthly publication which has a 'life' of four months and, during that period, accumulates eight readers per copy.

During April, one would measure eight readers per

6.1

Simulating readership data surveys in Great Britain – a review of current practices

TABLE 1

Month of Publication	Month of readership											
	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec			
January	1											
February	2	1										
March	2	2	1									
April	3	2	2	1								
May		3	2	2	1							
June			3	2	2	1						
July				3	2	2	1					
August					3	2	2	1				
September						3	2	2	1			
Total readers per copy	8	8	8	8	8	8						

TABLE 2

Month of Publication	Month of readership											
	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec			
January	1											
February	2	1										
March	2	2	1									
(April)	(3)											
May		3	2	2	1							
June			3	2	2	1						
July				3	2	2	1					
August					3	2	2	1				
September						3	2	2	1			
Total readers per copy	8	6	6	7	8	8						

in May, June and July and the position would not return to normal until August.

Although this is a childish simple model, and no publication would behave like this month after month, there is no doubt that the application of rule 3 does result in some cases of understating what the readership would have been in the months following the interruption of the circulation of a title.

The problem is that each title is a unique product, and behaves in its own unique way. Apparently similar copy, but five of these would be readers of earlier issues which they had read during April. The April edition would

get its final reader in July.

If the April issue were lost, adjustment would restore the three readers during the April period, so there would still be eight readers per copy measured in April. But following rule 3 above, May, June and July would not be scrutinised (assuming normal publication) and the pattern would be as **Table 2**.

The loss of the April issue would affect the readership publications, aimed at the same target, do not necessarily collect the same number of readers per copy, nor have the same reading life. It is difficult, therefore, to see how rule 3 could have been written in such a way as to cater for individuality of each title on the survey.

Let me now turn to the method of adjustment. First of all, the period of interruption is identified. This consists of data collected from the day after the title should have appeared to the day on which the title reappeared. The minimum periods of fieldwork scrutinised, therefore, would be one day for a daily newspaper, seven days for a weekly periodical and 28 days for a monthly periodical.

The AIR claims for the affected title during that period are compared with those during the previous 13 complete 'clean' weeks. By complete weeks, I mean seven days ending Saturday. In other words, if a daily paper failed to appear on a Thursday, the data collected on Friday would be compared with the 13 weeks ending the previous Saturday.

By 'clean' weeks, I mean consecutive weeks excluding any week in which adjustment had already occurred.

If the AIR during the affected period is less than that of the control period, then sufficient 'readers' are created to increase it to the same level. When creating average issue readers, only respondents who have claimed some positive frequency of reading the affected title are considered.

This probably works quite well for a publication whose circulation and readership are reasonably stable. But, consider a publication that is increasing its readership; it could happen that the claims during the affected period are as large or larger than those during the preceding 13 weeks. In such a case, no adjustment would take place (see **Figure 1**).

One might, from the curve, have anticipated a claimed penetration in week 39 over 4.0%. The observed figure is 3.7% but, since this exceeds the average for the preceding weeks of 3.0% no action is taken.

The same thing can, of course, happen in reverse when, during a period of declining readership, adjustment is made to bring the level for the affected period up to the average of the preceding weeks. Obviously, in both cases the result is less than satisfactory.

However, I understand that a working party is currently studying these problems.

6.1

Simulating readership data surveys in Great Britain – a review of current practices

FIGURE 1

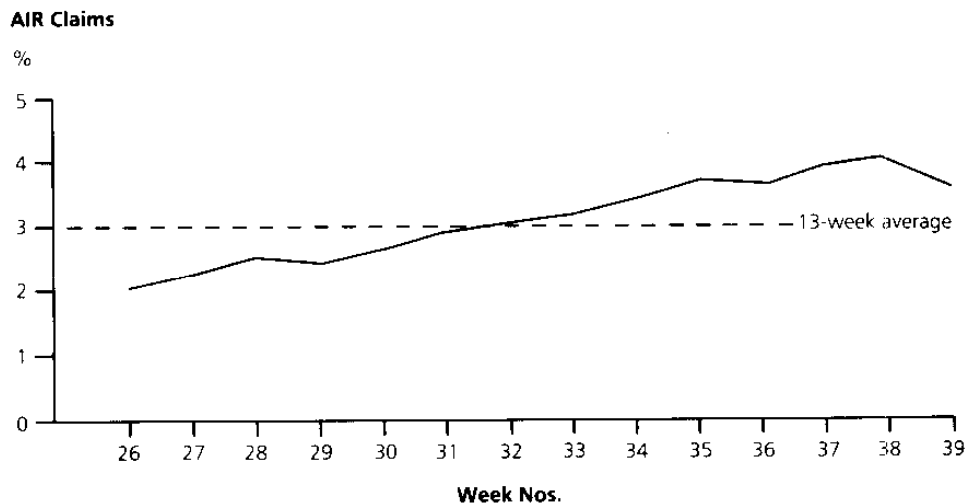


TABLE 3

	<i>NRS</i>	<i>TGI</i>
Sample size	30,000 adults 15+	24,000 adults 15+
Sample type	Probability sample drawn for the NRS	Random location omnibus sample
Response rate	c75%	60% of attempted placements
Interviewing technique	Face-to-face interview with mastheads	Self-completion questionnaire, no mastheads

COSMETIC WEIGHTING

Earlier in this paper I said that the National Readership Survey is the print-media 'bible'. The Target Group Index also collects data on readership, using the same recent reading questions as those in the NRS. There are, however, fundamental differences in the ways in which the data are collected (**Table 3**).

Not surprisingly, the two surveys produce different readership estimates. Perhaps surprisingly, the estimates are not all that far apart. During 1979/80, over 80 publications were measured on both surveys. A large number of those can be categorised as 'all adult' publications. Others are categorised as predominantly women's magazines or men's magazines. When we compared the results from the two surveys, we found the

results shown in **Table 4**.

I think you will agree that the results are not surprisingly different. Over half the titles measured on both surveys were within $\pm 0.9\%$ and three-quarters were within $\pm 1.4\%$.

However, readership figures are printed to the nearest thousand, and quoted that way (with no lip-service to sampling error). And 1% of all adults represents 430,000 readers!

It is not the purpose of the TGI to replace the National Readership Survey, but to provide product/media data. It was, therefore, agreed with the media owners to bring the TGI readership estimates into line with those of the NRS. This is what I term 'cosmetic weighting'.

I should say that, while in theory it is possible to force the desired result however large the disparity between the

6.1

Simulating readership data surveys in Great Britain – a review of current practices

TABLE 4

	Percentage	Penetration difference	
		All adult readership	Men or women readership
TGI higher	3.0 or more	1	3**
	2.0-2.9	6	1
	1.5-1.9	2	—
	1.0-1.4	8	2
No difference	0.5-0.9	8	4 } 6 } 1 } 4 } 5 } 20
	0.1-0.4	7	
	0	1	
	0.1-0.4	5	
	0.3-0.9	6	
TGI lower	1.0-1.4	1	3
	1.5-1.9	2	1
	2.0-2.9	2	1
	3.0 or more	1*	—
		50	31

* This magazine was not published for 12 weeks, so the comparison shown is between unadjusted TGI data and adjusted NRS data.

** These three are 'Skin' magazines when the face-to-face interview with a female interviewer probably results in underestimates of readership.

estimates, in practice it is only possible to do so without widely distorting the product and non-print media data in the process because the estimates from the two surveys are as close as they are.

The method used for this weighting involves the application of Taylor's Theorem for a function of many variables to the expression for the item totals in terms of a given set of survey data and initial weighting factors. (These initial weighting factors comprise projections to population totals for sex within age with regions.) If there are N items to be weighted, a set of N linear equations is obtained from which a new set of item totals can be calculated. This provides a new set of equations from which yet another set of weights can be calculated, and so on.

The method provides an iterative convergent process, provided certain precautions are taken:

- (a) the calculated differences must be small for Taylor's Theorem to hold true. This is ensured in the program by restricting the differences at each stage to such a magnitude that the process remains stable.
- (b) only positive weights are generated.
- (c) the iterative process tests the sum of the squares of

the differences between the observed item totals and the desired 'set' totals. The end of the process is reached when the differences between these values are negligible. (d) the items chosen must be independent so that, for example, if a set of items sum, say, to the universe (or any other total), then a second such set must not sum to the same total. If this restriction is not observed, then the N equation will not be linearly independent and a solution cannot be obtained.

Obviously, the larger the number of items, the larger the number of equations and, therefore, the longer the operation becomes. We set a limit of 140 items within each sex. Of these c 55 are demographic items, c 60 are readership items common to both sex and c 25 are readership items set for men or women only.

The result is that each respondent in the sample carries his own individual weight. This weight applies to all the data for that respondent – readership, product usage, leisure activities and so on – not just to the data we have set out to control. In theory, because of the constraint we place on the maximum weight applied, these weights could range from 0.10 to 19.99. In practice, 90% of respondents have weights ranging from 0.50 to 4.99.

This might appear to be an alarmingly wide range of weights. However, the survey has now been published for 13 years and this system of weighting has been applied for 12 of them. During that period, not only have subscribers seen readership figures with which they are familiar (the whole point of the exercise) but the product data have matched very closely those from other sources, or, where improvement was required, this was remedied by altering questions rather than any concern with the weighting. Indeed, they also match closely the estimates from the TGI before this stage of weighting is applied.

ASCRPTION

For media purposes, the value of the TGI is the direct link between media exposure and product usage. Between one TGI report and the next, there will be three publications of revised readership estimates from the NRS. It is fairly common practice for users of the TGI to take account of the revised readership estimates and apply these to schedule analyses of TGI data.

It is uncommon, however, to look at TGI data for a period of less than one year. There are two reasons for this:

- (a) the cost of computer time is relatively low – therefore, the difference in cost between, say, one quarter's and four quarters' data is insignificant.
- (b) the reason for using TGI rather than NRS is because one can select a target group based on product or brand usage. Therefore, to include sufficient users in the base

6.1

Simulating readership data surveys in Great Britain – a review of current practices

will generally mean using the full sample rather than a part of it.

Between November 1978 and November 1979, three important titles were not published – *The Times*, *Sunday Times* and the *Sunday Times Colour Magazine*. This caused a fair amount of chaos in the readership of the quality daily and Sunday press.

Bearing in mind that the TGI fieldwork period runs from April to March, one can see that, for the 1979 report, we had eight months 'good' data for these titles before they ceased publication. We had, obviously, no idea when they would start publishing again. We had to assume, however, that this would occur during the 'life' of that TGI report, ie between July 1979 and July 1980. To omit them altogether from the report would mean reducing the value of the TGI not only to Times Newspapers, but also to their rivals.

Publishing data on the printed reports was not a problem. One needed only to expand the data from eight to 12 months. This could easily be done across all product fields. Such action, however, would be of no benefit to tape users who would find that one third of the year carried no data for these titles.

We therefore decided to ascribe readership claims to people interviewed during the period when the titles were

not on sale. This ascription was to cover not only an estimate of average issue readership, but also of claimed frequency of readership.

The method we had decided to adopt involved matching respondents interviewed between April and November with those interviewed between December and March and then, having found our 'pairs' reproducing the frequency and average issue claims for the three titles from the first group to the second. The problem was how to control the selection of respondents.

There were obvious socio-demographic biases in the readership of the three titles (**Table 5**).

But, in addition, we needed some control which would match approximately the same lifestyle in the two parts of the sample – type of home, leisure habits, durable ownership and so on.

The list of such items from which we could choose was large and there was the problem that we might have more control cells than respondents. For the period of 'good' data, we had approximately 16,000 respondents and, among these, we were likely to find the results given in **Table 6**.

In other words, less than one respondent in five would read any of the titles and we needed to find a 'match' for half of these among the remaining 8000

TABLE 5
Profiles

	Readers of			
	<i>All adults</i> 100%	<i>The Times</i> 100%	<i>Sunday Times</i> 100%	<i>Sunday Times Magazine</i> 100%
Sex				
Men	48%	60%	53%	53%
Women	52%	40%	47%	47%
Age				
Under 35	37%	46%	47%	50%
35 or over	63%	54%	53%	50%
Social grade				
AB	16%	48%	43%	41%
C1C2DE	84%	52%	57%	59%
Household income				
£7500 or more	11%	36%	30%	28%
£5500–£7499	13%	15%	20%	21%
Under £5000	60%	40%	38%	41%
(Not stated)	(16%)	(9%)	(12%)	(10%)

6.1

Simulating readership data surveys in Great Britain – a review of current practices

TABLE 6

	Respondents who	
	Ever read	Read the average issue
The Times	c1100	c 350
Sunday Times	c2500	c1350
Sunday Times Magazine	c2800	c1500
Any of the three	c3000	c1900

respondents (the last four months of the survey).

We had recently been experimenting with ACORN* A Classification of Residential Neighbourhoods. Briefly, this was a 36 cluster solution of 40 variables, data collected on the 1971 census of population. The original purpose was to identify types of small administrative areas – wards and parishes – with varying degrees of social deprivation, such as overcrowding, high unemployment, lack of sanitary amenities, immigrant areas and so on. Obviously, when identifying such neighbourhood types, the more affluent areas were also identified.

Since our sample was based on wards, we were able to apply the cluster codes to our sample and then see how they helped discriminate between levels of product usage and media exposure. In a large number of cases, it proved to be very powerful indeed – not least for the quality press. I have appended a table showing an analysis of the readership of *The Times* and *Sunday Times* by the 36 clusters (Appendix A). In summary, this is shown in Table 7.

*(1) *The utility to market research of the classification of residential neighbourhoods* – Ken Baker, John Birmingham and Colin MacDonald (BMRB) – 1979 Market Research Society Conference Papers.
 (2) ACORN – A Classification of Residential Neighbourhoods – Baker, etc. ADMAP May 1979.

TABLE 7

	Proportion of adult population 100%	The Times 100%	The Sunday Times 100%
I 25, 28-30, 32, 34	19%	31%	36%
II 1, 2, 5, 7, 10, 11, 13, 15, 18, 20, 21, 26, 31, 33, 35, 36	43%	51%	42%
III 3, 4, 8, 9, 12, 16, 17, 19, 22-24	39%	18%	22%

Group I contains the six clusters where the readership penetration of one or both of the papers is significantly higher (at the 99% level of confidence) than the national average and Group III contains 11 clusters where one or both is significantly lower.

Since the cluster solution also discriminated between users of many products and services – holidays, table wine, banking, credit, high priced durables, etc – we decided to use combinations of clusters rather than individual products and services as a further control.

We therefore constructed a 240 cell matrix.

Sex	2	within
Age	2	“
Class	2	“
Income	3	“
Acorn clusters	10	“

The two parts of the sample were divided into these cells and the penetration of the titles was discovered in the first part of the sample. The number of respondents to whom readership should be ascribed in each cell in the second part of the sample was then calculated. Using a random number generator, the required number of respondents for each cell was selected from each part of the sample and, after matching, the readership data were reproduced from the first to the second part.

In general, this worked very well. The demographic profiles, ownership of products and uses of services for the two parts of the sample matched quite closely. Unfortunately, the story does not end there. The titles recommenced publication during November 1979. For our 1980 report, therefore, we had four months real data only, and we undertook to ascribe data to the other eight months. In the event, our segmentation into two class groups – AB-v-C1C2DE – was not good enough. Too many of the ascribed readers came from the C2's rather than the C1's. This had the effect of apparently weakening the case for selecting these titles rather than their competitors on the basis of the published product/media data. (Data illustrating this are available if requested.)

6.1

Simulating readership data surveys in Great Britain – a review of current practices

CONCLUSION

As I stated at the outset of this paper, no researcher wants to generate estimates as a substitute for real data. It is only in the conduct of continuous surveys where the historic data will be used for forward planning that it may be necessary to resort to adjustment or ascription. I am less

concerned about cosmetic weighting because, as far as the British TGI and NRS are concerned at least, one knows that it does not distort the data while providing figures with which the industry is familiar. But it is true that both the adjustment on the NRS and the ascription on the TGI, when they are necessary, could use some improvement.

APPENDIX A

Readership penetration of *The Times* and *Sunday Times* within ACORN clusters – TGI unweighted data, April 1977–March 1978

Readership						Readership					
Cluster number	Actual sample	The Times %	The Times index*	Sunday Times %	Sunday Times index*	Cluster number	Actual sample	The Times %	The Times index*	Sunday Times %	Sunday Times index*
1	546	2.9	91	8.4	89	19	1203	1.4	44 ⁻	3.8	40 ⁻
2	360	1.9	59	6.4	68	20	263	2.7	84	8.4	89
3	458	1.1	34	4.8	51 ⁻	21	275	2.2	69	6.5	69
4	1211	1.4	44 ⁻	5.0	53 ⁻	22	421	0.2	6 ⁻	8.6	91
5	1112	2.7	84	10.5	112	23	279	0	0 ⁻	2.9	31 ⁻
6	684	3.1	97	10.2	109	24	178	0	0 ⁻	8.4	89
7	63	1.6	50	9.5	101	25	490	7.8	244 ⁺⁺	15.3	163 ⁺⁺
8	1688	2.5	78	6.7	71 ⁻	26	137	5.8	181	12.4	132
9	1236	1.5	47 ⁻	6.8	72 ⁻	27	464	1.9	59	5.8	62
10	1391	3.0	94	9.1	97	28	409	19.6	612 ⁺⁺	35.2	374 ⁺⁺
11	231	0.9	28	6.5	69	29	754	6.9	216 ⁺⁺	18.4	196 ⁺⁺
12	1058	2.0	63	5.7	61 ⁻	30	1369	4.4	138	12.6	134 ⁺⁺
13	774	3.2	100	7.8	83	31	1320	3.6	113	11.4	121
14	525	4.0	125	10.3	110	32	942	4.5	141	13.4	143 ⁺⁺
15	330	4.2	131	10.6	113	33	385	3.6	113	9.4	100
16	803	0.9	28 ⁻	4.4	47 ⁻	34	523	8.4	263 ⁺⁺	26.8	285 ⁺⁺
17	605	0.5	16 ⁻	4.1	44 ⁻	35	437	3.4	106	11.7	124
18	584	2.1	66	6.0	64	36	190	3.7	116	10.0	106

* The indices are derived by the formula $n \div N \times 100$, where n is the percentage penetration of the title in a cluster and N is the national penetration of the title. The indices are based on actual (unweighted) data. The national penetrations were:

The Times -- 3.2% Sunday Times -- 9.4%

++ and -- denote penetrations within clusters which were significantly different from the national level at the 99% confidence level.