Michael Brown
Consultant
London, UK

4.1 Looking back at UK frequency scale development

This paper examines – necessarily briefly – three phases in the British development of reading frequency scales; fuller details are provided in the IPA publication covering the experimental research on which this paper draws. (1)

To put this work into context, it should be borne in mind that Britain has never gone far along the road of using frequency claims as a basis for average issue audience estimates – although such application of the data was discussed in the early days. The aim is therefore limited to obtaining probabilities for use in projecting cumulative audiences and frequency distributions of issue contacts. For this purpose, a reasonable viewpoint is that it is most important for a frequency scale reproducibly to segment the population into sub-groups which are as homogeneous as possible relative to their true reading behaviour. The relationship of the observed probabilities to the nominal labels on the scale positions is of secondary importance, seen from this viewpoint.

Reverting to the UK programme of developmental research, 20 exploratory, extended interviews were carried out in 1964. The objective was to see whether, in readers' own words, there appeared to be any habitual way of thinking about frequency of readership, which could provide indications of how a scale ought to be designed, if it was to be intelligible to most people.

Some of the findings from this stage are well worth repeating, since they underline the basic difficulties in this area of work. At the same time, it is to be noted that there may well be genuine differences between countries: how people talk about their reading behaviour does depend on factors which vary market-by-market. For example, there were no German- or Dutch-type reading circles in Britain in 1966, no more than there are today; again, the UK ratio of news-stand sales to subscriptions was and is very different from the American figure.

From the results of these exploratory interviews was observed, first, a universal tendency for readers to talk in terms of a time scale – either in the form of an actual estimate of frequency, such as 'twice a week', or verbalised in one of the conventional expressions of frequency over time, such as 'occasionally' or 'sometimes' or 'rarely'.

Secondly, there was a strong tendency to speak in terms of reading *occasions*, rather than of copies read. For magazines and excluding, obviously, very regular readers and primary household members – awareness of *publication* frequency was low. For monthlies, the problem was particularly acute. Quite a lot of readership was spoken about in relation to other events – going to the hairdresser, for example, or visiting with relatives or friends. Further questioning on such reading occasions made clear – not surprisingly – that the occasions could not readily be translated into copies seen: even when only one issue was involved, it was not always the current one. Quite a substantial proportion of monthlies' readers, it was therefore concluded, were probably just not able to say what proportion of *copies* they saw, however skilfully a frequency scale might be designed – they would always be obliged to answer in terms of the frequency of their reading *occasions*.

The second phase of the work described here took place in November 1964. It comprised the testing – and evaluation, rather than 'validation', in a strict sense – of four possible scales, partly derived from the exploratory work, partly based on judgement.

The first was a verbally expressed time scale. For example, for monthlies the question ran "I see . . . . . . every month/most months/every few months/once or twice a year/never."

The second experimental scale was purely numerical. For example, for weekly magazines the wording was "In an average month I see this number of copies: 4/3/2/1/none."

The third alternative was an abstract, Stapel-type scale with nine points, anchored at "Every copy" at one end and "Never see it" at the other.

The fourth experimental scale was purely verbal and common to all publication groups: "In the last few months I have seen nearly every copy/most copies/a copy now and then/hardly ever see it/not seen it."

One or other of these four scales was applied to 44 titles on a sub-sample of 100 informants. Then in the second part of the same, semi-structured interview, reading behaviour was probed for just one out of the four test titles – a daily newspaper, a Sunday, a weekly magazine and a monthly; the information collected here included the dating of the last reading occasion.

From these data, each scale was assessed on four criteria: the shape of the frequency distribution of respondents across scale points and the 'separation' of the points; the extent to which the scales seemed to conform to respondents' own views of their reading habits; whether the scales could stand on their own as providing an assessment of reading probabilities; and the apparent accuracy of the frequency claims. In this latter respect, the data from the unstructured part of the interview were used to rate an initial frequency claim as 'very likely to be correct', 'fairly likely', 'fairly unlikely' or

4.1 Looking back at UK frequency scale development

'very unlikely'. Clearly, subjective interpretation was here involved, to some degree at least.

The detailed results from this second phase of the research programme will not be covered here, but are contained in the IPA report cited earlier; but three general points are noteworthy.

First, all of the four experimental scales produced credible patterns of responses – with the possible exception of the purely verbal one, where there tended to be an uncomfortable 'hump' in the middle of the frequency distribution, breaking the usual U shape – but this might have been due to inelegant choice of the labels on this scale's positions.

Second, each of the four alternatives seemed to have its advantages and disadvantages – and (again not surprisingly) problems were greater with the two test magazines than for the newspapers and particularly so for the monthly.

Third, it was decided, on balance, that the numerical scale offered the best possibilities – although, on a personal viewpoint it does seem that, rereading the evidence, the decision must have been a narrowly balanced one.

Two events followed. The numerical scale was added to the main NRS questionnaire from May, 1965, for some, but not all, of the titles then covered. Two modifications were made to the experimental version: a 'less than one' position was added between 'one' and zero, for daily newspapers and weeklies, to bridge a rather wide gap of nominal probabilities; and the monthlies' scale was extended from five points to seven, by making it refer to the number of issues seen in the past six months, rather than the last four.

The other event was a relatively extensive validation of the numerical scale, with which study the remainder of this paper is concerned. The fieldwork took place in June 1965.

The interview comprised two parts. First, the numerical frequency scale – varying between publication groups – was applied to 29 titles; effectively, there was no preliminary screen-in question. Second, each informant was shown four consecutive issues of each of two test titles – the same daily newspaper, Sunday, weekly magazine and monthly used in the earlier stage of research described above. His or her readership of these issues was checked using the editorial interest approach on full issues, not skeletonised ones.

The test titles to be shown to a given respondent were selected on the basis of a fairly complex grid to ensure that, as far as possible, adequate numbers of middle-of-the-scale frequency claims were checked.

The results from this study can be examined in a number of ways; space allows only the selection of a few which seem important, on personal judgement.

First, it can be asked how frequency claims – on the numerical scale – stacked up against frequency assessed via the TTB measurement of actual issues. These data are shown in **Table 1** for just one of the test titles – the weekly magazine.

As can be seen from the table, the general correlation between claimed frequency and what might be termed 'through-the-book frequency' is reasonably good – particularly as judged from the top and bottom lines; but, looking down the columns of the table, if the TTB data are taken at face value, it is clear that the people claiming, for example, three issues in the average month hardly comprised an homogeneous group in terms of their actual behaviour. In this respect, the results for the newspapers tested were, generally speaking, better than the ones shown here, while those for the monthly – the Digest – were worse.

A second way of looking at the data is to compare the

TABLE 1
(publication: Woman)

	frequency claim (issues in an average month)					
	4	3	2	1	<1	0
Unweighted base = 100%	180	83	159	142	66	341
TTB claims (issues out of 4)	%	%	%	%	%	%
0	9	23	32	60	71	95
1	4	22	21	22	25	3
2	4	19	27	9	2	1
3	9	19	9	5	—	—
4	74	17	11	4	1	2

TABLE 2
(publication: Woman)

	Frequency claim (issues in an average month)					
	4	3	2	1	<1	0
'Nominal' probability	1.00	0.75	0.50	0.25*	0.13	0.00
TTB based probability	0.84	0.47	0.36	0.18	0.08	0.03

* Calculated as 0.5/4.0.

probabilities implied by the frequency scale positions, taken at face value, with the figures derived from the TTB results. **Table 2** shows this comparison, again for just one of the four test titles.

It should be borne in mind that the empirically based probabilities are, in effect, weighted averages with fairly small bases and thus subject to relatively wide confidence limits. Nonetheless, the evidence for material *over-statement* of frequency is strong – unless, of course, the TTB levels themselves were far too low (as some contributors to this symposium would aver).

Taking a third view of the data, three different average issue audience estimates can be compared: a recent reading one, from the NRS data themselves for the relevant period; and two frequency based estimates, from the frequency scale responses themselves and from the TTB results. These data are shown in **Table 3** for all the four titles tested.

The correspondence between columns 1 and 3 is very close except for *Reader's Digest*. Taking that case first, there are two possibilities. The recent reading level could be an overestimate (although, to be fair, Belson's evidence would point in the opposite direction).

TABLE 3
Average issue readership

	Recent reading* %	Frequency based nominal (scale) probabilities %	TTB probabilities %
Daily Express	30.8	33.7	30.6
The People	39.3	43.5	39.3
Woman	27.3	33.3	27.9
Reader's Digest	20.5	20.6	14.5

* *(NRS Jan June '65).*

TABLE 4
(publication: Reader's Digest)

	TTB based probability
Average, four test issues	0.145
Average of issue*	
1	0.144
2	0.139
3	0.146
4	0.151

* *In issues periods, 'O' representing the current issue.*

Alternatively, the TTB figure could have been low; some supplementary evidence on this point is shown in **Table 4**.

The differences between the probabilities shown above are not significant, due to small bases; but they provide at least a suggestion that issue age *is* a problem, despite the Allensbach results quoted at this symposium, and that, in the research reported here, one, two or even three month old issues of a monthly were too 'young'.

Reverting to the titles other than *Reader's Digest*, **Table 4** illustrates a case on *one* TTB method (not 'the' method, which doesn't exist) virtually in complete accord with *one* RR approach; which, on personal judgement, is a rather interesting by-product of what started out as an examination of British frequency scale development.

REFERENCE

1 Corlett, T and Osborne, D W (1966). *The development of reading frequency scales* London: Institute of Practitioners in Advertising.