Uwe Czaia
Czaia Marktforschung, West Germany
Rolf Speetzen
Axel Springer Verlag AG, West Germany

9.4 CURRENT TRENDS IN THE WORK OF THE AG.MA

SECOND THOUGHTS ON FREQUENCY SCALES AND MULTIPLE PAGE EXPOSURES – NEW EXPERIMENTS IN GERMANY

Introduction

The national Media Analysis has initiated a number of experiments to adjust it as *the* media planning tool to the changing needs, to move the establishment of media consumption nearer to reality. Some of the results, as well as the results of additional surveys sponsored by some publishing houses in close co-operation with GWA, a German version of the 3As, were transformed into an establishment tool within the MA, especially for the so-called partnership model.

In addition we noticed intensive discussion mainly within committees and bodies outside the MA on how to align the 'Standard' adequately to the modified requirements of the media users as well as of the media owners.

The demands on the concept of the MA were called for, by agencies and clients regarding the present methods and systems earlier this year, establishing the basis for further discussions in the near future. The voluminous catalogue covers expectations in relation to:

– the 'technique' of data establishment and the 'practice' (for example the construction of samples, weighting, correct representation of smaller regions)

– the establishment of media users and the calculation of probabilities

– qualitative descriptions of userships, ascertaining indicators for exposures to the advertisement.

For each of the topics expert ratings and experiments were proposed the first results of which are published. We would like to restrict this paper to two aspects:

– reading frequency

– amount read, multiple page exposures.

Reading frequency

Originally the MA used a 12-point scale as a prompt for estimating the subjective reading frequency of the respondents:

Of the last 12 issues of this publication I have read or leafed through ...

1 only one issue
2
3
4
5
6
7
8
9
10
11
12 all issues

The result of using this scale lead to the conclusion that it could not reflect real media consumption patterns, or in other words, it required an abstraction level far too high for normal respondents. The odd fields were claimed far below average. Prompt categories nine of 12, or 11 of 12, attracted hardly any responses.

Without further analysing the causes of these results through experiments into semantic comprehension, the ability of recollection or the active and passive stock of words, the problem of inconsistent claims was (presumably) solved by a contraction into a 7-point scale:

1 (only one) issue
2 – 3 issues
4 – 5 issues
6 – 7 issues
8 – 9 issues
10 – 11 issues
12 (all) issues

This means that the prompt aid was changed, but the questioning remained the same, a copy of the previous one. Thus in principle all results remained the same.

In addition to the poor filling-in of the odd categories we find below average response distributions in the centre of the scale. On top of this, there are extreme differences between the frequency based and calculated AIR (K1) and the recency based AIR (LpN), depending on the media categories. This is further evidence for the fact that the 12-point scale is limited as an adequate tool for establishing the frequency of reading. The lower the relation between reader and publication, the lower the reader loyalty, the more inaccurate or more inexact the claims.

We should learn from this and should not ask for more differentiation than can be achieved.

Apart from this we have to consider the different dimensions of questioning. On the one hand we try to establish reading within the last publishing interval, on the other hand we trace the frequency of reading over 12 times the publishing interval. K1 and LpN drift apart when the reading of a publication is increasing or decreasing, if the one score refers to just the last publishing interval and the other refers to 12 times this interval. Therefore it is an important to bridge the gap between reading frequency establishment and time of the interview. But K1 and LpN are drifting apart in the results, because: 'How many out of 12 issues were seen', can only be rated as a *frequency estimation*, which is therefore formulated in the question wording.

On the other hand, the claim of reading within the last publishing interval, the recency method, is an attempt to recall actual behaviour in an exactly determined time span, and to reflect this actual behaviour. Therefore it seems more than appropriate to move towards a recollectable behaviour in establishing the reading frequency, to determine the reading within the last four publishing intervals and not twelve.

If this is correct it should show in the results. The differences between K1 and LpN should decrease by using a 4-point scale instead of the usual 12-times the publishing interval.

To demonstrate this we have topline results from two representative surveys. One was conducted by Infratest and sampled in the Düsseldorf area. It was sponsored by Axel Springer Verlag. The other was conducted by

Contest-Census and Infratest in Franken (Franconia) and sponsored by Süddeutscher Verlag. The fieldwork period of both the surveys was spring/summer 1988. The main results were analysed and commented on by Siegfried Geiger. Our remarks are based on his report.

Both studies include results for monthlies, weeklies and dailies. The broadest audience (read or leafed through at least one issue of the title in question within the last 12 publishing intervals) was first confronted with the usual 12-point scale for frequency estimation. Immediately after that question and still before establishing the average issue readership they were confronted with an additional 4-point scale, which reads for example for weeklies:

And how was this in the last four weeks? How many of the four issues of ... which appeared within the last four weeks have you read or leafed through?

0 none
1 one issue
2
3
4 all issues

In this case we need 0 as a separate category, because the frequency estimation was determined among the broadest audience. Both surveys make evident that the differences between LpN and K1 are decreasing if we try to determine the reading of four issues within four times the publishing interval instead of the usual 12 publishing periods. This is true for monthlies, weeklies and dailies! (Table 1).

The total is based on 12 titles. Table 2 should demonstrate even more clearly to what extent the differences between the two readership claims LpN and K1 can be diminished by changing the frequency question and moving it towards a reflectable time span.

Table 1

Differences between LpN and K1 (in broadest audience) – weekly magazines

Title	LpN %	K1 12-point %	K1 4-point %
A	45.6	51.1	48.6
B	57.6	61.6	57.6
C	41.3	47.7	43.1
D	45.9	49.1	48.7
E	78.2	81.3	78.9
Total	61.0	66.1	63.1

If we choose a stringent yardstick for the consistency of the responses and exclude for example all those who claim 10-12 of 12 in the first case and 'none' or 0 in the second case, the decrease of the differences is not quite so extensive but still significant. We see for example a reduction from 16 to 12 in the case of daily newspapers.

For comparison the questions with the 4-point scale were exactly the same for all publishing intervals, for weeklies, fortnightlies and monthlies in these experiments. The covered time span was thus:

1 month for weeklies

2 months for fortnightlies

4 months for monthlies.

One month, two months or eight weeks are somewhat defineable time spans. However, four months is an extremely unusual period.

Table 2

Differences between LpN and K1 (Indices)

	Index K1/LpN	
	12-point scale	4-point scale
20 monthlies	18	3
12 weeklies	8	3
21 dailies	16	-1

Index:((K1 – LpN)*100)/LpN

Therefore one should consider different frequency scales (4-point and 6-point scales) for further developments of this method. This is not disturbing, because the reading frequency is just one aid for segmentation and cumulation, to achieve the best differentiation based on the selection criterion LpN. In addition it is desirable to keep the differences between K1 and LpN as small as possible.

The 4-point scale was tested for daily newspapers in the survey in Upper Franconia. However, the recommendation for the MA of the 90s provides a 6-point scale for dailies, covering the previous week ie the six workdays prior to the interview.

SUMMARY

A shortened frequency scale related to a time span which the respondents are able to remember can contribute towards diminishing the differences between K1 and LpN. In addition to this aspect on the data processing side, the shortened frequency scale has another advantage, namely streamlining the interview to make the interview easier for the respondents, thus reducing resistance in answering the questions.

Amount read – multiple page exposures

One of the central aspects of the MA 90 concept can be seen in the establishment of indicators, which supply hints going beyond the usual vehicle contacts towards advertisement exposures or advertisement contact dosages.

The establishment of advertisement coverage within the existing or future national media analyses with the primary task of measuring exposure to individual advertisements seems to be hopeless if we consider the practicability and the feasibility of the task, unless the methods and data processing are completely changed. The usual units such as average issue readership will be called in question.

The discussion and the experiments within the MA (results will be published after this Symposium) as well as additional experiments such as 'Yardsticks for exposure quality' restrict the topics to two dimensions: amount read and multiple page exposure, thus coming closer to the aim of measuring advertisement contacts, which is in principle possible for TV, utilising the meter system.

The effectiveness of questions regarding the amount read and the multiple page exposures was tested extensively within the scope of a joint venture between the German 3As (GWA) and some 20 publishing houses, who sponsored a follow-up of 'Yardsticks for exposure quality', a survey which was presented at the Salzburg Symposium.

In addition to the questions determining media exposures, broadest audience, frequency estimation and AIR, the exposure qualifying criteria were established for all titles within the broadest audience. Furthermore, the respondents were asked to state for all titles of which they qualified as broadest audience: the issue

last read, and how much of that particular issue was read in total, how much was opened twice or more times. All these questions followed immediately after the establishment of the vehicle contacts. Here is the question phrasing in an English translation, (perhaps this translation cannot quite keep the tone of what is between the lines).

"It happens that one has not always the time and opportunity to read a magazine as much as one would like to. Of the last issue of ... which you have finished reading or which is not at your disposal anymore, how many pages of the total have you opened altogether to read or to look at something?"

The prompt aids were as follows:

(1) Almost no page opened to read or to look at something

(2) Only a few pages opened ...

(3) A quarter of the pages opened ...

(4) Half of the pages opened ...

(5) Three quarters of the pages opened ...

(6) All, almost all pages opened ...

The question on multiple page exposures is different from the well-known MPX approach:

"While reading or paging through a magazine, some pages are not opened at all, some pages one only sees once, some other pages one opens maybe twice or three times. How many pages of... have you opened twice or more often to read or to look at something?"

The prompts:

(1) No page opened twice or more often to read or to look at something

(2) Only few pages ...

(3) A quarter of the pages ...

(4) Half of the pages ...

(5) Three quarters of the pages ...

(6) All, almost all pages ...

It was checked, whether the recognition technique (TTB, reading and noting studies: editorial and advertising) was suited for differentiation and evaluation of advertisement exposure probabilities on a large scale, such as a representative sample of 4,200 respondents, not only subscribers to or buyers of a particular issue of individual titles, but the whole reflection of the total readership. It is obvious that the requirements for the field work were extremely stringent: all interviews had to be conducted within four weeks and two sweeps. The research company was INFRATEST.

It had to be considered whether this extraordinary effort was justified, and whether the assumptions could be verified or falsified. It was very likely, in this case, to achieve pseudo correlation, mainly due to the technique applied, validation through recognition.

There are various disruptive factors, some related to the advertisements, others related to the readers, such as:

- selection of ads
- creative approach
- selected product fields
- interest in products by reader groups
- product consumption
- attitude towards advertising

or external factors, such as

– confusion about which ads have been seen

– confusion about which magazines have been seen

– campaign awareness

– competitive advertising.

According to these factors, we noticed error rates in the recognition scores. In an additional test with readers of a weekly magazine we achieved error rates of 10-17 % (advertisement seen), by adding additional ads into the test issue, which were not inserted into the original issue. It demonstrates very clearly that the 'coverage of advertisements' cannot be established by a uni- or a few-dimensional approach.

One can only determine indicators for likely correlations. If this can be established sufficiently often, ie for the separate titles and the separate issues of the title, one can conclude that the technique is practicable.

Contrary to the usual procedure we would like to start with the result, based on a report by Siegfried Geiger for the German AG.MA:

– there is a positive correlation between the responses to the amount read scale and the noting scores for the editorial section according to the reading-noting study.

– the responses to amount read and noting of particular articles are not dependent on the pagination of the magazines. The correlation could be established with rather 'thick' as well as rather 'thin' issues.

– this is also true for the different publishing intervals.

It is not surprising to find this confirmation for the editorial part. The number of disruptive factors is much smaller than for advertisements, if we think of individual campaigns and competitive advertising for the individual brands.

– however, 'amount read' is also suited for contributing to the explanation of advertisement noting scores.

These results differentiate between the different print media categories, as expected, but, it should be noted always in one direction!

All results on amount read are in principle confirmed by the responses to the multiple page exposure. Here too, a clear relation between 'recognition' and the 'amount read' through repeated reading of the pages seems to exist.

The responses to the scale 'amount read' and the recognition scores of editorial articles show a very good match, ie the scale 'amount read' could be used as a predictor (Table 3).

Of every test issue 15 double page spreads with advertisements were tested. Between 'no article in detail' and 'read almost all' the actual noting scores were between 23 and 79% and not, as in theory, between 0 and 100%.

Looking at the figures (Table 3) we notice that 73% of the respondents have claimed to have read (almost) all pages and should therefore yield a noting score near 100%. In fact, they achieve only 79%.

Table 3

Average noting of editorial part (at least noticed headlines/photos etc)

	Amount read %	Recognition %
(Almost) No page	27	23
Only few pages	39	40
About a quarter	47	51
About half	59	60
About three quarters	65	69
(Almost) All pages	73	79

Table 4

	Average noting of ads
Amount read	%
(Almost) No	23
Few	20
A quarter	25
Half	31
Three quarters	41
(Almost) All	50

Base: 13 reading-noting studies.

Even on the base of 4,200 interviews the cases for individual issues of a magazine vanish. The base for comparisons between titles is far too small. In another separate survey for the TV supplement RTV simular results were be established. The scores range from 7-80%.

Table 5

	Average noting of ads
MPX Response	%
No page	23
No page 2 x	38
Few	45
A quarter	42
Half	44
Three quarters	58
(Almost) All	60

For advertisements the same relation can be shown, though on a lower level, indicating the same correlation between 'amount read' and 'noting' (Table 4).

The level increases clearly for those respondents who claim to have seen some parts more than once, according to the multiple page exposure response (Table 5).

Of course, it seems impossible to establish in every large scale survey the amount read as well as the MPX and in addition, at least for subgroups of the respondents among the broadest audience of a particular title, a reading noting study for validation. Therefore, the responses on amount read have to be weighted.

This could be done by including the title-specific results and calculating the error-rate seen in the recognition results. That means reducing the total noting level by the average error score. But such an approach is hardly practicable and

would depend too much on the imponderables of the recognition method. The individual categories were thus weighted with the values as follows:

(Almost) No	= 0.05
Only few	= 0.10
A quarter	= 0.25
Half	= 0.50
Three quarters	= 0.75
(Almost) All	= 1.00

For multiple exposures the same values are valid, however, in addition to the total amount read. For example: approximately seen half of the pages plus a quarter more than once: 0.5 + 0.25 = 0.75.

By separate print media categories we determined the following results:

Base: LpN = 100	Amount read index	MPX index
Dailies	82	91
Weeklies	78	95
Fortnightlies	75	90
Monthlies	73	91

SUMMARY

The results seem plausible, because they reflect the different functions of the media, such as for example the superior topicality of daily newspapers. Nevertheless, more comprehensive tests are required, preferably based on independent samples, before these scales can be introduced as standard indicators for exposures to advertisements in the media planning procedure:

– a crucial point is the capacity of the respondents through the longer interview as well as the requirement on the ability to abstract in the case of infrequently read publications.

– if the reading events are not determined individually, the second event will get the same value as further events. The additive calculation will be shortened in favour of the multiple exposure. The establishment of individual reading events is not practicable for separate media categories within the broadest audience on large scale representative surveys. If the establishment would be limited to the AIR, the calculation of the probabilities based on the broadest audience would be made difficult.

– for example, if a response would be approximately half for amount read as well as MPX, the weight would be 0.5 + 0.5 = 1.0. However, the second response can only refer to half of the half. Therefore, the weight should be 0.75. It is essential to check the semantic understanding.

– an additional survey of the Axel Springer Verlag AG demonstrates: the frequency of handling a publication increases with the decreasing quantity read. Because of the fact that the responses to the reading quantity and MPX are usually consistent, it has to be checked empirically, which parts will be opened more than once, whether an additive link is justified or impartial to the media. An over-estimation of exposures through the linking is very likely.

The model presented here in brief is probably a major and decisive step towards the establishment of 'coverage data for advertisements'; however, a number of questions has yet to be answered, to show the general validity.