

PERSONALISED MEDIA LISTS

Katherine Page, Ipsos-RSL

Synopsis

This paper outlines a new way of reducing the number of titles any one respondent is asked about, while maintaining, or even increasing, the overall media list.

In essence, PML[®] is a technique to filter which titles a respondent is asked about, based on a small number of demographic and topic interest questions which are asked at the beginning of the interview. The complexity of the filtering is such that PML is only feasible using on-screen prompting.

As well as being shown all the titles they have a high propensity to read, respondents are also shown a sample of titles they have a low propensity to read. Thus the majority of readership claims are collected directly from the respondent, and there is a basis to replace the relatively small proportion of 'missing' claims.

Many different variations of the techniques used for PML are possible. This paper outlines a prototype, which will no doubt be refined as development progresses, along with the data already available.

Introduction and background

The concept of 'Personalised Media Lists'[®] was born out of a long-running debate as to how to reduce the 'Title Load per Respondent' (TLR)* for the British NRS.

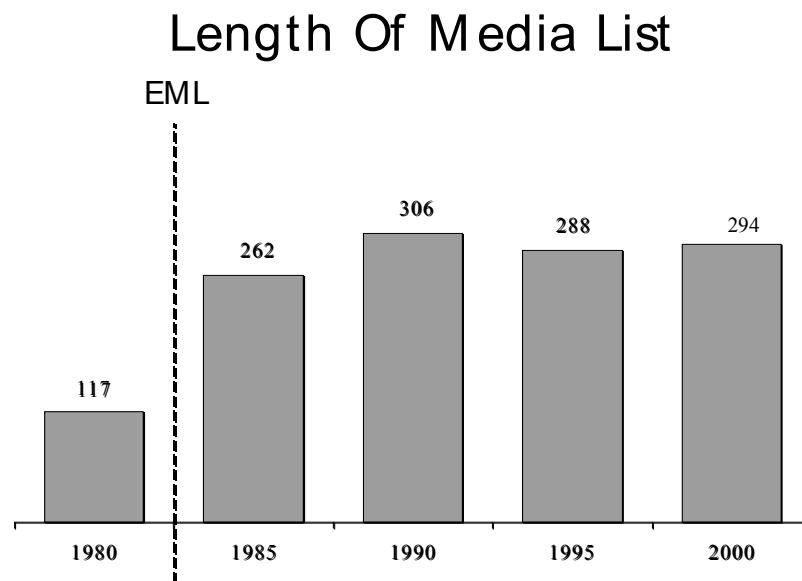
At present, each respondent is asked about around 300 newspapers and magazines, plus a selection of newspaper sections, depending on which newspapers they read. It is only possible to accommodate such a large number of titles because of the EML technique, which was introduced in 1984.**

The introduction of the EML allowed the media list to increase substantially. There were further increases, but for the last ten years the number of individual titles measured has been fairly consistent.

[®] Registered to Ipsos-RSL

* This phrase was coined by Michael Brown

** Initially titles are presented to the respondent on cards, which show six titles (EML cards). The respondent is asked to sort these cards identifying those which show at least one publication they have read or looked at in the last 12 months. In this way the majority of EML cards will be discarded (after a further check) before title by title questions begin.



Nevertheless, the length of the media list has long been a subject for debate. On one hand, the newspaper publishers are concerned that it is too long, and that there is an imbalance between newspapers (of which the respondent is shown 54, including a selection of regional newspapers) and magazines (237). On the other hand, some users would prefer to see the media list extended, to include titles and sectors not currently covered.

One solution to a media list deemed 'overlong' is to consider splitting the list between two or more sub-samples of the total survey and using fusion as the method of transferring readership data from one sub-sample to another. Work has been carried out in the UK exploring this technique. In 1995 the Press Research Council commissioned experimental work to assess how effective such a fusion process would be¹. In the event, the technique was not adopted by the British NRS, but it has been used in other countries. For instance a four-way split of titles has been used on the Dutch SummoScanner, as Wim van der Noort and Costa Tchaoussoglou explained in Florence².

Regardless of the efficiency and quality of the fusion process, however, split samples have a number of major drawbacks. The standard errors of estimates obtained on a sub-sample will of course increase substantially relative to the total sample. This has important implications for the trading currency. Likewise, the sample of readers available for individual titles to analyse will also fall. The British NRS measures many small magazines for which this would be a problem. Of course, overall samples may be increased to counteract these effects, but this will increase the cost of the survey.

Furthermore, the media list may seem rather arbitrary to the respondent. It may be, for instance, that the respondent's particular interest is in motorbike magazines, but these happen not to be included in the version of the media list he is shown, while other less relevant titles are.

At the same time that the media list length debate resurfaced in the UK in the late 1990s the possibilities of new interviewing techniques were also being considered. Of particular interest was the technique of showing all visual prompt materials on-screen, rather than on paper. This technique was pioneered in France, with the development of Double-Screen CAPI (DS-CAPI). DS-CAPI was introduced as the data collection methodology for the AEPM magazine survey in January 1999.^{3,4}

There are a number of advantages to on-screen prompting, not least of which is that the computer can be programmed to carry out much more complex routing and filtering than is possible when the interviewer has to manage paper show materials. In other words, on-screen prompting offers much greater flexibility and control as to who sees what than is feasible with traditional showcards.

It was with this in mind that Ipsos-RSL began developing ideas as to how the media list could be filtered according to demographics and topic interest. These led to the concept of Personalised Media Lists (PML), which was first presented to NRS Ltd in September 1999.

When NRS Ltd published the Specification for the new NRS Tender in July 2000, one of the requirements was indeed a reduction in TLR. A limit was set as to the number of magazines the respondent could be asked about. Tenderers were required to propose their solutions.

Summary of methodology

There are five main elements of the PML technique.

1. A small number of demographic and topic interest questions are asked before the readership interview commences. On the basis of these questions the computer software selects which magazine titles the respondent will be shown.
2. The media list of magazines has already been segmented into clusters according to a small number of demographic and interest variables. If a respondent is selected to see a particular cluster of titles, he/she will see all the titles in that cluster. As magazines which are directly competitive are likely to be in the same group, they will therefore be selected in a like-for-like manner.
3. For each cluster of magazines, past Survey data has been analysed to identify those variables which are most powerful in predicting whether a respondent is likely to have read any title within that cluster in the past year. Three strata are defined for each cluster of magazines according to likelihood to read: likely to read; unlikely to read; and very unlikely to read.
4. All respondents in the 'likely to read' stratum will be shown all magazine titles in that particular cluster. Only a sample of those in the other two strata will be shown these titles.
5. The sample collected in the lower two strata of respondents is used to estimate and 'replace' the small proportion of readership claims which have not been collected directly. This is achieved by a series of mini-fusions.

In line with client requirements, the method proposed has been developed for magazines only. All respondents will be asked about all daily and Sunday newspapers on the media list.

This paper will examine each of these elements of the methodology in turn, with data based on retrospective simulations and a test which is already underway. Many different variations of the techniques described are possible, and no doubt the methodology will develop further.

Demographic and topic interest questions

In order for the computer software to select which magazines each individual respondent will be shown on-screen, it will be apparent that it is necessary to collect some data before the readership questions can begin. A few simple demographic and topic interest questions are needed to filter who will see which EML screens.

Clearly it is not desirable for this to be a lengthy or complex battery of questions, which does have a bearing on which variables can be chosen. We consider, for instance, that it would not be advisable to collect social grade information at this stage in the interview, so it is obviously not possible to use social grade in the filtering process. After analysing past NRS data and experimenting with a range of variables to predict readership, the following are proposed:

- sex
- age
- household type (i.e. presence of adults and children)
- marital status
- future plans
- topic interest

Of course, many different combinations are possible and there is an element of judgement in the choice.

Although topic interest is usually the most powerful variable in predicting readership, a combination of demographics and topic interest allows for a more powerful discrimination. This combination also ensures that all respondents in the core demographic target group will be shown the titles, regardless of their specific topic interest. For instance motoring magazines can be shown to all men aged 15-34, if this is the main advertising target, regardless of whether or not these respondents say they are interested in reading about motoring.

Asking about demographic and topic interest before readership represents a major change for the British NRS, and, although not unprecedented, is not the usual procedure for a readership survey. Although the proposed questions were relatively short and straightforward, it was decided to check that asking them first would not have a negative impact on the readership estimates. Since April 2001 a split sample test has been running on the NRS, ahead of the introduction of DS-CAPI and further PML testing.

On half the sample the topic interest and demographic questions are asked before the readership questions, on the other half they stay in their usual positions i.e. topic interest is asked directly after the readership questions, and the demographic questions are asked as part of the classification section midway through the interview.

Data from the first three months of this test indicate that there is no bias in the readership estimates according to whether the questions are asked before or after topic interest. Details can be found in Appendix A. Similarly, there is no overall bias in the level of claims by topic interest according to where it is asked. Indeed the average number of topics claimed per respondent (eleven) is the same for both halves of the sample. Details can be found in Appendix B.

Clustering magazines

Rather they select which titles are shown to respondents on a magazine by magazine basis, the media list of magazines is segmented into a number of clusters. The PML system will then determine which of these clusters the respondent will be shown. If a respondent is selected to see a cluster they will be shown all the magazines in that cluster. This means that competitive titles are shown on a like for like basis.

Hierarchical cluster analysis has been used to identify which magazines have the closest relationships in terms of the profile of their Read Past Year (RPY) readers, using the range of variables described in the previous section.

Of course this analysis suggests a variety of solutions, but initial work concluded with a 16 cluster solution. We are expecting these clusters to change, however, when data from the new list of topic interests currently being tested on the NRS becomes available for detailed analysis.

Cluster		Number of titles
1	Women's magazines	29
2	Younger women's magazines	13
3	Parenting	5
4	Lifestyle	19
5	Older women's magazines	7
6	Older titles	6
7	Teenage	13
8	Football and fast cars	9
9	Men's, music and film	11
10	Men's	9
11	TV	10
12	Men's interests	32
13	High flyers	8
14	Hobbies	5
15	Gardening, country and classics	18
16	Trader magazines	3

As can be seen, the clusters vary in subject and size (from five to 32 titles).

Predicting likelihood to read

For each cluster of magazines, past Survey data is analysed to identify those variables which are most powerful in predicting whether a respondent is likely to have read any title within that cluster in the past year. It is these variables which will be used to filter which magazine clusters the respondent is shown.

RPY rather than average issue readership (AIR) was used for two reasons. Firstly, it enables us to look at the largest possible group of readers for that cluster. Secondly, a positive RPY claim is the filter to all subsequent readership questions, including the crucial recency question.

CHAID analysis was used to identify which variables (demographics and topic interest) are the strongest in explaining readership. These variables are of course different for each cluster, and also vary in number by cluster. The power to explain readership is more pronounced in some clusters than others, depending on how tightly targeted the readership is. Expecting or having young children, for instance, is a strong predictor when it comes to describing the readership of parenting magazines. The variables describing readership of more general women's titles are less powerful.

Having defined those variables which best predict readership, the next step is to use them to stratify the total sample. Again, this is done separately for each cluster of magazines.

We have defined three strata for each cluster:

1. HIGH - most likely to be a reader
2. LOW - unlikely to be a reader
3. VERY LOW - very unlikely to be a reader

In deciding how to stratify respondents we have had to balance two considerations. On one hand we wish to maximise the difference in readership penetration between the three strata in order to make the scheme as efficient as possible. On the other hand, we want to ensure that as many as possible of the readers are questioned directly about the titles concerned. As we intend to show the titles to all respondents in the High stratum, but only a sample of those in the Low and Very Low strata, it follows that most readers should fall into the High stratum.

An example stratification is shown below:

Men's, music & film (11 Titles)					
	Description of Strata	% of unweighted NRS sample	% of respondents who have read at least one title in the past year	% of respondents who are average issue readers of at least one title	Average Issue Readership penetration (at least one title)
HIGH	Everyone aged 15-24 Men aged 25-44	26%	72%	78%	30%
LOW	{ Women aged 25-34 interested in arts/music etc. Men aged 45-54 interested in arts/music etc.	12%	15%	13%	11%
VERY LOW	Other men				
LOW	Other women	62%	13%	9%	2%

Base NRS October 1998-September 1999

As can be seen from the above, just 26% of the NRS sample yields 72% of RPY readers for this magazine cluster and 78% of AIR readers.

The figures above, however, are based on net readership of the 11 titles in the cluster. Looking at the average issue readership penetration of individual titles across each of the strata, the skews become ever more pronounced. Examples are shown below for four magazine titles: FHM, one of the leading men's magazines, and three music magazines.

		% of respondents who are Average Issue Readers of			
	Description of Strata	FHM	Q Magazine	Mixmag	Select
HIGH	Everyone aged 15-24 Men aged 25-44	19.5%	3.4%	3.1%	1.7%
LOW	{ Women aged 25-34 Interested in arts/music etc. Men aged 45-54 interested in arts/music etc.	5.9%	1.5%	0.4%	0.3%
VERY LOW	Other men				
LOW	Other women	0.7%	0.2%	*	0.1%

Base NRS October 1998-September 1999

This stratification uses fairly broad topic interest data about general interest in 'arts' including music. A specific question about interest in rock/pop/dance music is currently being tested on the NRS and should provide more discriminating data.

Sampling respondents from the three strata

The interviewers' software is programmed both with the magazine clusters, and the respondent stratifications for each cluster. After the initial demographic and topic interest questions have been asked, the programme will make an instantaneous selection of the magazines the respondent will be asked about.

All respondents in the High stratum will be shown the titles in that cluster. Although respondents in this stratum will account for the majority of readers for the titles concerned, there will of course be some readers in the other two strata. Thus a sample of respondents in the Low and Very Low strata will be taken, in order to provide the basis to model any 'missing' readership claims. Typically the sampling fraction in the 'Low' stratum is 0.5 and 0.25 in the Very Low stratum.

Splitting those respondents unlikely to be readers into two strata, Low and Very Low, improves the efficiency of the sample fractions chosen. In particular, the sample taken from the Very Low strata can be minimised because the probability of finding readers is so low, (hence variance in the estimates obtained is also low).

It will be apparent that using this procedure the number of titles/screens shown will vary from respondent to respondent unless further controls are introduced. Some respondents will be selected to see more titles/screens than others by virtue of their demographics or interests. Even if two respondents have identical demographics and interests they are still likely to see a different number of screens because of the random selection of clusters for which they are in the Low/Very Low strata. Furthermore, the magazine clusters vary in size.

Having simulated the PML selection, the distribution of TLR per respondent (assuming six titles per EML screen) without any further control would be as follows:

Number of titles including newspapers	Number of magazines	% of NRS respondents	Cumulative % of NRS respondents
Less than 144	Less than 90	3.5	3.5
144 - 168	90 - 114	10.1	13.6
174	120	5.8	19.4
180	126	5.5	24.9
186	132	5.9	30.8
192	138	5.4	36.2
198	144	7.3	43.5
204	150	7.8	51.3
210	156	8.1	59.4
216	162	8.2	67.6
222	168	7.4	75.0
228	174	4.3	79.3
234	180	4.7	84.0
240	186	3.7	86.7
246	192	3.0	90.7
252	198	3.3	94.0
258	204	1.7	95.7
264 - 300	210 - 246	4.3	100.0

The debate as to whether the number of titles shown should be constrained to an equivalent number is one of the key issues to be resolved. The argument in favour of ensuring that respondents are shown an equal or near equal number of titles is that titles whose estimates are based predominantly on respondents with a light TLR may be at an advantage. Such an argument is based on the underlying concern about TLR, namely that the longer the media list, the greater the respondent/interviewer effects and the lower the readership estimates.

Possible interviewer manipulation of the filter questions to ensure a 'short' list is also a potential concern, although we consider that the complexity of the filtering and sampling procedures make this unlikely. Because of the random sampling of the Low and Very Low strata it will be impossible for the interviewer to predict the outcome of a certain combination of answers to the initial demographic and topic interest questions.

On the other hand, PML not only reduces TLR but increases the likelihood that the titles which are shown are the ones most relevant to the respondent. Furthermore respondents who do read a wide range of titles are likely to be the most able and efficient in responding to a long list, and may expect to be shown their full menu of titles. Conversely, a respondent who reads little may be the most daunted and/or bored by a long list.

At this stage, the extent to which the effect of personalising the relevance and length of the media list may offset the effect of presenting media lists of varying length is speculation. However, it will certainly be one of the most interesting features of PML to test.

Proportion of directly collected readership claims

The majority of readership claims will be collected directly. Using past NRS data, we have been able to simulate the PML process (using the prototype design) and estimate the number of readership claims which would be collected directly for each title. This varies of course from title to title, depending how well the title fits within the magazine cluster/PML stratification. The original estimate was that on average 89% of unweighted readership claims would be collected in the usual way. Directly relevant topic interest was not available for all titles, however, and it is expected that the average will rise to well over 90% of claims being collected directly once PML is re-designed using the revised topic interest question.

Crucially, however, these are substantially higher proportions of directly collected readership claims than would be possible if the media list was simply divide arbitrarily between different sub-samples of the Survey. If the media list is split between two half samples than the number of directly collected claims per title will be 50% of what it was on the total survey, assuming no increase in overall sample size.

Any missing claims will be modelled using a series of mini-fusions. Assuming that there are 15 magazine clusters, for instance, 30 separate mini-fusions will take place (one for each of the Low and Very Low strata for each cluster) with donors passing all their readership claims for magazines in the relevant cluster to recipients in the same stratum. In this context, it should be noted that the amount of data which will be attributable to the fusion will be small, i.e. no more than 11% of all readership claims.

Impact on Standard Errors

As the readership estimates will no longer be based on a full sample of NRS respondents, there must be an impact on the standard errors of the estimates produced. Again, it has been possible to estimate the impact on standard errors using past survey data. On the basis of the prototype PML design, the average increase in standard errors is 15%, though again there is variation by title.

Increase in SE compared to current NRS design	% of published magazine titles (excluding TV titles)
Less than 10%	23.3
10 – 19%	52.9
20 – 29%	17.5
30 – 39%	6.3

This compares well to what would happen if the media list was simply split between two half samples. That would lead to an increase in standard errors of over 40%.

As already discussed, we are anticipating improvements in the PML design when the revised topic interest data are available, and these should help minimise the increase in standard errors further.

There is another key point to make, however, which is that readership estimates are rarely used for media planning purposes on the basis of the total survey population. It is more usual to run schedule analysis, cost ranks etc. on the basis of specific target markets. Assuming that these target markets correspond broadly with the magazines' target audiences, it is highly likely that all respondents in the target market will be questioned directly anyway about the magazines concerned. For instance, using PML all women will be asked directly about their readership of general women's magazines such as Woman's Own, Hello! and Cosmopolitan. Not all men will be asked, hence the increase in standard errors, but for all analysis conducted on a female target group there will be no increase in standard error. In other words, there is no loss of precision where it matters most.

Testing

Initial field tests to check that the PML filtering procedures are practical in the field have already been carried out, although these were not of an order to generate quantitative readership data. Once a full six month sample of the new topic interest data are available (based on around 18,000 respondents), the intention is to re-work the PML magazine clusters and filter definitions with a view to a quantitative test. The British NRS will switch from CAPI to DS-CAPI in January 2002, and the intention is to conduct a test of PML on half the NRS sample from July 2002 onwards.

Conclusions

The commercial pressure to measure many publications within one survey is strong, and perhaps the greatest influence of all in shaping methodology. This pressure must be balanced against the potential implications of respondent overload for 'reliable' measurement of readership, particularly as those implications have been shown to be greater for some types of publication than others. Just how long a media list is 'too long' is a moot point, however.

What we can say, though, is that *if there requirement to reduce the title load per respondent*, (or increase the overall media list without further increasing the load per respondent) PML seems to offer a promising option, which minimises the inevitable impact on standard errors and sample sizes. Of course, many different designs and variations of the techniques described here are possible. There is no doubt that there will be further developments as we move into full-scale quantitative testing.

References

1. Carpenter, R and Wilcox, S (1995) Data Fusion in the British National Readership Survey – An Experiment. Worldwide Readership Symposium, Berlin.
2. Tchaoussoglou, C and Van der Noort, W (1999) Divide and Unite – Splitting the SummoScanner and Data Fusion. Worldwide Readership Symposium, Florence.
3. De Langhe, E and Le Van Trouc, O (1999) The CAPI Double Screen Questionnaire in the French NRS: From Clay Model to Practice. Worldwide Readership Symposium, Florence.
4. Marx, J-L (1999) The New Magazine Press Readership Survey in France – Preliminary Results. Worldwide Readership Symposium, Florence.

Appendix A

Split Sample Topic Interest Test

	NRS April-June 2001	Readership First	Topic Interest First
Unweighted sample	8,901	4,479	4,422
Est. Population 15+ (000s)	46,818	46,818*	46,818*
Gross Readership:	%	%	%
National Dailies	69.1	68.5	68.2
National Sundays	79.7	81.3	77.5
General Weeklies	54.9	53.8	56.7
Women's Weeklies	51.1	51.4	51.5
General Fortnightlies	4.2	4.3	4.3
Women's Fortnightlies	5.0	5.1	4.3
General Monthlies	122.9	116.9	128.4
Women's Monthlies	107.4	104.4	110.4
Bi-monthlies	3.0	3.3	2.9
Women's Bi-monthlies	19.6	19.5	19.6
Quarterlies	1.8	1.9	1.8
Women's Quarterlies	5.3	4.9	5.6

* N.B. The two halves of the sample have been separately re-weighted to match the Survey profile.

Appendix B

Split Sample Topic Interest Test

	NRS April-June 2001	Readership First	Topic Interest First
Unweighted sample	8,901	4,479	4,422
Est. Population 15+ (000s)	46,818	46,818*	46,818*
Topic Interest:	%	%	%
Sport	49	48	49
Cars/Motoring	36	36	36
Food and Drink	54	55	52
Home Ideas, DIY	51	52	50
Gardening	45	44	46
Fashion/Clothes	44	43	46
Film and Video	59	59	59
Classical Music	19	19	19
Rock/Pop/Dance Music	39	39	39
Arts/Books/Theatre	37	36	37
Photography	13	15	11
Personal Finance/Investment	31	31	31
Business/Company News	25	25	26
Travel and Holidays	60	59	60
Science/Technology	31	33	31
Computers/The Internet	31	31	31
The Natural World/Geography	43	43	42
The Countryside	44	44	43
Beauty and Personal Appearance	33	33	32
Health and Fitness	53	54	51
Education	37	37	37
Baby/Childcare/Parenting	17	16	17
Women's Pages Generally	42	43	41
Caravanning	5	6	5
Items for Sale (e.g. Cars, Bikes, Houses)	41	41	42
Jobs/Appointments	33	33	33
Celebrity News and Gossip	53	55	52
TV Programmes	86	87	86

* N.B. The two halves of the sample have been separately re-weighted to match the Survey profile