AN IMPROVED METHOD OF COLLECTING AND PROCESSING READERSHIP DATA

Peter Masson and Dr Paul Sumner, Bucknull and Masson

The dominant readership data collection method for over 2 decades has been 'Recent Reading' derived from a single interview survey. A Binomial convolution model (or some approximation thereto) is then used to calculate the usual schedule statistics (Net, Gross, average frequency/FD). This has required in addition the collection reading frequency.

We do not challenge the utility of this readership model. We do say that in this complex multi-media age it is woefully inadequate.

There are 4 principal inadequacies:

1. Readership is modelled as occurring all at once.
2. No account is taken of the repeat reading of a single issue. Gross reach is underestimated for all magazines.
3. The measure reports issue exposure and not advertisement exposure
4. The independence assumption which underpins the Binomial Convolution model is never tested. In some cases it is clearly wrong (parent and supplements, daily reach of dailies).

Do these issues really matter?

1. The time dimension.

If we are to progress in the evaluation of the effect of given doses of advertising it is essential that we know exactly 'when' the advertising exposure takes place. Current data collection techniques for TV, Radio, Outdoor and Internet provide us with exposure data by day part by day. Singularly (of the major media) magazines are unable to provide this data (newspapers can be modelled to provide reach by day of week (1.2.)). This severely limits the assessment of the contribution of magazines in the evaluation of a mixed media campaign across time.

We will need readership data for magazines at least at the on a day by day basis (because the rate of 'forgetting' (3) is to be measured across hours and days rather than weeks) but preferably to the day part level. Time of magazine reading during the day is an important input to planning media in time juxtaposition to maximise synergy effects. We also need it at this day part level to establish multiple pick ups (see below) of a specific issue.

This lack of time related data for magazines exposure has been a major cause of the apparent poor performance of print in advertising recall studies in comparison to TV. In many cases only a fraction of the print delivery will have taken place before the recall test*.

2. Repeat reading

The Recent Reading model takes no account of repeat reading occasions which give rise to additional Opportunities to See an advertisement. Why have publishers been prepared to give up these additional OTS when a repeat passage past a poster site or a repeat visit to a Web page clocks up another OTS?

3. Advertising exposure and not issue exposure

The answer to the above question may be something of a trade off. 'I know that my issue audience OTS is greater that the advertising OTS. By ignoring my repeat 'pick ups', which will bring the reader into contact with most pages before the magazine is discarded, the issue audience estimate is probably therefore a good estimate of the advertising exposures I am offering'.

A convenient argument but it is clear from proportion of issue data that there are large variations between titles and the proportion of issue read is well below the 100%. We really need to qualify the (average) advertising page exposure explicitly **and** the multiple pick up explicitly (by time).

Again advertising page exposure and multiple pick ups will vary a great deal within target and reading frequency groups. We will need to collect and apply the data at the individual informant level.

4. The independence assumption

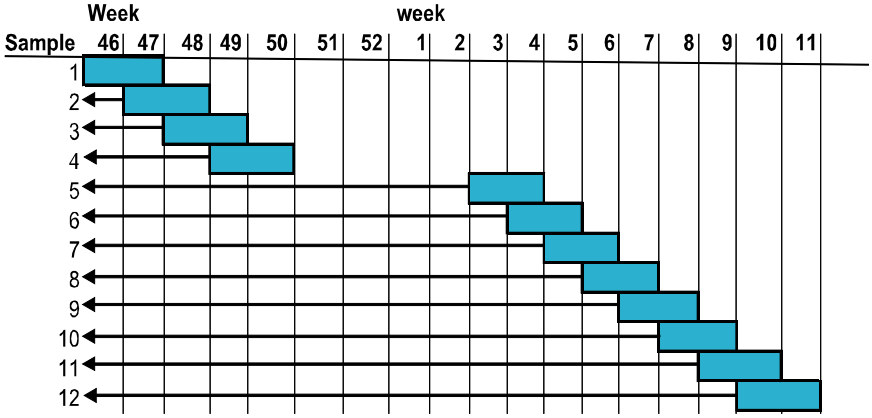
The data for all other major media (TV, Radio, Internet, Poster) is or will be in the form of panel data (of varying durations from 1 week to continuous). For the period of the panel no modelling is required to produce the Reach and Frequency estimates (apart from compensating for panel drop out). These statistics can simply be counted. All the limitations of the independence assumption can be removed. It makes sense therefore to try to collect data for magazines in the form.

However there remains the issue of evaluating campaigns of longer duration than the panel. While TV and Internet tend to be continuous, Radio and Poster tend to be short term, perhaps only of one or two weeks duration.

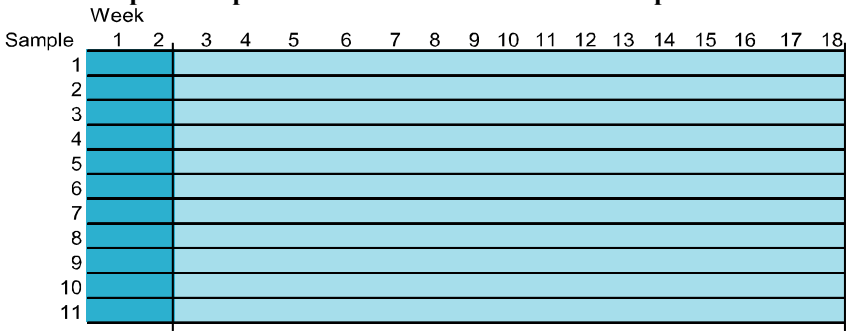
Our approach, explained in some detail later in this paper for the Belgium MAGlab data (4), is to 'expand' the short term panel data to 'n' weeks and 'count' the reach and frequency statistics for any campaign period. This extends our work for the creation of TV and radio 'Virtual' diaries from day after recall data (5).

These techniques have an application for any short term panel (whatever the data collection method) where there is a succession of 'waves' across time as indicated in the following schematic.

A. The fieldwork samples and timing



B. All sample 'collapsed' into weeks 1 and 2 and then 'expanded to 'n' weeks



The 'waves' from each sample may be analysed separately (if sufficient in size) to provide regular day by day across the period top line reporting. However they can be combined together (for any consecutive period) to provide a larger sample (for detailed target group analysis) representing an average (of specific issues or ratings) for the period. Then this combined sample can be 'expanded' so that each informant has (readership data) for every survey period.

Working with such averages of specific issues (or ratings) is more convenient in the media planning process and means that magazine publishers are not involved with pricing on the basis of specific issues with differing audience levels.

The Belgium 'MAGlab' survey design.

1. 12 random samples (sufficient to produce a net return) of 250 were drawn

2. Each sample was staged to start one week later than the previous (with a gap in the Christmas/New Year period).

3. Respondents first completed the 'standard' recent reading, reading frequency and proportion of issue read questions administered by CIM (Belgium JIC) readership survey, along with identical classification questions.

4. Respondents then completed a 14 consecutive day diary reporting their reading events. All reporting started on a Sunday.

5. Each day of the diary enabled respondents to report on up to six titles or different issues of the same title that had been read that day. Respondents first wrote in the title (the diary contained a check list of titles, in typed form without mastheads) and the date of issue or issue number. There were 69 titles on the prompt list in the Flemish questionnaire and 58 in the French questionnaire. Respondents were specifically asked to include reading of titles out of home and to make a note of the issue numbers/dates when reading.

6. They then reported whether the title had been looked at or read during each of the 7 listed day parts and the time spent reading in each of these day parts on a 7 point scale the lowest point of which was 'less than 2 minutes'.

7. Finally they reported the proportion of the issue read during the course of all the reading occasions during that day.

8. In the data file each reading event for each informant over the two week period was recorded as an age (in days) since the publication release date (namely the day the first copy of that issue came off the binding line)

From 2 week Panel Diaries to a long term Virtual Diary

The descriptions below are considerably simplified for clarity of exposition. However, these simplifications are in the area of the exact description of each sub-panel and its dating. The simplifications do not affect the construction of the VDiary, nor the model required for VDiary processing.

This section describes the processing of the informant provided S-diaries, which comprised the diary panel, to produce the Virtual Diary that is the input to the new model of readership. The diary panel, for the purposes of this explanation, is assumed to consist of 12 sub-panels.

Thus we began with 12 two-week sub-panels. If we call the week that the first sub-panel (Diary A) commences week 1 then we had the following pattern of diaries:-

Diary Letter	Week Commencing
A	1
B	2
C	3
D	4
E	5
F	6
G	7
H	8
I	9
J	10
K	11
L	12

That is, Diary n week 2 is Diary n+1 week 1.

Processing for weekly publications

The processing had to be done for each weekly publication separately. The processing scheme was the same for each weekly publication.

We created a 6-week virtual diary entry for every informant in samples (diaries) A through L. Every informant only had 2 weeks actual data on which to build this VDiary entry. In fact we concentrated on the first week of each S-diary. The information in S-diary wk2 was used only for information on turnover between issues and to provide information about the performance of the issue of wk1 in wk2. **The 6-week Vdiary had entries for one issue only of the publication in question. Multiple issues are dealt with in the readership model.**

The only people (at this stage) who had Vdiary entries created for this publication are those informants who made at least one claim for any issue of the publication in wk1 of the S-diary .

The 2 weeks actual S-diary data gave the information on the first two weeks performance of Issue 1 of the VDiary, for every informant. In this way we produced an 'average' week as the basis of the VDiary. This is an important point. The VDiary record is a diary of an 'average' period, not of specific weeks. There was in fact no necessity for the sub-panels to be placed in consecutive weeks. The 12 sub-panels could have been spread over six months, or a complete year.

Suppose Diary F (say) starts in week 47. Then Week 1 of the VDiary will only carry over all references to Week 47, but not to previous issue weeks. The second week of Diary F (actual week 48) will carry over all references to week 47, but not weeks prior to 47, to the second week of the VDiary. The second issue in the S-diary (wk2 of the S-diary) is ignored at this stage. We had however to maintain a check to see if the first week of the S-diary contains any mention of issues prior to the start week of the S-diary.

Weekly Vdiary - Issue Week 1 entries that first appear after Week 1

We needed a six position marker **A** as part of the VDiary, for each weekly publication. This marker indicated whether this publication issue (the issue appearing in wk1 of the Vdiary) was read at all in wks 1 to 6. This reading might have been repeat reading (read in previous week) or read for the first time in this week (even though the issue appeared earlier).

It can be that although an issue appears in week n it is **first** read by an informant in a later week. For the week following the issue date there is no problem. Every S-diary can be examined to see if any of issue week 1 of the S-diary is first read in week 2.

Of course it is perfectly possible to read a publication (say) in wk1, not in wk2 but read again in wk3. First reading 2 weeks or more after issue date is very small.

We knew the VDiary entry required for the second week, it was just appearances of issue wk1 in wk2 of the S-diary. The VDiary entries for the third and subsequent weeks are obtained from the S-diary entry. Suppose we are entering an S-diary which has week 47 as its first week. If there is an entry for Issue 45 (not 46) in S-diary wk1 then this issue 45 data becomes wk 3 of the VDiary. Similarly if there is an entry for wks 44 or 43 or 42 in S-diary wk1 then these entries become the VDiary wks 4, 5 and 6.

Readership Growth over Issues

Typically a weekly magazine (say publication X) will retain the majority of its readers from week to week. However a few drop out to be replaced by the same number of new readers. (It is the same number because we want to retain the same total 'average' readership from week to week). We have from the two weeks of each S-diary a start pattern – there are some readers in S-diary wk1 not in wk2 and vice versa. The **new** readers in wks3 etc. will steadily decline. We have to predict the number of new readers not in the previous VDiary weeks for every week n (>2). It is probable that for weeklies there are no new readers (of the Wk1 issue) after week 6: all the readers of the publication will appear in one of the first six weeks. We may have to change this limit from 6 weeks as more evidence accumulates. From the evidence so far, first reading more than four weeks after the issue date is tiny or non-existent.

We needed to keep a second marker showing the presence or absence of each informant in weeks 1 to 6. In fact this marker sequence may be used repeatedly in the model to predict presence in wks 7-12, 13-18, 19-24 etc. An issue wk1 informant (those for whom there was an initial VDiary created) may be in any of the 6 weeks. A second week 'new' can only be present in wks2-6 - and so on until a sixth week 'new' can only be in wk6.

The wk1&2 start pattern is the most important parameter (because obtained from the S-diary and not estimated) for this growth curve. Let us say the growth is G2 from a start (wk1) readership of R% to a total wk1&2 readership of G2*R%. (That is growth G2 = 1.mm)

We now had to estimate the end point of the growth curve (after 6 weeks at the moment and perhaps permanently). This process now described is provisional.

From the CIM preliminary interview for the actual data we were able to calculate a probability of reading X.

The only people who can be readers of X (in any week) are those who have a non-zero probability for X. We will set this as the 6-week limit to the growth of readers of X. – call it G6*R%. (That is growth G6 = 1.nn with nn>mm .)

Let F = (1.nn-1.mm) . Then G3 = 1.mm+0.55*F and G4 = 1.mm+0.76*F and G5 = 1.mm+0.89*F. This formulation is provisional and may need change, but it accords with the data we have so far..

Thus the number of new readers we need in Wk3 is 0.55*F*R%. The number we need in Wk4 is 0.21*F*R% . The number we need in Wk5 is 0.13*F*R%. The number we need in Wk6 is 0.11*F*R%. (The percents are of the total population.)

Thus at each additional week from 3 to 6 we had a number of new readers to allocate and the same number of old readers to discard. In wk3 we have readers in wks1&2 to discard from; in wk 4 readers from wks1,2&3 can be discarded, and in wk5 readers from wks1 to 4. The choice of discard (or donor) restarts, in other words a donor for wk n can also be a donor for wk(n+1).

We have described above how we decide **how many** new readers there should be at any week. We now describe **how these new readers are chosen.**

The process used was (a variant of) an ascription process. The donor set for the ascription of wk n issue was those who had seen the publication wk1 issue in week 1. We recorded which weeks the publication was seen at all.

The donee set (at wk n) was those with a non-zero probability but who had not seen any previous week's issues. We used a simple sex&age distance function:-

Donors must be used only once in this process. In wk3 we discard from wk1 and replace discards with new readers in wk3 chosen by the ascription process. In wk4 we discard from (choose donors from) all of those who were in wk1 **including those wk1 entries who were earlier discarded for wk 3**.

We should make it clear that the treatment of wk2 new readers was rather different from wks3-5 'new'. We actually know who is new in wk2 from the S-diary. We have to do an ascription to match these wk2 new readers with the closest wk1 informant. Then the chosen wk1 informant donates their VDiary to the new wk2 informant.

Processing for monthly publications

For monthly publications we have to proceed rather differently since the S-diaries do not cover a complete issue period for such publications. Future research designs should attempt to remedy this for at least a proportion of respondents.

What we were aiming for was a complete 20-week temporary diary (TDiary) for each informant/publication giving the complete reading of **one** issue over the TDiary period. (Note: 20 weeks is a provisional period. It may be necessary for the TDiary to be more than 20 weeks if there is any significant reading of issues more than 5 months old. From the evidence so far there is no case for extending the TDiary beyond 20-weeks.

We can think of **all** issues read by the informant in **one month** as the average issue of that monthly publication. More correctly the i**ssue readership** is the total readership of a single issue in its lifetime, and the **average** issue readership over a period is the mean issue readership of all issues appearing during that period. In this project we achieve a good approximation to the average issue by building our TDiary from a sample of different issues corresponding to the S-diary dates.

Each informant gives reading data for a set of issues of a number of publications. Let us take the example of the information for publication A. The informant may have read the most recent issue, or any older issue.

For example say, on the first day that informant I starts their S-Diary, the most recent issue of publication A is 9 days old. Although the informant may have read this issue on the day it was released, we only captured data in the S-diary for the period when it is 9 days old through to 22 days old (the last day of this S-Diary.) This data forms days 9 to 22 of informant I's TDiary.

Now, let's consider the issue prior to the most recent issue. By the time informant I starts their S-Diary, let's say that the previous issue is 40 days old. This data then forms days 40 to 53 of informant I's TDiary of the most recent issue. This is a crucial point in understanding the TDiary construction: the reading of an **older** issue D days after **its** issue date is taken as a proxy for the reading of the **current** issue D days after **its** issue date.

Extending this to older issues, we could thus extrapolate the TDiary as far as we liked. However, this TDiary will be incomplete because, for example, we have no data for the TDiary on days 0 to 8. We fill in the gaps in informant I's TDiary by ascribing the data from a matching respondent in a *different* sample group. (Since each sample group starts on a different date, the gaps in the data in the corresponding TDiaries will fall at different points in the lifetime of our TDiary issue.) In fact, to fill in all the gaps for any informant, we had to ascribe two matching respondents, since a month is longer than 4 weeks, and so each gap will be greater than 2 weeks, which is the length of an S-Diary.

The ascription used the same simple parameters as previously used for the weeklies. We had to use donors the minimum number of times. That is we only started repeating donors when we had used all once. The ascription was performed separately for each sample group. Two separate ascription passes were performed for each sample group. In each ascription pass, the recipient set was all readers of publication A in the sample group. The donor set was all readers of publication A in some other sample group.

We chose the donor sample groups as follows:

- For the first ascription, we selected the sample group that a) filled in the most gaps in the recipient's TDiary; and b) resulted in a TDiary whose remaining gaps were able to be filled in from a respondent in another sample group;
- Chose any sample group to fill in the remaining gaps.

VDiary for all monthly readers

So far we had a twenty week TDiary, but only for those people who read the issue of a publication published in the first TDiary month. We now had to extend this to readers of the publication in six months. The situation is similar to that of the weeklies in that some people cease to read to be replaced by the same number of new readers. The number of new readers declines from month to month.

Unlike for weeklies, we did not have an observed second point on the growth curve of new readers. We only have the first month readership and an end point, readers in one year (in this case). Suppose the readership of one issue is R%. Then the readership in 12 months is G*R where G is a growth factor. (Note G will usually be between 1 and 3 but it is possible (though very, very unlikely) for it to be as much as 12.

Now, with this growth factor and with the aid of some fairly complex, but straightforward, maths we can predict the number of new readers at any point between one month and one year.

We have described above how we decided **how many** new readers there should be at any month. We now describe **how these new readers are chosen.** It is very similar to the process used for weeklies.

Again the process was an ascription process and uses the same parameters as for weeklies. The donor set for the ascription of month n issue was those who have seen the publication month 1 issue in month 1. We recorded which months the publication was seen at all.

The donee set (at month n) was those with a non-zero probability but who have not seen any previous month's issues. We used the same simple sex&age distance function as for weeklies.

Donors had to be used only once in this process. In month n we discarded from month1 and replaced discards with new readers in month n chosen by the ascription process. In month(n+1) we discarded from (choose donors from) all of those who were in month1 **including those month 1 entries who were earlier discarded for month n**.

The Second Stage Model – Page Exposure

The first stage of the new Press model is a sophisticated count of the VDiary. Note that for this Press model the VDiary is extended so that the whole period of the schedules to be evaluated can be **counted**.

The aim of the second stage model was to refine the model (from a count of OTS) to page exposure measures. This was done using information from the 'time spent reading (per daypart)' and 'proportion of issue read' questions. The detailed results of the first stage – day parts, days, weeks and cumes – had to be available to this second stage.

Page Exposure

For each day respondents will have recorded in the VDiary (for up to 6 titles/issues of a title) the following:-

1. The titles and the issues read that day.
2. The times of day they read each issue (7 day parts). This is also a partial measure of multiple pick ups. But of course an informant can pick up a magazine several times during any one day part and this is not recorded.
3. The hours/minutes they spend reading each issue in each of the day parts. This is Page Exposure measure (PEX) 1
4. The proportion of the issue that they read including all their reading occasions of that issue that day. This is PEX 2.

PEX1 can only be applied as a reduction to PEX2 in narrowing to a day part from a complete day. We could simply have reduced the proportion seen in a day part in the ratio that the time in that day part bears to the total time in the day. But this is much too severe a reduction. We should set a length of time of day part reading (we used 20 minutes) above which the informant is assumed to have seen the same proportion of the magazine as in the whole day. Then shorter (than 20 minutes) day part reading reduces the proportion in ratio to 20 minutes, not in ratio to the time spent in the whole day.

We thus had for each day and each day-part a proportion P of the issue read on that day. It is important to note that this proportion converted an 'opportunity to see' measure (reads some of the magazine) to a page exposure measure (definitely reads X% of pages). We now had to consider how to estimate the proportion read for a number of days.

Suppose the proportions on a sequence of days are P_1, P_2, …. , P_D . There are two extreme positions. First we could assume that proportion read in total, P_T, is given by:-

$$P_T = \max (P_1, \dots , P_D)$$

In this, very unlikely, scenario (extreme 1) the informant always reads whatever they had read previously, only increasing the total proportion read if the day proportion is bigger than any previous day's proportion. We only mention this scenario to set an extreme.

At the other extreme (extreme 2) we could set P_T as:-

$$P_T = \min\{1.0, \Sigma(P_1, \dots ,P_D\}$$

That is the informant always reads new, unread parts of the magazine until the whole magazine is read. This is a likely scenario for, say, a programme magazine.

A middle position would be when we assume that the proportion read on any one day is distributed between read-before and new in the ratio of 50:50. Thus if 70% had been read before and this day's reading was 40% then this 40% would be split 40*50 = 20% new and 40*50 = 20% already read. The total proportion read becomes (70+20)% = 90%. Of course the total proportion read would be capped at 100%.

This 'equal' (i.e. day's proportion split as 50:50) split was the initial default level for all publications. However we introduced another parameter K for each publication with range from 0 to +100%. Then new: old was split in the ratio K:(100-K). A high value of K (approaching 100) indicated a magazine that is approaching extreme 2. Low values of K (less than say 40) would be very unlikely, in fact we couldn't think of any examples.

The value of K was set by examining the (total) proportion of issue read within frequency of reading groups claimed by the panel respondents in the initial interview. An average value for K was about 70. Further amendments to K can be expected following the examination of the TTB data

Calibrating the panel recent reading claims to the CIM 'currency' level.

The fit between the RR claims of panellist and the published (2004 CIM) was overall very good. The Gross Reach claims from the (2749) panellists for all titles measured was 23.3m compared to the Gross Reach from the (10,383) respondents of CIM 2004 for the same titles at 25.8m (11% higher). There was no reason for them to be exactly the same as the fieldwork periods were different as were the methodologies of recent reading claims and panel counts of exposure to any issue in a one week period. However it was decided that the AIR results from the panel should be calibrated to the CIM AIR in such a way that the day by day panel results also reflected the adjustment in AIR. The procedure was as follows:

If the panel score was higher than the CIM, then we changed the panel AIR to NO for sufficient YES panel AIR answerers such that the next YES answerer chosen to move to NO would take the panel score below the CIM score. Any panel AIR claim changed from YES to NO also had any VDiary entries blanked for this informant/publication. The YES answers could be chosen at random from all YES answerers (in the relevant language region).

If the panel AIR was lower than the CIM score then the processing required was a little more complex. It required a simple ascription process. The donors for the ascription were the current informants who had answered YES to the panel AIR question (for this publication).

We went through the donor list informant by informant with each donor selecting (by the ascription process) a donee. As many donees were chosen as were required so that the adjusted panel AIR just exceeded the CIM AIR. If the list of donors did not generate enough donees then we returned to the beginning of the donor list and continued selecting donees.

Donees were chosen from those who had a frequency claim of 1,2 or 3 (and obviously had answered NO to the AIR question) for this publication. Only if this list of frequency 1, 2 or 3 claimants was insufficient (only a tiny handful of cases) did we seek from those who had frequency claims of 4 or 5.

Frequency claimants 1-5 always yielded enough donees.

NOTE: The process not only changed the panel AIR claim from NO to YES, it also copied any diary entries from the donor to the donee.

We used three simple ascription parameters Sex, AgeGp and Class for the Ascription distance function.

Benchmarking the results

As all panel respondents also answered the 'standard' recent reading question (now adjusted to CIM 2004) it is possible to make a direct comparison of their 'recall' of reading responses against their day by day (virtual) diary responses for reading any issue of the title in an average issue period. To get a good representation of 'any' issue read we have to examine a period in the diary where multiple issues are available for reading. For the following analysis we took 35 days for weeklies and 140 days for monthlies.

Here are some examples for groups of weeklies of the relative positions of the claims made by panel members:

Table 1	Panel members		
Weeklies '000's	RR Claim	Panel Count	Average Single issue
Type	Last week	of AIR at 35 days	Net Reach at 35 days
TV /news	100	112	98
Women's	100	117	99
Total	100	113	98
Index		100	87

Panel counts of AIR were some 13 % higher for these titles than the initial Recent Reading for reading in the last week. While the precise AIR count on the Vdiary panel varies according to the time period measured it is consistently higher than the RR claim.

The single issue (average of issues measured in the fieldwork period) was just 2% lower than the RR claim although this differed between individual titles from 26% lower to 8% higher. The French language (Fr) titles were on average for single issue net reach 7% lower than the RR claims while the Dutch language (Fl) titles reported 2% higher.

Table 2	Panel members		
Monthlies	RR Claim	Panel Count	Average Single Issue
Type	Last month	of AIR at 140 days	Net Reach at 140 days
Special interest	100	76	69
Women's	100	99	88
General interest	100	110	99
Index	100	85	76
Index		100	90

In the case of the monthlies the panel count of average issue reach is lower by 15% on average than the recency claims of the panel members. This would fit with the evidence that RR tends to overestimate monthlies. However there are notable differences between the publication type and some extreme differences between specific titles with some titles achieving only 50% of their RR claims and others 20% higher. The limited sample sizes and that only two weeks of data was available for monthlies (not a complete average issue period) will be contributing to the title by title variations.

The relationship of the single issue cume audience (after 140 days) to the panel 'count' of AIR at 90% is very similar to that for the weeklies (87%)

New metrics for Magazines

Since panel respondents report daily on the reading events of each issue by 7 days parts we have a measure of daily audience (net and gross) and the accumulation of these contacts (net and gross) over time. In the second column of the table below the average number of 'pick ups' for the net audience is reported. For the TV/news publications this averages around 5 but the women's titles have only half this level.

Table 3 Weeklies	OTS (pick ups) for Single issue Net Reach	APX % of single issue read	OTS for Av page exposed	Index Gross Av Pages Exposed/ RR AIR
TV /news	5.2	78	1.8	148
Women's	2.6	78	1.5	155
Total	4.2	78	1.7	150

In the third column (although provisional at present) is the predicted proportion of issue read from the APX model described above along with, in column 4, the number of exposures to an average page for each person in the net audience. This brings the TV news magazines back more in line with the women's weeklies in terms of average page OTS (compared with issue OTS). In the final column the gross average page exposure audience is shown as an index against the initial Recent Reading claim of reading in the last week. On average this is 50% higher.

Table 4 Monthlies	OTS (pick ups) for Single issue Net Reach	APX % of single issue read	OTS for Av page exposed	Index Gross Av Pages Exposed/ RR AIR
Special interest	2.6	77	1.6	86
Women's	2.6	69	1.7	104
General interest	4.0	65	1.9	121
Total	2.8	73	1.7	94

Repeating the analysis for monthlies shows a slightly higher level of repeat reading (2.8 OTS) to the women' weeklies (2.6 OTS) but a lower proportion of issue read (73% compared to 78% for the weeklies) but overall with the same OTS for an average page exposed. Relating the resulting the average page exposure audience to the Recent Reading claim shows a level 6% lower overall.

The Panel (Vdiary) data shows therefore a significant difference in the relationship between weeklies and monthlies compared to Recent Reading

Audience structure over time

The Vdiary panel allows us to observe the composition of the audience of a title over time. As a title adds net readers over time so the frequency of reading and demographic profile changes. The following example is for a TV/news weekly title.

Cume net	WK1	Wk1-2	Wk1-3	Wk 1-5	Wk2 (new)	Wk3 (new)	WK4/5 (new)
Always	696.7	806.5	862.8	883.5	109.8	56.3	20.7
Regularly/sometimes	134.2	181	210.6	226.6	46.8	29.6	16.0
Occasionally	142.9	216.2	251.8	270.6	73.3	35.6	18.8
All	973.8	1203.7	1325.2	1380.7	229.9	121.5	55.5
Profile% cume net							
Always	71.5	67.0	65.1	64.0	47.8	46.3	37.3
Regularly/sometimes	13.8	15.0	15.9	16.4	20.4	24.4	28.8
Occasionally	14.7	18.0	19.0	19.6	31.9	29.3	33.9

In week 1 72% of the accumulated audience are 'always' readers. In week 2 under 50% of the additional accumulated audience are 'always' readers and by week 4/5 this has fallen to 37%.

We are now also able to report the average daily (or actual day by day) net (at least one 'pick up') and gross (including multiple pick ups) audience for each title.

	Wk1	Wk2	Wk3	Wk4	4 wk
TV News title	**Av Day**	**Av Day**	**Av Day**	**Av Day**	**Av day**
Net daily reach 000's	289.7	297.2	54.6	18.5	660.1
Gross daily reach 000's	374.7	385.3	62.8	20.5	843.2
Net frequency profile %					
Always	80.4	82.1	66.1	59.8	79.4
Regularly/sometimes	9.3	8.5	17.0	26.3	10.1
Occasionally	10.3	9.4	16.9	13.9	10.5
Gross frequency profile%					
Always	82.5	82.1	67.6	58.6	80.6
Regularly/sometimes	8.8	9.4	15.6	28.8	10.1
Occasionally	8.7	8.5	16.8	12.6	9.3
Av daily pick ups					
Always	1.3	1.3	1.2	1.1	
Regularly/sometimes	1.2	1.4	1.1	1.2	
Occasionally	1.1	1.2	1.1	1.0	

In the first 2 weeks this title is delivering just under net 300,000 OTS (to the issue) and a gross OTS of 380,000 on a daily basis. These contacts comprise over 80% 'always' readers. In weeks 3 and 4 average daily contacts are very much lower and the composition of readers has changed with a much higher proportion of less frequent readers (35-40%).

Since target group profiles often differ between the frequency of reading groups there will be a changing target group structure by age of issue. The following example for the same TV/news title shows its changing structure by social grade.

Average daily audience profile by social grade				
% Net	**wk 1**	**wk 2**	**wk 3**	**wk 4/5**
SC 1/2	16.7	16.0	14.0	12.6
% Gross				
SC 1/2	16.4	14.6	13.9	11.7

The Social Class (1 and 2) profile declines consistently with the age of issue.
The lower profile of Social class 1 and 2 in the Gross audience implies a lower average number of 'pick ups' during each day for this group.

Clearly this continuing level of daily (net and gross) reach demonstrated by the Vdiary panel data and the changing audience profile is something to be exploited by both planner and seller

Audience accumulation

With the VDiary panel estimating (net and gross) reach accumulation of 1 or more issues and the frequency distribution of those exposure is simply a matter of counting. Reporting can take place by or within any specified time period.

This can be compared with the current methodology using a Binomial convolution model with Recent Reading and Reading Frequency data to estimate audience accumulation

	Index of net reach cume				Average OTS			
Insertions	**1**	**3**	**6**	**12**	**1**	**3**	**6**	**12**
Binomial model								
V diary panel 'count'								
Weekly								
TV/news	*100*	*127*	*140*	*150*	*1*	*2.4*	*4.3*	*8.0*
	100	137	160	160	7.3	12.5	21.3	42.7
Women's	*100*	*144*	*169*	*189*	*1*	*2.1*	*3.6*	*6.3*
	100	169	213	213	2.7	4.8	7.7	15.3
TV/news	*100*	*147*	*173*	*191*	*1*	*2.0*	*3.5*	*6.3*
	100	160	200	200	4.3	8.1	12.9	25.9
Women's	*100*	*189*	*249*	*300*	*1*	*1.6*	*2.4*	*4.0*
	100	227	321	321	2.2	2.9	4.2	8.3
Monthly								
Member	*100*	*130*	*140*	*149*	*1*	*2.3*	*4.3*	*8.0*
	100	149	169	169	2.8	5.6	10.0	20.0
Health	*100*	*187*	*233*	*264*	*1*	*1.6*	*2.6*	*4.5*
	100	242	282	282	2.7	3.4	5.8	11.6

In the table above a range of titles were selected with different rates of cume. Results were indexed against the one issue cume successively for 3, 6, and 12 insertions (in consecutive weeks/months in the case of the Vdiary panel). The first line, in italics, reports the results from the binomial model.

While the binomial reflects the differences in growth between the titles it always reports a much lower rate of cume even at 12 insertions. The Vdiary found no data to support continued growth of any significance after 6 insertions and thus shows no further growth for 12 insertions.

Further since the Binomial model has only Recent Reading/frequency claims it does not reflect the multiple 'pick up' and severely under-reports OTS at each level

Conclusions

The way in which print delivers its exposure opportunities is complex (and virtually identical to the Internet). We do not believe that the simplification provided by Recent Reading/Binomial Convolution model, with which we have lived for the past 4 decades, can continue to satisfactorily serve the interests of Magazine publishers.

This MAGlab project based on a good sized panel sample confirms, at least for weeklies, that the recent reading claim is a reasonable proxy for the reach of an average issue across its life. But magazines are severely under-representing their audiences in terms of the gross contacts delivered - and deprived of demonstrating the advantage that their contacts are staged over time provides (continuity of exposure). At the same time we can recognise that the composition of the audience delivered changes across time.

The Vdiary panel allows us to count and value (according to time spent and proportion of issue read that day) each successive contact with the publication such that we can provide a good estimate of the day by day contacts (OTS) delivered by an average page within the issue.

Magazines can now demonstrate to its advertisers when and how many advertising contacts it delivers, which has long been the case for TV, and move themselves one step up the ARF media model. This in itself will make the interpretation of advertising recall/image studies more realistic knowing precisely the print audience delivery for each time point of the research.

As far as schedule and multi-issue evaluation is concerned the VDiary panel is a straightforward 'counting' operation (complicated only by the APX calculation) against which Binomial Convolution model can be seen to provide rather poor estimates of audience accumulation.

The costs of such panel data collection are significantly higher than for a conventional single interview Recent Reading surveys. We can foresee routes to amortizing these panel data costs over longer time periods using modelling and calibration. In the first scenario we can re-calibrate the Vdiary panel data to new (updated) Recent Reading results renewing the panel data only every other or every third year. In the second scenario we can slightly extend the data collected on the 'currency' RR survey so that we are able to attach the V'diary panel data record (using calibration) to the respondents of the 'currency survey'(6). This would be updated with each new 'currency' survey with the panel data itself being updated periodically (every 2 or 3 years).

Text references

1. ***Callius P, Lindberg I*** It all started with a Phone call 'daily reach of dailies', Worldwide Readership Research Symposium 2005
2. ***Lindberg I, Masson P, Sumner P.*** Measuring the 'daily reach of dailies' and newspaper supplements, ESOMAR/ARF Week of Advertising Measurement, Cannes 2002
3. ***Callius P, Masson P***, Time related measures of relative 'effect' - an essential ingredient for Multi-Media campaign evaluation, Worldwide Readership Research Symposium Venice 2002
4. ***Patrick Hermie, Trui Lanckriet, Véronique Couvreur & Koen Lansloot,*** The WAR Project. A small step up the ARF media research ladder, a giant step forward in magazine planning. Worldwide Readership Research Symposium 2005
5. ***Elms S, Masson P***, Multi-media audience measurement Admap 40th Anniversary Issue 2004
6. ***Callius P, Lithner, A, Svanfeldt S,*** The creation of a single source cross media Internet panel. ESOMAR/ARF Week of Advertising Measurement, Montreal 2005