

Magazines deserve time – the sequel

Patrick Hermie, Trui Lanckriet and Stef Peeters, PhD., Mediaxis, Belgium

If you should start a national, large-scale study on magazine audience accumulation, which method would you use? In this paper we will answer this question. Or at least try to. We report on the methodology and results of a Belgian experiment that compared the two methods of accumulation research and the ways to integrate accumulation data in mediaplanning.

In the age of accountability and ROI, it is more than ever essential to understand what the impact is of advertisements in magazines. Millward Brown (Pincott, 1991) established the importance of the time factor in explaining and predicting the impact of magazine advertisements. The integration of this topic in today's ARF Print Research program, confirms the need for sound accumulation data that can be integrated in the magazine planning practice. The current paper reports on a methodological experiment that builds on the pilot studies on audience accumulation presented at the 1999 World Wide Readership Symposium in Florence by Peeters, Debeer and Lanckriet (Belgium) and by Julian Baim, Martin R. Frankel and Joseph Agresti of MRI (USA).

1. A brief reminder

In order to observe the accumulation of readers in time, "first time reading" of an issue should be measured. To that end, Peeters, Debeer and Lanckriet (1999) and Baim, Frankel and Agresti (1999) used completely different methodologies.

The American pilot study used a panel of respondents who kept a diary on magazine reading for one week. The panel duration was limited in an attempt to lessen the "respondent burden" and to increase the respondent's commitment to record the diary thoroughly. The diary was fairly simple. For every title read, respondents had to fill in the issue date on a daily basis. They also indicated if they had read the issue for the first time and whether or not the magazine was read at home. This "real-time recording of readership" allows daily observation of first time reading instances. As respondents were free to fill in the issue number themselves, no maximum time limit was imposed. Even very old issues can be read for the first time. The respondents also have the possibility of filling in several issue numbers for the same title. In this way, parallel reading was taken into account.

Although a panel allows to collect single source data over time, it is extremely difficult to ensure the reliability and validity of its outcome as it favours in-home reading, assumes active and correct memory recall and heavily depends on respondents' attitude and thoroughness. In our 1999 experiment, we therefore used short face-to-face interviews. Our objective was to create a flexible methodology that could be repeated at any given time for any given title, entailing a simple procedure for interviewees and minimising memory requirements. In the face-to-face interview we combined recent reading and through-the-book methods. Recent reading questions, identical to the Belgian Readership Survey (CIM), were used to measure total readership and declared magazine reading in the reference period (Average Issue Readership). A complete through-the-book approach was used to measure the audience at the issue level: all readers of a title (not only readers during the last week/month) received a complete copy of the last six issues while answering readership questions.

While the results of both studies were remarkably similar on all important variables, some methodological issues clearly remain. For one thing, both pilot studies had to deal with participation and representativity issues. Given the volume of the required material in our study (six issues for 33 titles), interviews were carried out at central location after street recruitment. This explains the observed over-representation of regular readers. As this variable influences accumulation directly (Napier and Douglas, 1998), this parameter was weighted before modelling.

Using a random digit telephone sample in their initial test, Baim, Frankel and Agresti (1999) reported that participation acceptance was lower than 50%. Only 30% of respondents who had agreed to take part actually returned the diary. There was an over-representation of elderly females. In the second test, a panel was made up based on a national NPD panel. A letter was sent beforehand followed by the diary to 50 panel respondents. MRI got 33 useable diaries back. Once again there was an over-representation of elderly females. In the full-scale study, the sample was stratified to ensure oversampling of men and younger, employed adults of both sexes. Prior to the diary mailing, respondents were sent a letter describing the study and asking for their cooperation. The letter contained a brief description of the respondent's task, instructions on how to fill in the diary and an incentive. MRI also tried to contact all respondents the weekend before they were due to start their diary recording. This phone call served both as a reminder to start the diary on Monday and to answer any questions on diary instructions. The completion rate of diaries was 51%. Slightly more than half (52%) of the respondents were men showing that the oversampling successfully counterbalanced the anticipated differential response rate.

Identifying first time reading instances is not evident. Baim, Frankel and Agresti (1999) calculated accumulation curves on the basis of the diary entries of first time issue reading. In our study, however, the use of reported first time reading proved to be impossible since the 'first time read' question led to misinterpretations. New issues were sometimes not identified as read for the first time: this kind of time framing is too difficult or our question was badly worded. For this reason, we based our curves on the estimated accumulation of total readers.

2. The sequel

2.1 Objectives

The most important objective of our previous experiment was to create a flexible methodology which could be used at any moment for any title with an easy procedure for respondents that does not require them having to search their memory too much. This is a necessary condition to carry out a large-scale audience accumulation research project.

Although this objective was largely met, there are a number of points that could be improved. The methodology appears to be fairly simple for respondents and was not a great burden on their memory. The logistic handling of a full through-the-book approach, on the other hand, was rather heavy and, contrary to our expectations direct measurement of first-time reading was not reliable resulting in a more complex accumulation modelling. The straightforward diary approach of Baim, Frankel and Agresti (1999) seemingly offers an attractive alternative. The first objective of this experiment was to gain more insight into both the through-the-book and panel method. Which measure is the most valid and reliable?

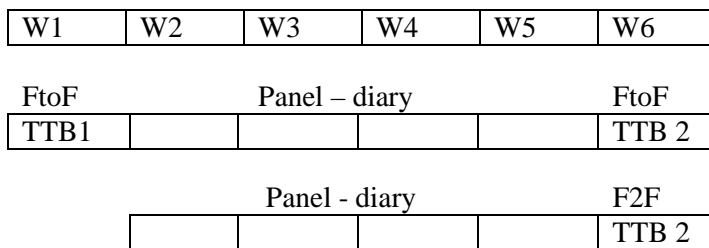
Our second objective was to find out which one can most easily be repeated on a large scale? It is possible to simplify the through-the-book method. On the other hand, a panel might prove to be a more practical alternative after all. Baim, Frankel and Agresti (1999) demonstrated the viability of a one week panel. Yet, a longer lasting panel might improve the efficiency and economics of accumulation studies. Remember that the German Axel Springer Verlag AG, used a panel of 6,377 people who kept a diary for a six weeks period (1992). A longer lasting panel would also allow direct measurement of replicated reading.

Comparing both methods and combining both objectives led to a somewhat more complex experimental design.

2.2 Method

Figure 1 shows a schematic presentation of the design. Half of the respondents were interviewed face-to-face (in home) in week 1. All respondents were then asked to keep a diary for four weeks (weeks 2-5). Afterwards, all respondent were interviewed face-to-face in week 6.

Figure 1: Experimental design



W = week

TTB = Through-the-book (TTB1= prior to the panel, TTB2= following the panel)

F2F = face-to-face

The field work took eight weeks to complete. In order to generate observations over a sufficient period of time, the panel was divided into three waves each starting one week later. In this way, the diaries covered six weeks instead of four.

The respondents were recruited by telephone and selected on the basis of their reading patterns (“ever read” and “read last week/month”, worded exactly as in the Belgian Readership Survey, CIM). Participants had to read at least one of 6 selected Dutch titles (4 weeklies and 2 monthlies).

The face-to-face interviews contained one question at title level "Have you read this magazine during the last 7/30 days". The other questions were asked at issue level. This means that the six most recent issues of a title were presented to the respondents. They were asked if they had read or leafed through a specific issue and if so, whether they had done this during the last 7 days (30 days for a monthly). The subsequent question “And have you read this issue before?” replaced the “unreliable” question "Did you read it for the first time during the last seven days/ last 30 days?" from our previous study.

The diary concept is a very simple one, directly based on the study by Baim, Frankel and Agresti (1999). A pre-test¹ showed that keeping the diary was not only easy but also an enjoyable experience. The respondents filled in the diary for a period of four weeks and took it with them wherever they went. This was made possible due to the diary's user-friendly format. Based on the experiences of Baim, Franklin and Agresti (1999) and our own pre-test, the reliability of the diary format was ensured.

Figure 2 : The diary

The logistic burden of the face-to-face methodology could be lessened by working with stripped issues instead of complete issues. In his book, Michael Brown (1999) quotes a number of sources that conclude that readership figures for stripped issues do not differ significantly from non-stripped issues, if the stripped issues contain a certain percentage of editorial pages. At the same time, Brown points out that some studies conclude otherwise. For the sake of comparability with our previous pilot study, we continued with the complete through-the-book method in this experiment.

3 Results

3.1 The sample

The experiment was limited to the Dutch-speaking part of Belgium. In this way, the number of titles to be examined could be limited more easily and men could be included in the sample.

2300 people were identified by phone as readers of at least one the six titles. Out of these, 668 (29%) were willing to start the procedure. Compared to the participation acceptance of approximately 50% obtained by Baim, Franklin and Agresti (1999), this seems on the low side. However, demands on respondents in this study were more important. Respondents were asked to keep a diary for 4 weeks and were explained that an interviewer would visit them twice (to explain and to collect the diary and in most case ask supplementary questions). Eventually, 75% out of these 668 respondents started the diary after the explication of the interviewer. 72 people were eliminated, as interviewers could not contact them after all.

Out of 495 respondents who started, 415 diaries were returned. The non-response noted in our experiment was due to the loss of the diary, boredom, having to stop when keeping the diary was forgotten several times and was then given up altogether, a lack of willingness to cooperate from the start and personal reasons. An 84% return rate is a very satisfactory result, proving that the diary itself was fairly easy to keep. This exceptionally high response is undoubtedly to a great measure due to the fact that the diaries were delivered to the respondents at home and later collected by interviewer. In their experiment, Baim, Frankel and Agresti (1999) noted a completion rate of 52%. In their case the diaries and instructions were sent by post and respondents were phoned once more just before the experiment got underway.

¹ Ten people kept the diary for a one-week period and were then interviewed about ease of use, the feasibility/complexity and the willingness to carry this out four weeks in a row. Furthermore, questions were asked about their response behaviour (directly after reading or leafing through a magazine, a few hours afterwards, etc.).

Eventually, some diaries could not be used for comparison of the two methods because respondents had not completed the first or the second questionnaire or because the diary itself was only partially completed. The additional drop out results in a usable response of 69%, lower than the 84% return rate but still considerably high.

The profile of the respondents who had kept the diary correctly and those removed from the sample varied only significantly on professional activity. More retired people dropped out. Looking at each week separately, we observe an equal distribution of the number of editions and titles read during these 4 weeks. This suggests that respondents have been thoroughly performing their task during the entire period.

The sample was evenly distributed throughout Flanders. Women, middle aged (35-54 years) and better-educated people were somewhat over-represented. Contrary to our previous study, regular readers were over-represented in only one title. In terms of source of copy, however, primary readers were generally over-represented. While this means that accumulation data may be slightly overestimated, we decided not to weight the sample because the main objective of this experiment remains the comparison of two different methods within one and the same sample.

3.2 Accumulation

3.2.1 Comparing the methods

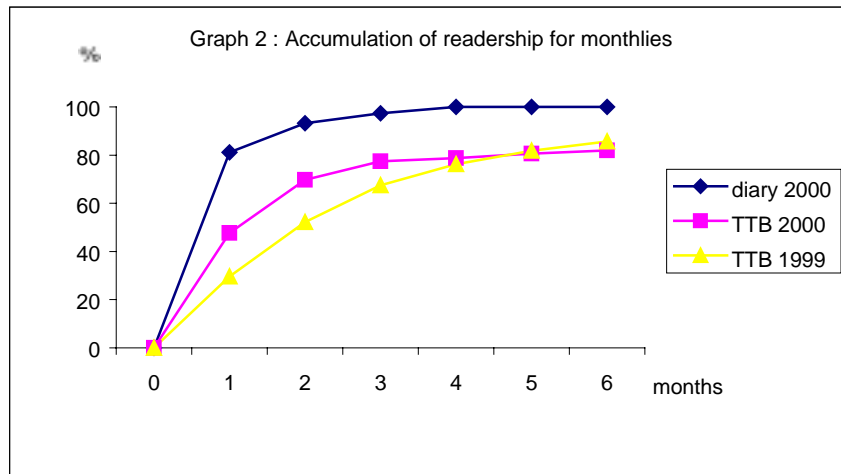
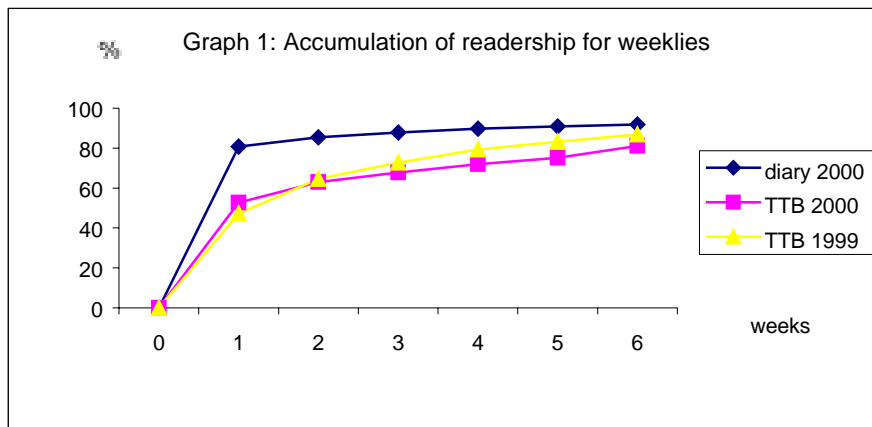
In our previous study (Peeters, Debeer, Lanckriet, 1999), we calculated accumulation curves on the basis of total issue reading question. Answers to the question "Have you read this issue for the first time during the last seven days?" turned out to be unreliable. While we speculated that the problem arises from the limits of episodic memory, inadequate wording could not be ruled out as an explanation. That is why we in this experiment we used an alternative question "Have you read this issue before?". Analysis on this new question again shows that it is very difficult for respondents to reconstruct the timeframe of reading moments. 26% of weekly issues that exist less than 7 days are declared to be read before. 14% of new monthly issues are wrongly declared to be read before the last 30 days. We must conclude that there is no reliable way of measuring first time reading based on recall.

The calculation of the accumulation curves of the through-the-book method was therefore based on the total issue reading question, as in our previous study. The question "Did you ever read or leaf through this issue?" allows to estimate the cumulative audience for a specific issue after 1, 2... to 6 reference periods. The number of readers of an issue divided by the total number of readers on title level, is an estimate of the proportion of total readership reached after one week. When we add those who read the previous issue, but not issue X, we have an estimate of cumulative readership after two weeks, and so on. This paper reports only observed accumulation data (before modelling).

The curves of the diary method are based on the first time reading of any issue read during the lifetime of the panel. As the date of publication had been noted, it was very easy to ascertain the issue date. Furthermore, the respondents indicated when they had read the issue for the first time. The number of first time readers per issue is divided by the observed number of first time readers of all the issues. An advantage of the diary method is that the data can be collected per day resulting in more detailed accumulation data.

It is interesting to note that the diary method also has its flaws. On a control on the entries, we found that in 12% of the cases respondents did not indicate a first time reading moment for the most recent issues they reportedly read. This means that respondents either simply forgot to indicate a reading moment as a ‘first time reading’ or forgot to report the real first reading moment that occurred earlier. Fortunately, these ‘mistakes’ can be corrected by assuming that the first reported reading moment for the most recent issues is effectively the first time reading. Clearly, also the diary method has to deal with inaccuracies.

3.2.2 Velocity and level of accumulation



The curves obtained by the though-the-book method (TTB 2000) method show that at the end of one reference period, the monthly magazines reach 48% of their total readership, weeklies reach 53%. For weeklies the actual curve is very similar to the curve observed in 1999 for the same titles. For monthlies, the average of 48% at the end of one month is much higher than the 36% observed in 1999. However, the comparison is somewhat tenuous. The TTB 2000 curve is an average of 2 magazines whereas the 1999 curve is based on 19 monthlies.

Only one monthly (Fit & Gezond) appeared in both studies and its reach after one month was comparable (29% in TTB 1999 versus 34% in TTB 2000). The difference is mainly due to “Touring”, a magazine for members of a large drivers association. This magazine accumulates faster than any other monthly magazine. At the end of one month already 65% of its total readership is reached. In 1999, no members magazine was studied.

The curves observed through the diary method show that both weeklies and monthlies reach 81% of their readership at the end of one reference period. This is very similar to the results obtained by Baim, Frankel and Agresti in 1999. The comparison of the level of accumulation after one reference period between the diary and the through-the-book method shows an important difference. This difference is of course due to the difference in calculating the curves. In through-the-book the issue readership is expressed as a fraction of declared total readership, while in the diary method the first time issue readership is expressed as a fraction of observed first time issue readership of all issues. The curves of the diary method accumulate to the 100% reach by definition. In the through-the-book method the accumulation of the issue readership by definition does not reach the total readership figure of the title (the alternative calculation of accumulation expressed as a fraction of AIR is discussed below).

An interesting methodological point is the role of forgetting in the through-the-book method. In our previous paper we stated that using magazine covers as a memory aid minimizes the problem of forgetting but probably doesn't exclude it. By comparing the results of the first and second interview for issues questioned on both occasions, we can now quantify the effect of forgetting. Our analysis is based on the total readership (ever read) question on issue level, as this is the only question that does not contain any timeframe. Some respondents (“drop outs”) declared that they had read the issue in the first interview but failed to mention it in the second interview. For the weeklies, the average drop out rate was 17%, for monthlies 14%. Drop outs may be explained by the forgetting of particular reading moments. However, other “noise” factors may be at work. During the second through-the-book interview, 12% of the reported readings of one of last 4 issues, involved issues that were not mentioned in first interview. While some of them may involve real reading moments between the last diary entry and the second interview, it is likely that false positive identifications are made, e.g. on the basis of issue confusion.

3.3 Influencing factors

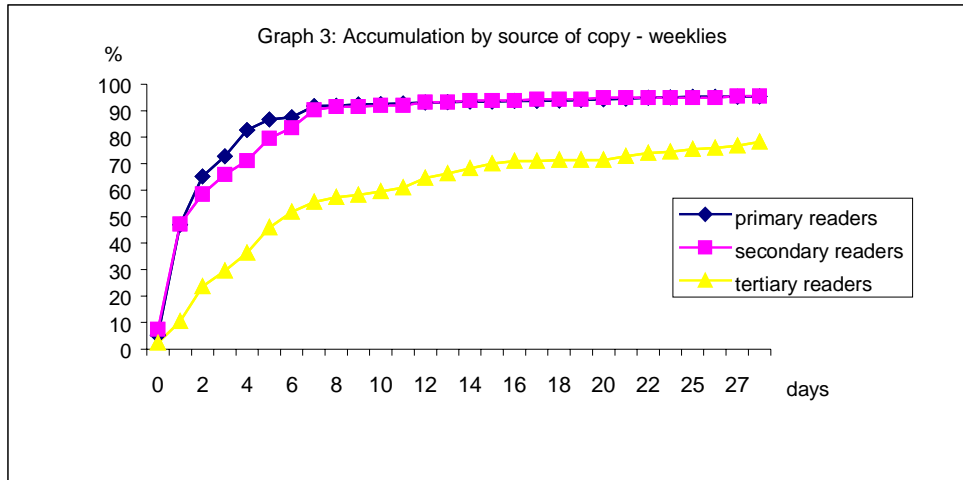
The following analyses are based on the diary observations for weeklies. The number of observation for monthlies was considered too small to analyze in detail influencing factors.

3.3.1 Sociodemographics

Peeters, Debeer and Lanckriet (1999) could not examine the effect of gender on accumulation of readership since their sample included women only. This experiment involved both men and women but gender has no significant influence on the accumulation of readership for weeklies ($F=0.888$; $p = 0.771$). As in the previous study, neither social classes ($F = 1.236$; $p=0.8$), nor age groups ($F=1.215$; $p=0.8$) have a significant influence on accumulation.

3.3.2 Source of copy

In our previous study, source of copy was not asked in order to limit the questionnaire. In this experiment, we asked respondents for each magazine the source of copy in general (i.e. not specifically for each issue). The difference between primary and secondary readers is not significant and only visible within the first week (see Graph 3). Primary and secondary readers do accumulate significantly faster than tertiary readers ($F=4.262$; $p = 0.000$).

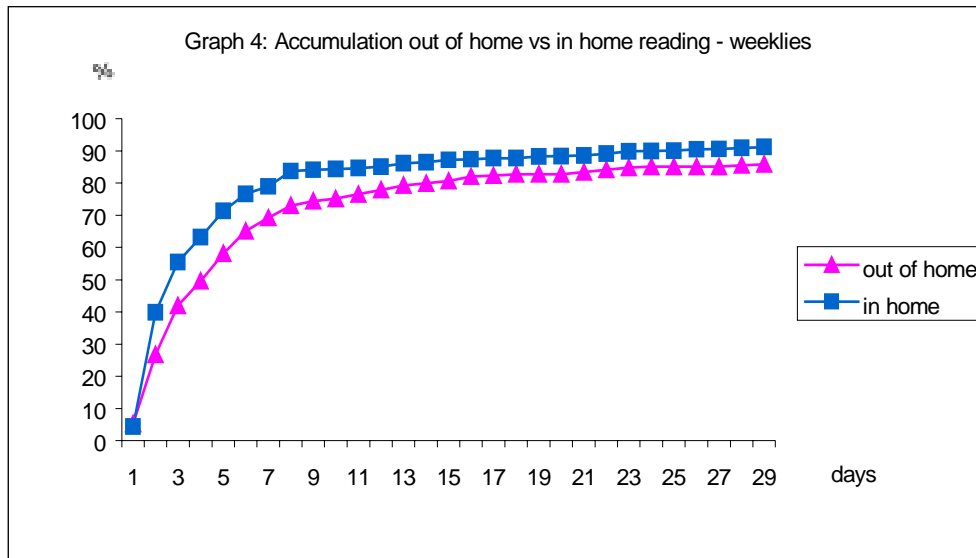


3.3.3 Frequency of reading

Peeters, Debeer and Lanckriet (1999) stated that frequency of reading is one of the most important determinants of accumulation. This observation is confirmed in this experiment ($F=2.780$; $p=0.0$). Regular readers of weeklies accumulate much faster than occasional readers.

3.3.4 Out of home reading

A major advantage of the diary method is the possibility to measure out-of-home reading. Using the through-the-book method it is unlikely that a question relating to place of reading per issue would provide reliable data. Monthly magazines are more likely to be read out of home (28%) than weeklies (13%). Comparing accumulation of in-home and out-of-home reading shows that, for weeklies, in-home reading accumulates somewhat faster, but not significantly ($F=0.030$; $p=0.863$) (see Graph 4).

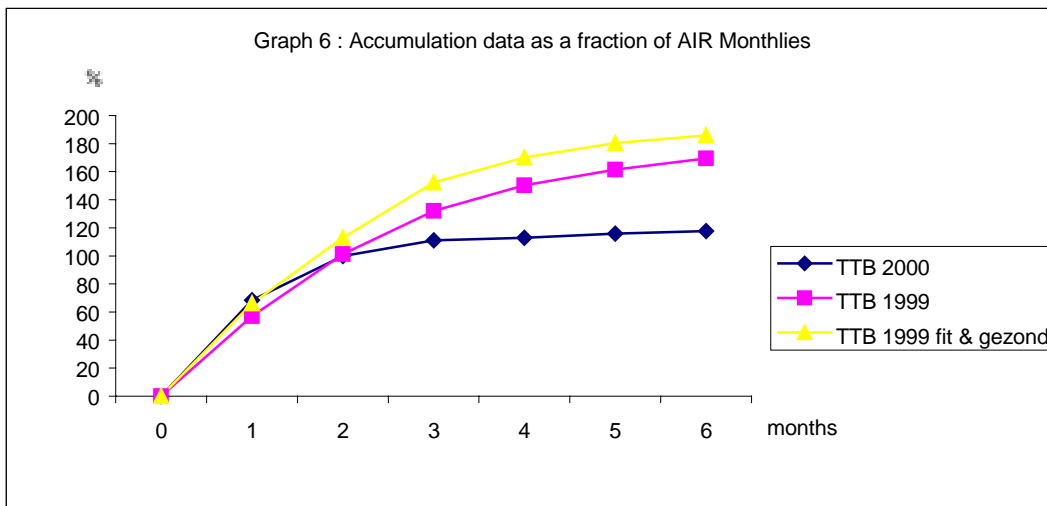
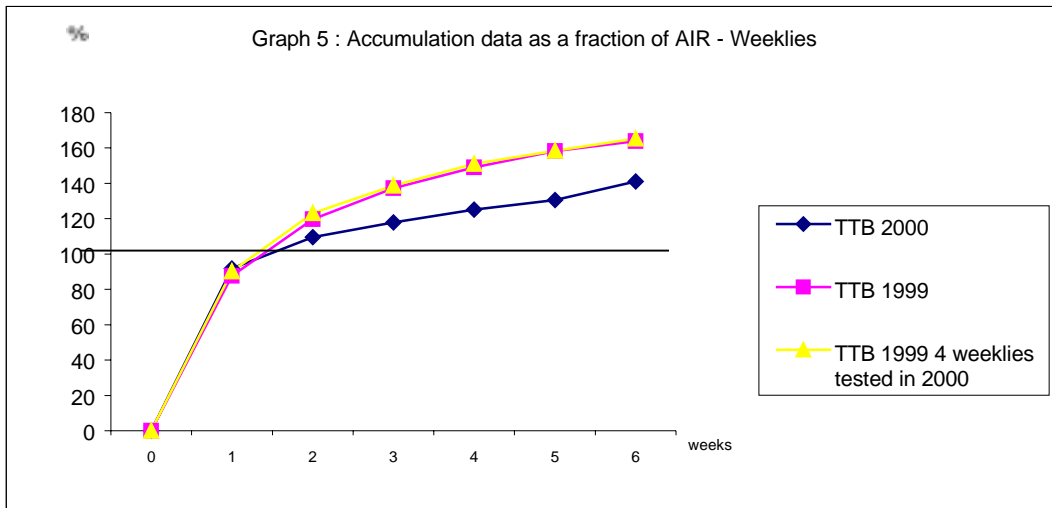


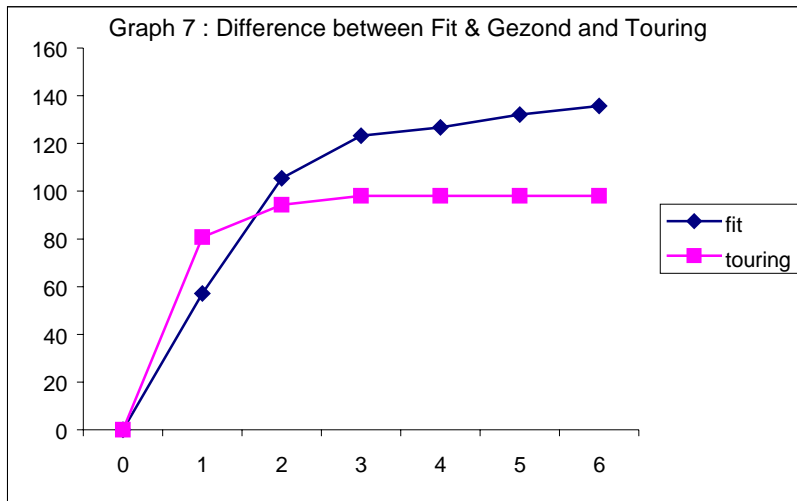
4 Integration in media planning

4.1 Through-the-book data

In order to link through-the-book accumulation data and readership data, Peeters, Debeer and Lanckriet (1999) recalculated accumulation as a fraction of the declared readership during the last week for weeklies, last month for monthlies. This declared readership is the classical Average Issue Readership measure (AIR) used in most audience survey throughout the world. In practice, the number of people that have ever read the most recent issue is simply divided by the number of people that declared to have read the magazine (any issue) during the last reference period. This implies that the AIR readership equals 100% and that accumulation curves are presented as percentages of AIR. Peeters, Debeer and Lanckriet (1999) reported huge differences between magazines but overall accumulation ended well above the reported AIR : on average 64% for weeklies, 69% for monthlies.

In 1999, we stated that the observed proportions were probably overestimated. This hypothesis seems to be confirmed for the weeklies (see graph 5). The comparison of the 1999 and 2000 results for the four weeklies included in both studies, showed that the difference dropped from 65% to 41%. For the monthlies, the observed difference with classical AIR is only 17% (see graph 6). Again this result is mainly due to the atypical accumulation of Touring. The accumulation curve of Touring clearly differs from the curve of Fit & Gezond, a women's monthly focused on health (see graph 7).





4.2. Diary data

In a diary panel method without preliminary recent reading interview, the notions of total readership and average issue readership as traditionally used in mediaplanning are not present. A direct link between diary and Average Issue Readership data AIR can be made assuming that overall issue accumulation corresponds to AIR.

Conceptually, this is not a sound solution. AIR doesn't take parallel (first time) reading into account whereas in reality this is an important phenomenon. In this experiment, we observed on average 4% of people who indicate in their diary reading more than one issue of a title the same day. When considering the reference period : 8% of the readers of weeklies report reading more than one issue for the first time during the reference period. We have less observations for monthlies, but on average 14% of the readers read 2 issues for the first time during the reference period of 4 weeks. This implies that logically, diary accumulation data and recent reading audience data (AIR) cannot be equaled.

It is worth remembering that diary data allow far more detailed analysis of reading behaviour than the traditional recent reading data. They allow an overall analysis of issue readership within the reference period. Both for weeklies and monthlies, 19% of issues read during a reference period is older than one reference period. They also allow an analysis of replicated reading. Some opinion leaders in the industry have argued that the only correct solution to adequately measure both the average audience and accumulation, would be to replace the traditional readership surveys (face-to-face, recent reading) with a diary survey. This fundamental change of methodology, however, would inevitably have a strong impact on all results and readership figures, thus changing the currency used in our business.

5 Conclusion

This paper covers a detailed comparison of the two basic methods of accumulation research and different ways to integrate accumulation data in mediaplanning.

In a previous pilot study (Peeters, Debeer, Lanckriet, 1999), we chose a short, face-to-face interview using the full through-the-book approach in order to minimize demands on respondents. Interviewing the respondents at a central location also lowered costs considerably. In reality, this method is limited by the logistic burden and sampling biases. Nevertheless, results proved to be valuable. In the same period, Baim, Frankel and Agresti (1999) reported on an attractive alternative method using a one week diary on magazine issue reading. Results of both studies were remarkably similar. So we decided to confront both methods in one design and look for the “better” approach.

Our experiment involved readers of six specific magazines in the Dutch speaking part of Belgium. All respondents participated in a four-week diary panel on magazine issue reading. Half of them were interviewed face-to-face using the through-the-book method both before and after the four-week diary period, half was interviewed only afterwards.

A major issue in the 1999 pilot was the manifest unreliability of the ‘first time read’ responses. In order to rule out inadequate wording, we used another phrasing in this study. Unfortunately, the conclusion remains: our episodic memory is not equipped to answer silly “first time read” questions...

In choosing the through-the-book approach in 1999, we assumed that recognition of magazine covers minimizes forgetting. In this experiment, we compared the results of the face-to-face interviews before and after the panel for overlapping issues. 17% of the people who declared having ever read an issue of a weekly in the first interview, did not mention that issue in the second interview (14% for monthlies). This observation might be taken as an indication of forgetting, though other noise factors may be at work.

One of the problems in composing a diary panel method is the return rate. Having an interviewer hand over and collect the diaries, has a positive impact on the return rate but a negative impact on costs. Another potential problem is the quality of the survey in terms of the respondents motivation and accuracy. In fact, some of the diaries were incomplete. More important, a thorough control showed that in 12% of the cases respondents ‘forgot’ to indicate the ‘first’ reading moment for the most recent issues. Although these ‘mistakes’ can be corrected, this observation indicates that a diary method also has its inaccuracies.

An analysis of influential factors on the diary data for weeklies, confirmed the 1999 conclusions: no effect of age of social class, a clear effect of frequency of reading. This experiment led to additional observations: no effect of gender, a significant effect of source of copy (primary and secondary readers opposed to tertiary readers), no significant effect of out-of-home reading. Overall, both methods seem to measure fairly well the same reality.

A comparison of accumulation curves calculated on the basis of diary and through-the-book data, shows that the level and to some extent the speed of accumulation differ. In fact, the differences on question level are small. Differences essentially come down to the use of different denominators: total readers or average issue readership in our through-the-book approach, total declared first time issue readers in the diary approach.

In 1999, through-the-book accumulation data were linked to traditional recent reading audience estimates through the use of average issue readership. This “fusion” of measures from different research traditions, led to the annoying result that the final accumulation level was on average more than 60% higher than the traditional audience measures. The current experiment confirms this gap, albeit at a lower level. In our opinion, the screen-in method of these studies probably influences the gap. Yet, the consistency of this gap, and the important difference between magazines, suggest magazine-readership dynamics that go beyond our current understanding.

A pure diary setup does not include recent reading questions. Linking diary accumulation data to audience data, can be done but only on the assumption that overall issue accumulation corresponds to AIR. This is an attractive and very straightforward way of proceeding for those who don't mind the conceptual simplifications: traditional average issue readership excludes parallel reading, the diary data do take parallel reading into account.

Overall, there is no clear winner. The diary method is probably more economical and allows more detailed analyses of reading behaviour. As all methods, it has some inaccuracies. The question is whether these inaccuracies are randomly distributed over magazines and can be monitored or controlled. Another, more important question is whether or not the publishers, agencies and advertisers are willing to accept some degree of ‘conceptual inadequacy’ in order to integrate recent reading audience data and diary accumulation data in media planning.

References

Baim J., Frankel, M.R. & Agresti, J. *Magazine Audience Accumulation: Development of a Measurement System and Initial Results*, Worldwide Readership Research Symposium 9, 1999.

Brown Michael. *Effective Print Media Measurement – audiences ... and more*. Ipsos-RSL, Harrow, 1999.

Peeters, S., Debeer V. & Lanckriet, T. *Magazines Deserve Time – The Build-up of Magazine Audience over Time*, Worldwide Readership Research Symposium 9, 1999.

Pincott Gordon. *An Examination of Wear-out and Exposure over Time*, 1991.

Die Dimension Zeit in der Mediaplanung [The time dimension in media planning], Axel Springer Verlag AG, 1992.