

WHAT DO WE MEASURE? FOUR SHORT PUZZLING STORIES

Jean-Louis Marx, AEPM

All men make mistakes, but only wise men learn from their mistakes.
Winston Churchill

Introduction. Doesn't time fly!

Apart from Lucy, our distant relation who lived a few million years ago in the African Rift valley, I can't think of any being as closely scrutinized as the reader. How can one track him down? How can one recognise and identify him? How can he be construed from bits and pieces collected by scores of field interviewers?

A beginner would think that with a strict survey procedure, a properly built sample and professional interviewers, the reader can't escape him.

How innocent!

A few years later, complaints from publishers dissatisfied with their audiences, or else conflicting tests and analyses, are likely to calm down his enthusiasm. The researcher ought to keep questioning his procedures and methodologies. It's a matter of professional ethics. Therefore he is naturally led to wonder about the very subject of his work:

- What is exactly a reader?
- What am I really measuring?

Since 1993, the AEPM which I have the pleasure of directing, measures in France the audience of magazines. Among all the surveys and tests we have conducted in the last 10 years, I picked up four that to my mind epitomize the pitfalls of audience monitoring of print.

Let me tell you those four stories which are rather puzzling for a researcher.

Story N°1: In this story research finds there are too many readers around.

We thought that technology would spread automation of media audience monitoring, following the case of television. The radio was testing a recorder clock. The GPS technology would soon make it possible for poster advertising. What credit would one give to press audience monitoring relying just on memory?

In 1995, we decided to conduct with TNS-Sofres a test to check whether one could get the magazine pages scanned by their bar codes, by a panel of people who had been given a wand. The point was to test if the panellists would comply with such constraints. They were asked to scan the magazines at the beginning and at the end of each time they took one, the same way TV panellists do when they sit down in front of their TV set and when they move away. The test lasted two months and the sample for this experience had been limited to 20 people. The instructions were generally followed.

As we were anxious to assess the results we would get, we also asked the panellists to hold a readership diary during the first month of the experience. At the end of the test, they were asked to answer the standard audience face to face questionnaire.

We found the diary and the scan information matched pretty well. 14% of the readings reported in the diary were missing in the scan, either because the magazine had no bar code, or because it wasn't clear enough to be read by the wand.

However, there were discrepancies between the results of the scan and the face to face interviews. Here is the AIR readers chart comparing the two approaches.

Face to face/Scan comparison	39%
F to F moves back the last reading occasion	43%
F to F moves forward the last reading occasion	0%
Readership statement F to F and not scan	9%
Readership statement scan and not F to F	9%

The data suggests the face to face interview has a powerful telescoping effect. The respondent’s memory draws the event closer in time. As a consequence, there is a significant gap between face to face and scan for the number of magazines read within AIR and for those read the day before.

	AIR Number of mag. read	YESTERDAY Number of mag. read
Face to Face	7,3	1,9
Scan	4,6	0,9

Story N°2: In this story, research says readers are missing

The same year, we set out on checking a question on the quantity of reading, having in mind to release a page exposure indicator besides the audience one. The question was simple:

<p><i>How many pages of this issue did you read, browse through or look up?</i></p> <ul style="list-style-type: none"> - All the pages, or nearly all of them - About 75% - About 50% - About 25% - Less than 25% - None or hardly at all

Since memory is not a very reliable tool, we had cautiously restricted the question to magazines read the day before the interview.

The test was conducted by Ipsos who interviewed 393 people. The interviewees were invited individually at the institutes’ premises to take part in a survey. They were purposely left waiting for half an hour in a room fitted with four cameras, and they were offered magazines to keep them waiting. So we could precisely track the quantity of pages they read or browsed through, and compare it with their statement.

The next day, these guinea pigs were visited at home by an interviewer who asked them the audience questionnaire in the standard interview conditions.

The total reading time per person was of 32 minutes. Each interviewee had read an average of 3.7 magazines at Ipsos. We had noted 1453 readings. There was more than enough to calibrate our question on the quantity of reading.

But we hadn’t expected a great deal of readings would be missing in the audience questionnaire. Indeed, our interviewees declared the next day only 32% of the readings of the magazines they had actually read the day before. Two thirds had vanished.

The result surprised us and we had to stop the experience and the page exposure assessment. No element in the experiment we had conducted could account for the lack of the majority of the magazines read the day before in the interviewees’ declarations. We could only barely guess that interviewees answer the audience questionnaire depending on the perception they have of their reading habits rather than on the memory of what they actually read. Therefore the odd readings, as were those of the Ipsos test, were forgotten for the most part.

Story N° 3: In this story research finds interviewees are not sure about the size of their family.

Following the release of the first series of results, some publishers got concerned about the discrepancies between RPCs within the same family of magazines. In 1996, we conducted an analysis on the consistency of this indicator. The name of the game was to check out for each magazine the consistency of three numbers:

- primary readership (the number of readers belonging to a buyer or subscriber family)
- paid circulation for France
- the number of persons aged at least 15 (this is the reference population) in families of primary readers.

The consistency of the data means the RPC of a magazine among primary reader families cannot be greater than the number of persons aged 15 or more.

Let's take an example: a magazine has 1,500,000 readers, 900,000 of which belong to buyer or subscriber families (primary readers). Its paid circulation for France is of 300,000 copies. Its RPC among the primary reader families is consequently of 3.0, which means there are 3 primary readers (aged 15 or more, as studied by the survey) per paid copy within primary reader families. But the survey made a census of all the family members belonging to the interviewees' families. And we therefore know that in the families of primary readers of the magazine, there are only 2.5 individuals aged 15 or more. If we suppose that all the members of the family are readers of the magazine, we still have an excess of 20% of primary readers (3/2.5).

The analysis covered 110 magazines. We only found 18 magazines for which the data tallied, and 92 for which there were more primary readers than individuals in the reading families.

It was then useful to find out if the AIR figures the most consistent with the size of the family were more frequent in some cases. We also discovered that AIR results were all the more consistent as:

- the periodicity of the title was frequent
- its circulation was large
- its readership was regular
- it was picked up many times

The more a magazine is read regularly (whether it comes out more often, whether readers read all the issues, or whether they pick it up many times), the more accurate are its AIR findings.

What could be happening, then? One must discard the idea that interviewees can't get the size of their family right or that they are making up readings of magazines.

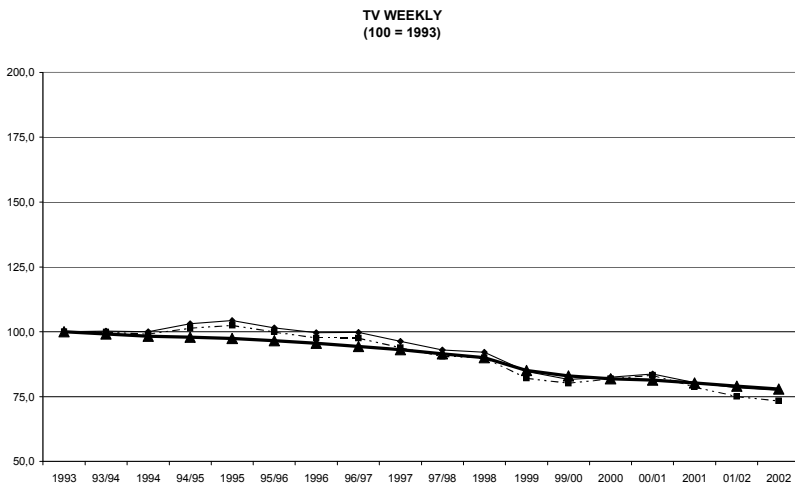
Several elements could explain these inconsistencies:

- There is the telescope effect we already mentioned: Statistically speaking, when an interviewee reads a magazine frequently, an answer closer in time is more likely to be an accurate one.
- As for the recency question, the source of copy question is not always very accurate. Are taken into account not only the magazines bought or subscribed to by the interviewee himself, but all those of the family. The reader of a magazine may very well not be aware of how another member of his family got a magazine. However statistically speaking, the larger the circulation of a title, the more likely answer of the interviewee concerning the purchase of the magazine by himself or another member of the family is to be accurate.
- Still, the AIR reflects readership of an issue which can date back to a while. RPC are based upon the average circulation of an issue and therefore discounts replication, which is the readership of an issue over several publishing periods. It would sound logical to take these extra readings into account for the RPC.

Story N°4: Why is it that the readership always drops when circulation increases, and never the reverse?

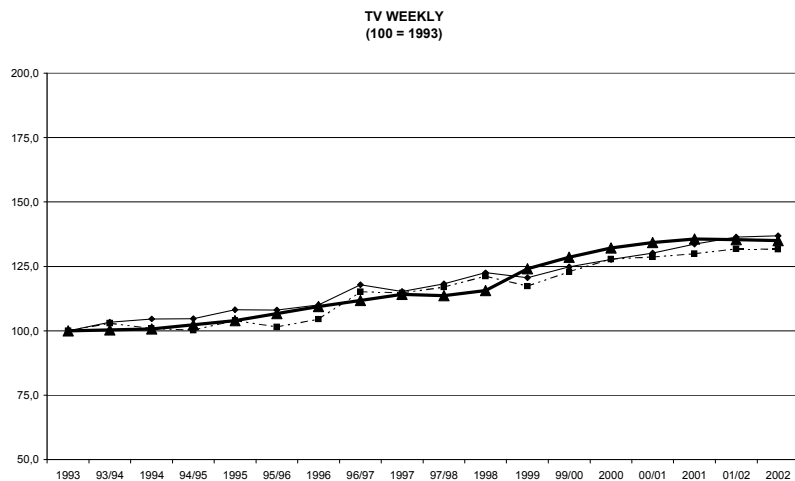
The issue of each series of results always triggers off phone calls from worried publishers. They cannot understand how their magazine can lose tens of thousands of readers when at the same time circulation increases regularly by a few points every year. Their reaction is generally to question the methodology of the survey which, to their mind, fails to reflect the how well their magazine is doing.

According to those phone calls at least, the case where the number of readers grows faster or declines more slowly than circulation seems to never happen.



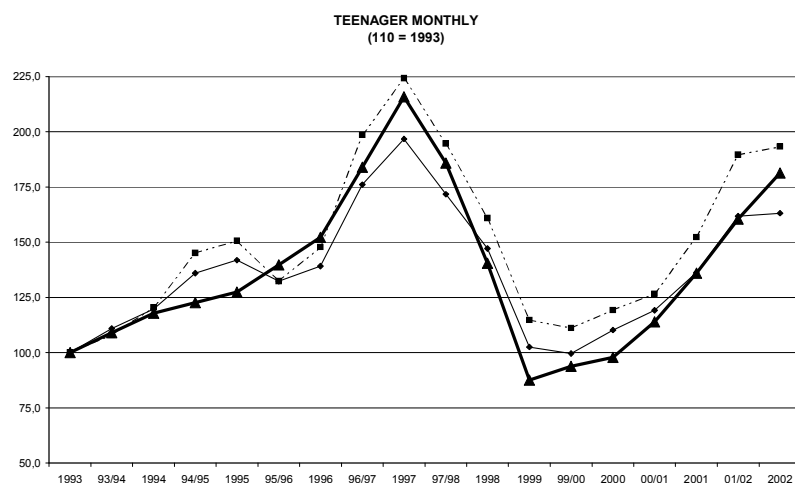
This peculiarity was worth looking into. Does the audience develop independently from circulation? Or on the contrary, is there some correlation between the two? With ten years data (1993-2002) it could be useful to draw the readership curves and to see if they were in keeping with the circulation curves.

The analysis covered 135 magazines whose results were published in 2002 and for which we had a series of at least four consecutive readership figures with a constant periodicity.



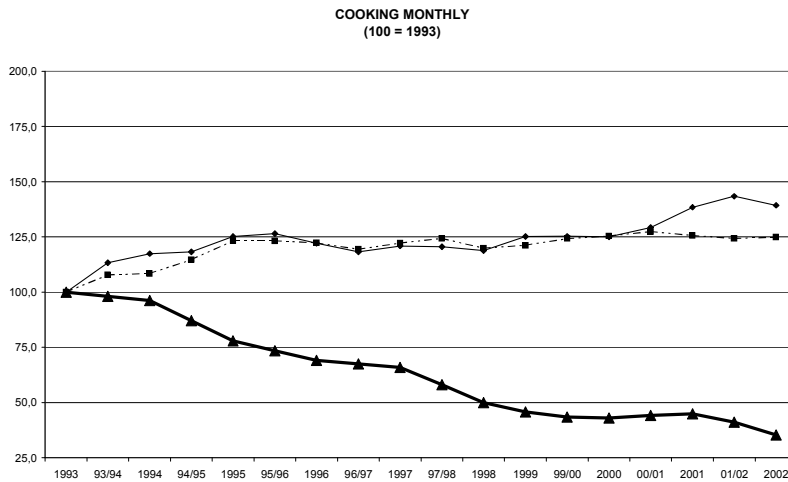
The rule is common sense. The audience is in keeping with circulation. For 103 of the 135 magazines within the scope of the analysis, the curves are parallel, as for those two TV magazines one gaining, the other losing readership.

The curve in heavy type is that of circulation. The light type one is the total AIR and the dotted one is the primary AIR.



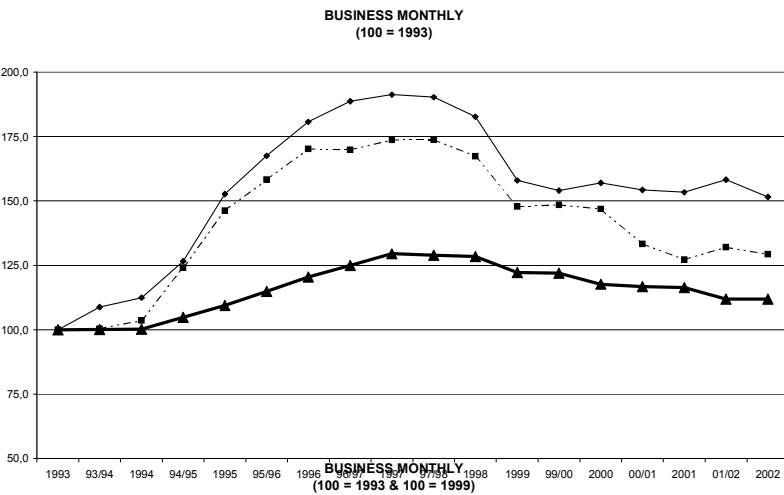
We can also mention this teenager monthly whose circulation and readership have had ups and downs over the period.

Publishers who worry about their results are focused on the last readership figures and tend to disregard the statistical error when reading a survey - in fact, a slightly negative trend often follows or announces a positive one.



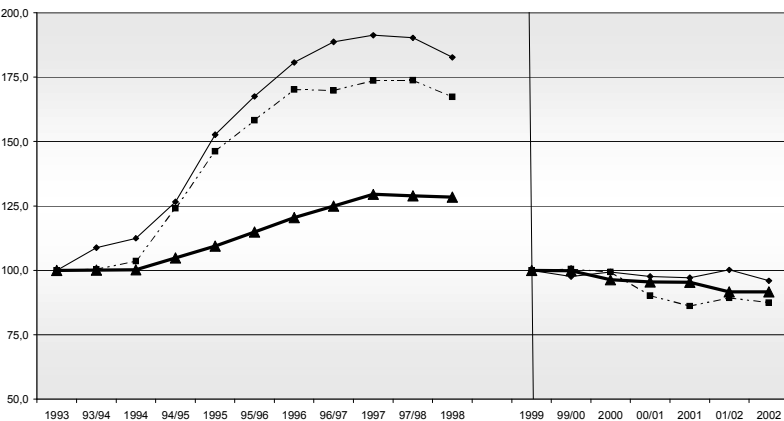
The comparisons between readership and circulation are comforting, since globally the readership study translates the circulation trends. However in some cases one can but wonder. As for the case of this cooking magazine which lost two thirds of its circulation in 10 years, but shows complete stability of readership in the past 8 years.

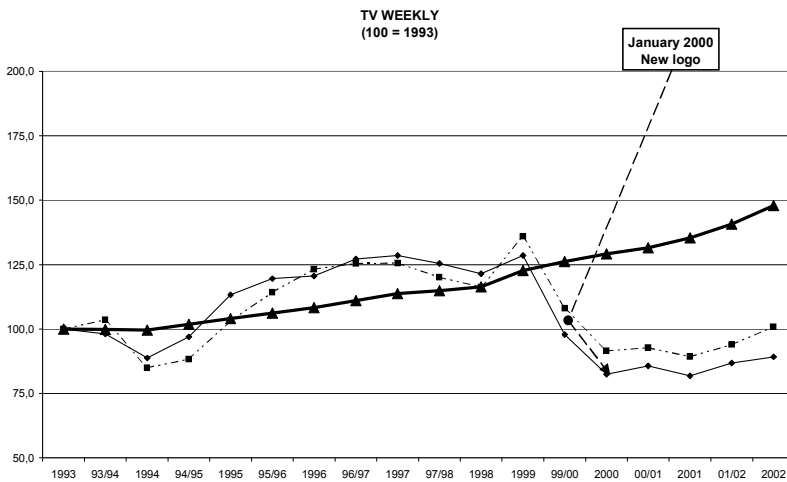
It must be that part of the readers of the magazine stopped buying it but still read it. They must have kept the old issues to build up their cooking encyclopaedia, as it were.



Another case conflicts with the rule. When a magazine is new, it takes time to reach its optimal readership level. As for the case of a business monthly launched in 1991.

Between 1993 and 1997, the RPC of the magazine increases. If we start a new curve from 1999, the readership and circulation trends are parallel again. Perhaps you need to wait until readers stock up with old issues. But it also takes time before the awareness of the magazines is sufficient to help occasional readers to identify it and to declare readership properly, instead of declaring an other one when answering an AEPM interviewer.





The brand image of the magazine is an important element. We mustn't forget the interviewees are asked to identify a title, being shown only its logo. This is what can happen when a magazine changes its logo:

This TV magazine changed its logo in January 2000. In one year we counted about 40% less readers .

Conclusion: Much ado about nothing

From all these examples, one can see that the key problem in readership studies is the lack of reliability of people's memory. There are neither cameras nor recording clocks in our interviewees' brains. When answering our questionnaires, they seem to rely on their habits rather than actual facts. Must one complain about it?

Could they tell us for instance, that two weeks ago they exceptionally went through a health magazine in the waiting room of their doctor? This is the kind of thing we ask them.

Can they really remember precisely when they first read the cooking monthly they looked into three days ago in their archives to find the calf's head recipe (the one "Uncle Jacques" likes so much), i.e. eight months ago when it was published? Is this the kind of thing we should ask them ?

Should they also be able to identify this automotive magazine which has just redesigned its corporate identity? They don't buy it, nor are they subscribers. But they never fail to flick through it whenever they can, as they did 2 weeks ago at a friend's, browsing the special car show issue which was six months old.

Must we draw the conclusion from these stories that Lucy escapes us, that the piecemeal information the interviewers collect in the field are not enough to fathom the reader? In a word that readership is a lottery?

It seems to me that we are misled if we are trying to discover a sort of standard reader who would be the origin of the species, as is our distant relation. There is no such thing as a standard reader. There are masses of them, as there are numerous occasions of reading the press, and as is the diversity of the cultural, information or entertainment range it has to offer. The description of these audiences and their relationship with the titles they claim to read is therefore essential.

But the readership study has two other key functions.

- First it is a barometer on the health of dailies and magazines, each new release of results bringing information on the trend of each title. It is therefore necessary to maintain a constant methodology to ensure results are comparable from one survey to another. But it is as important to train interviewers and to monitor their work so as to avoid biased results which could turn into disasters.

- The second function of the study, its *raison d'être*, is to be an exchange currency. The confidence of the various players of the market in their currency is of the utmost importance. Readership must be sterling. If the market is confident, one can then reverse the proposition: instead of wondering " is what we measure really readership?", we can state: "readership is, by definition, what we measure".