

ARE CHANGES IN READERSHIP PRECEDING CHANGES IN CIRCULATION?

Ingemar Lindberg - Sifo Research & Consulting AB / Montilo Media AB

Summary

The relation between circulation and readership has always been a matter of great concern. Not only to publishers, when changes (increases, that is) in circulation are not simultaneously and proportionally reflected in changes in readership, but likewise for media researchers in their subsequent endless 'firefighting'. The topic has been the subject for many a study over the last decades, particularly on modelling (see the work-in-progress comprehensive bibliography¹ being prepared for NRS in the UK by Michael Brown). The efforts seem to have brought limited comfort or progress as far as modelling is concerned.

Most authors have tried to build models with circulation as the independent variable. A few authors, like Jane Perry² and myself, have suggested that changes in circulation tend to lag behind changes in readership, and that circulation therefore might rather be considered as the dependent variable. This paper set out to investigate which comes first: the egg or the chicken.

This was done by regression analysis of circulation and readership of 48 publications over a 13 year period. The collection of the readership data used for the analysis, differs completely from other data being used in previous analysis by other researchers: (1) a direct *pure random sample* from the continuously updated official population register, (2) a *self-administered questionnaire* mailed to and from the respondent, (3) AIR being calculated from a non-equidistant semi-nominal 7-point *frequency scale (with miniature mastheads)* only.

The regression analysis shows a very high degree of correlation: 19 (40%) of the R(2) values are above .80, and as many as 30 (65%) are above .60. There are 8 (17%) cases between .30 and .59, and 10 cases where there is hardly any correlation, due to invalidating circumstances like changes in issue frequency, increased distribution of free copies, heavy editorial changes affecting structure of readership and rpc etc.

The egg and chicken problem: The analysis show some evidence that changes in readership do precede changes in circulation, rather than the reverse. The period by period readership changes are, however, probably too small in relation to the sampling error, and the time periods analysed (12 months) too long, to expect a clear-cut relation. The many external and internal factors affecting the circulation/readership relation, and in particular the continuous changes in rpc, should make it difficult - if not impossible - to get to a working model at all, for many publications.

Rather than trying to model circulation and readership data, the author suggests a 'dynamic' graphing method, mapping indexed circulation and readership data over time, the DYNAMAP™. This method should be a powerful diagnostic instrument to show the direction and inter-relation of circulation and readership (and rpc) in a time perspective (several years). With this map the publisher and the researcher can, hopefully, in fruitful and creative co-operation, try and mutually understand the reasons for the actual development, rather than debating if readership figures are right or wrong.

Publishers should consequently spend much more resources on detailed analysis of their subscription files and other sources of information, to see where their copies go. Only then can they begin to understand how reading is created and why the number of rpc develops the way it does.

Michael Brown¹ also relates to this in one of the 14 points in his Summary of conclusions: "...An adequate model of the environment of readership/circulation relationships is provided by considering (a) the 'supply' variables which determine how many reading opportunities a copy provides and (b) the 'demand' factors which determine whether a reading opportunity becomes a reading event".

¹ Brown, Michael "The relationship between readership and circulation" First working paper, 1996

² Perry, Jane "Some further thoughts on readership and circulation" Worldwide Readership Research Symposium VII, 1995

1. Background

The relationship between readership and circulation is a matter of constant interest and concern, for publishers and media researchers alike. The problem lies in the fact that circulation and readership figures do not always move in the same direction. The subject has been covered at every Symposium since the first in New Orleans in 1981, and in a large number of papers before that.

In the UK the NRS in 1994¹ initiated a major research project to systematically review "Readership, circulation and readers per copy". In his 'First working paper' Michael Brown gives the arguments for the study:

"...it was not unusual for NRS staff to have to spend considerable time and effort in meeting criticisms of the data on occasions of apparently inexplicable variations in readers-per-copy estimates, particularly when demonstrable circulation gains were not mirrored in comparable increases of average issue readership. If any more systematic, generalised explanation of the numerical relationships involved could be offered, it was suggested, then less resources might need to be devoted to 'firefighting' and, more importantly, general reliance on the NRS' estimates could be increased...". This argument is easily recognised by all media researchers.

One of the first steps in the NRS project was to make an overview of previous research in the area. As a separate document within the project Michael Brown is therefore putting together "The relationships of readership and circulation: an annotated bibliography". This 45-page document is already (June 1997) most comprehensive, but will be further extended.

2. The purpose of this paper

Over the years many authors have described the results of efforts aiming at building models for the prediction of readership data from circulation figures. The reason is of course that circulation data should be more accurate and reliable than readership data, because of the sampling errors and other shortcomings in readership surveys. The purpose of the exercise would be to get tools for 'correction' of readership data, to make them conform with the changes in circulation data (this practise seems to have been abandoned in later years). Circulation data have thus always been considered as an **independent** variable.

A few authors, like Jane Perry² and myself, have suggested that changes in circulation may rather tend to lag behind changes in readership, and that circulation therefore might rather be considered as the **dependent** variable. This paper initially sets out to try and shed some light on which comes first: the egg or the chicken?

3. Data available for the analysis

Circulation data

Circulation data from the Swedish ABC, which are considered to meet very high standards, were available for most calendar years in the analysis. In a few exceptional cases, where a single year's figure was missing, an average value of the two adjacent years was used.

Over the years some changes have been made in the definition of audited circulation, but these changes should not distort the analysis in any way.

For a few magazines a substantial part of their circulation is being sold outside Sweden. This should not affect the general findings of this analysis, but when analysing individual titles more closely, this phenomenon should of course be considered.

Readership data

The readership data for this analysis were taken from our media and marketing survey ORVESTO Consumer, conducted continuously since 1971. In the absence of a JIC, this survey serves as the equivalent to an NRS in Sweden.

The ORVESTO survey differs in most respects from other similar readership surveys around the world (For full details about the whole procedure, and the research, thinking, and arguments that lead to it, please see my second symposium paper³):

¹ Brown, Michael: as above

² Perry, Jane: as above

³ Lindberg, Ingemar "Self-administered questionnaires and frequency of reading: an optimal combination to measure readership?" Worldwide Readership Research Symposium VIII, 1997

1) The sample

The sample for this survey is a *random sample* directly drawn (in one step) from the continuously updated official Swedish population register. It is a systematic sample as far as age is concerned, since the register is in order of age and every n:th entry is selected. Because of the sample size (30,000 completed interviews/year) the relation male/female is also basically in perfect balance, as is the geographic distribution.

2) The data collection procedure

The information (media consumption as well as vast amounts of TGI, lifestyle and other classification data) is collected in one step (a procedure followed basically unchanged since 1971), by means of a *self-administered questionnaire*, mailed to and from the respondent. (In later years detailed information on TV and radio is also collected via a subsequent 8 days diary to the same respondents.) It is thus a *pure single source survey*. The possibility to collect data in this way (with response rates of 65-70% of the gross sample) is probably unique to the Nordic countries, and Sweden in particular.

3) Establishment of AIR (average issue readership)

In most countries the AIR figures are the result of a Recent Reading (RR) question (read in the last publication interval). This result is then used (as a 'true' value) to attribute probability values to the various groups in the frequency question being used.

When we started to engage in readership surveys in the late 60's we found it absolutely essential to be able to incorporate a considerable amount of classification data for target market selection purposes. In order to find the optimal combination of questioning technique (recent reading vs frequency of reading) and data collection procedure (personal, telephone or postal/self-administered), we made a major pilot survey⁴.

As a result of this pilot survey we found, that people were quite able to answer frequency questions about readership, in a consistent way, with self-administered questionnaires, with very high response rates. This basic procedure has been followed ever since. The establishment of AIR is calculated *directly* from the results achieved from the frequency scale, using apparent probabilities.

4) Using a non-equidistant semi-nominal 7-point frequency scale

Since 1984 we are using a 7-point frequency of reading scale, which is semi-nominal and non-equidistant. It is printed, in our questionnaire, together with mini-mastheads of the publication names, and the respondent is asked to state:

"Of the issues published I read:"

- No issues
- Almost no issues
- Approximately 1 out of 4
- Approximately 2 out of 4
- Approximately 3 out of 4
- Almost all issues
- All issues

		Av de nummer som ges ut läser jag:						
Antal nr/år (ca)		Inga nr	Nästan inga nr	Ca 1 nr av 4	Ca 2 nr av 4	Ca 3 nr av 4	Nästan alla nr	Alla nr
AFTONBLADET	vard 300 sönd 50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
EXPRESSEN	vard 300 sönd 50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

For each title the *number of issues* per year is stated (between the logo and the boxes) for the respondent's reference.

5) Restrictions/limitations with the data

Universe

The universe in this analysis is restricted to those 15-70 years old. (Since 1994 the upper age limit in the survey is 79 years). The limitation to 70 years may have a marginal effect on rpc for magazines with an ageing audience.

Titles not included in the analysis

It would have been desirable to include a lot more titles in the analysis, for which readership data were readily at hand. It was therefore a disappointment to find, that official circulation data was only partly available: sometimes individual years missing, only first half year's data available, or mixed first and second half year's data etc.

⁴ Cerha, Jarko " Selective Mass Communication", 1967

4. Regression analysis of circulation and readership data for 48 publications over a 13 year period

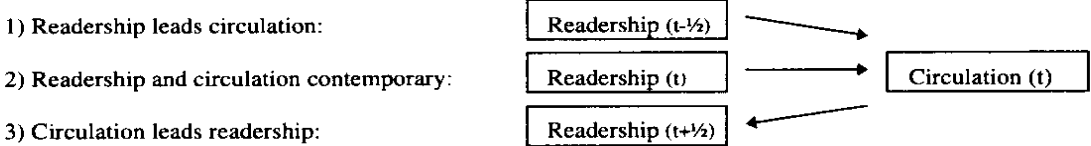
The analysis procedure

The procedure was to do a straight-forward regression analysis of circulation and readership data for 48 publications over a 13 year period (1984-1996). A linear model was used. This procedure is supported in Michael Browns¹ Summary of conclusions, p.7: "The relationship of readership to circulation is basically *nonlinear*; but a straight line model may well provide a perfectly workable approximation".

(It may be noted, that the next step in the NRS procedure will be to analyse the readership and circulation data for 81 periodicals over the EML period 1984-1991.)

In order to analyse the 'egg and chicken' problem the test design was to compare the R(2) correlation values as follows:

- 1) $C(t)/R(t-1/2)$ = Circulation figures for calendar year (t) vs Readership figures for the 12 months period starting 6 months before the calendar year
- 2) $C(t)/R(t)$ = Circulation figures for calendar year (t) vs Readership figures for calendar year (t)
- 3) $C(t)/R(t+1/2)$ = Circulation figures for calendar year (t) vs Readership figures for the 12 months period ending with the first half of the next calendar year



If the R(2) values were higher for 1) than for 2), it would prove the hypothesis that changes in circulation tend to lag behind changes in readership. If, on the other hand, the R(2) values were higher for 3) than 2), it would prove the opposite.

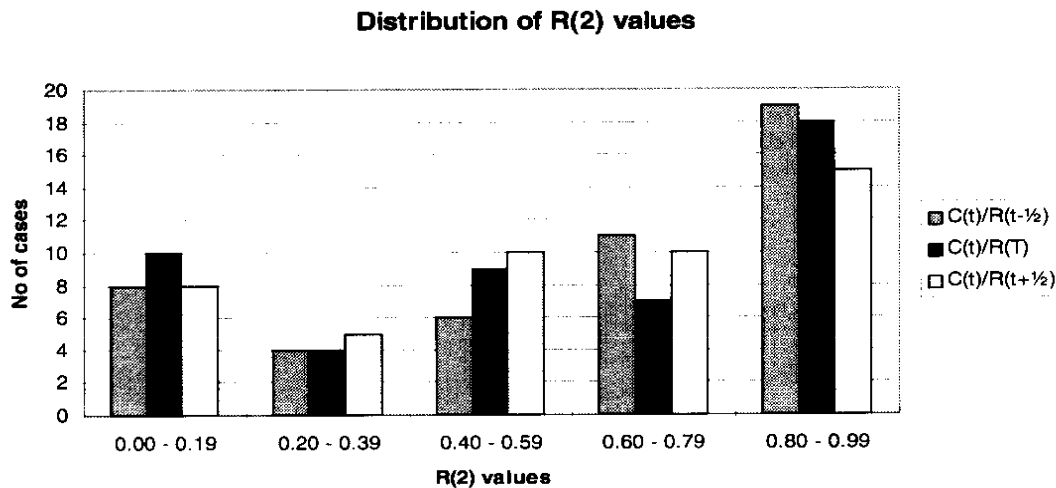
Results of the analysis

The results are shown in detail in Table 2 and in summary in Graph 1. The figures for the graph are as follows:

Table 1

R(2) values:		$C(t)/R(t-1/2)$	$C(t)/R(T)$	$C(t)/R(t+1/2)$
0.00 - 0.19		8	10	8
0.20 - 0.39		4	4	5
0.40 - 0.59		6	9	10
0.60 - 0.79		11	7	10
0.80 - 0.99		<u>19</u>	<u>18</u>	<u>15</u>
Total		48	48	48

¹ Brown, Michael: as above



Graph 1. Distribution of 48 R(2) values

The approx. 8 cases with no regression should be disregarded for various reasons, which leaves us with some 40 usable cases. In some cases the issue frequency has been changed one or more times. In others circulation has been boosted or the amount of free copies has been increased dramatically. External factors, like the sharp economic recession starting 1989, has made the readership/circulation relation to more or less to collapse for some titles etc.

The major impression is otherwise that the R(2) values are generally very high, also in relation to other known similar analysis.

The difference between the three sets of R(2) values is very small. There is maybe a slight tendency in favour of the circulation time lag idea (first and third columns), but hardly significant. The first column (readership leads circulation) shows 30 cases with R(2) values above 0.60, while only 25 in the second column (readership and circulation contemporary). However, if one compares the three R(2) values for each publication, they are very close in most cases.

This leaves us with the situation of no support for deciding which variable is dependent or independent, and the question in the title of this paper remains unanswered.

Why are there no differences?

The results bring us back to the real world. The 30 or so titles (out of 40) with the highest R(2) values are publications with stable development, either progressive or regressive, and no dramatic changes have been made either editorially or with circulation. **A regression analysis of the kind undertaken here, cannot therefore be expected to bring forward any significant differences between the three sets of R(2) values, which could give an answer to our question.** Less so then in the cases when the correlations are smaller, not significant or negative.

Overall this leads us to believe that the expectations to be able to build any workable model may be very small indeed. Guy Consterdine⁵ refers to this (in his San Francisco paper 1993) "... it is impractical to build a mathematical model which predicts the rpc for a given publication with useful accuracy...".

It also feels appropriate to quote a few points from Jane Perry² 1995 Berlin paper:

"...The publisher has much more control over his circulation figures than he has over his readership..."

"...There are two major causes of circulation changes: external factors and actions taken by the publisher himself..."

"...The relation between Readership and Circulation is linked more directly to the individual history and circumstances of particular titles..."

If modelling attempts do not seem to lead anywhere, is there any alternative?

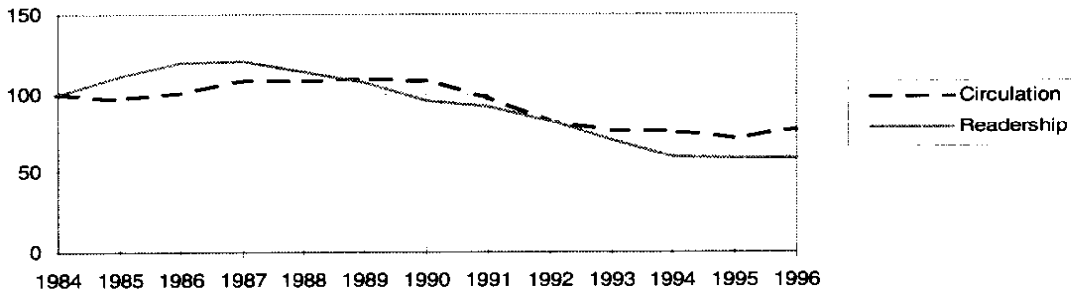
⁵ Consterdine, Guy "What determines readers-per-copy patterns for UK magazines?" Worldwide Readership Research Symposium VI, San Francisco, 1993

² Perry, Jane: as above

5) Instead of a model: The dynamic DYNAMAP diagnostic tool

In the early preparation work for this paper, graphs of the *indexed* relations (*over time*) between readership and circulation (base 100 for 1984) were made for all the 48 titles over the 13 years in our analysis. Generally the curves are quite smoothly following each other, while sometimes there are sudden shifts and outbursts.

Take a look at the 'conventional' way of presenting this type of information in this example:

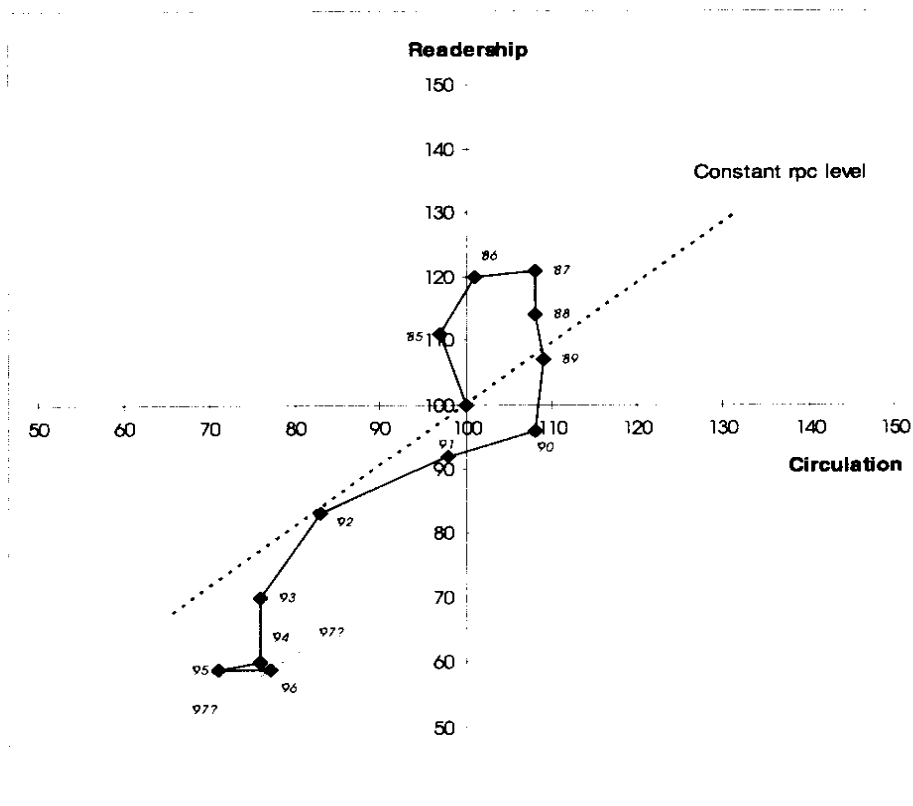


Graph 2. Readership and circulation for a business weekly; both indexed to 100 for 1984.

What does the graph really tell us? How can the information be used and interpreted? Is there a way to illustrate what is happening in a more dramatic and dynamic way? Some approach that can make the publisher and the researcher jointly try and understand why readership, circulation and rpc develop the way they do, rather than fighting about who is right or wrong.

Incidentally, when looking at graphs of *indexed* readership and circulation figures (*against each other*), while preparing the R values for the regression analysis, the idea of adding the time dimension to that kind of graphs was born. In doing so, very interesting patterns emerged, of the relation between readership, circulation, rpc and time, which seem to be a very promising approach to better understand the life and performance of a publication:

Business weekly



Graph 3. DYNAMAP™ of a business weekly. Readership and circulation indexed: 1984 =100.

The map starts at the crossing of the two axis, where circulation and readership are both indexed to 100 (in this example for 1984). By linking the year by year dots for circulation/readership indices, in chronological order, the pattern of how the 'road' for this publication is evolving over time emerges.

Note that the dotted line, at a 45° angle through origin, indicates at what combined values of readership and circulation the number of rpc remains constant (as compared to 1984 in this example). Points above the dotted line indicate increasing rpc levels, while points underneath show diminishing rpc. By choosing a different year for index 100 the curve will move in the map, while maintaining its basic shape.

Back to the example in graph 3: What does it show?

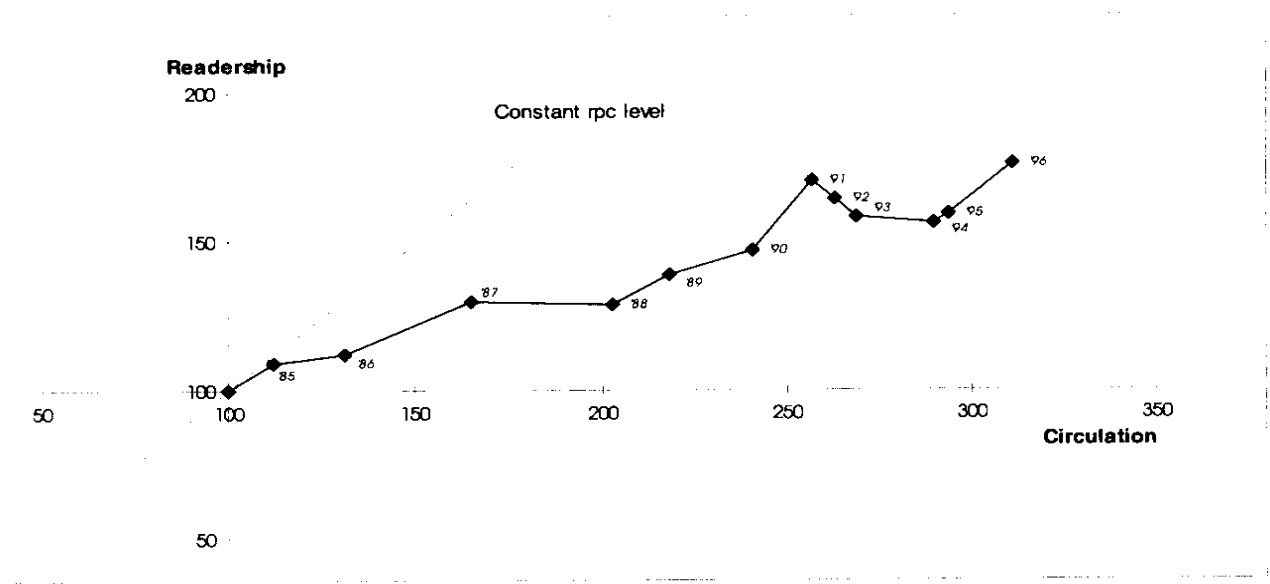
Readership was first nicely increasing up till 1986, while circulation remained stable. Between 1986 and 1987 readership was unchanged, while circulation increased. Between 1987 and 1990, readership decreased by some 20%, while circulation remained virtually unchanged (and rpc fell). From 1990 both readership and circulation have fallen sharply, basically at the same rate, thus keeping the level of rpc stable. Between 1995 and 1996 circulation recovered somewhat.

Why has readership and circulation developed like this? Economic recession? Need for editorial renewal? Are competing media getting stronger? Most importantly: where will it go from the present situation?

In this case we know, that radical editorial change took place in 1994/95. The image as somewhat of a slow 'business gossip' magazine was left for a concept of being a fast reporter of economy news. This has obviously had a positive effect on circulation in 1996, but not, as yet, on readership. So, where will it go from here? 'Up, up and away' or back to the previous downward trend?

Another example:

Morning business daily

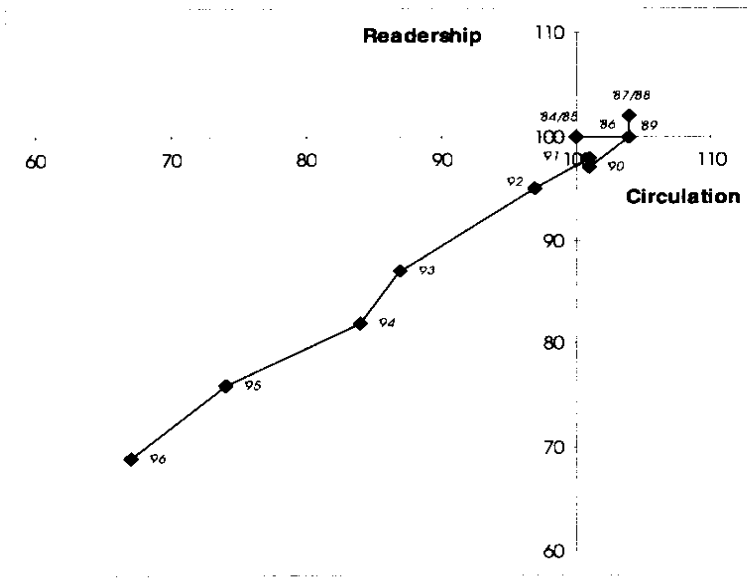


Graph 4. DYNAMAP™ of a morning business daily. Readership and circulation indexed: 1984 – 100.

This publication is a veritable success story, ever since its start in the mid 70's. The graph shows how it has grown over the years, with temporary minor losses in readership in the recession years 1991-94. Note the scale intervals!

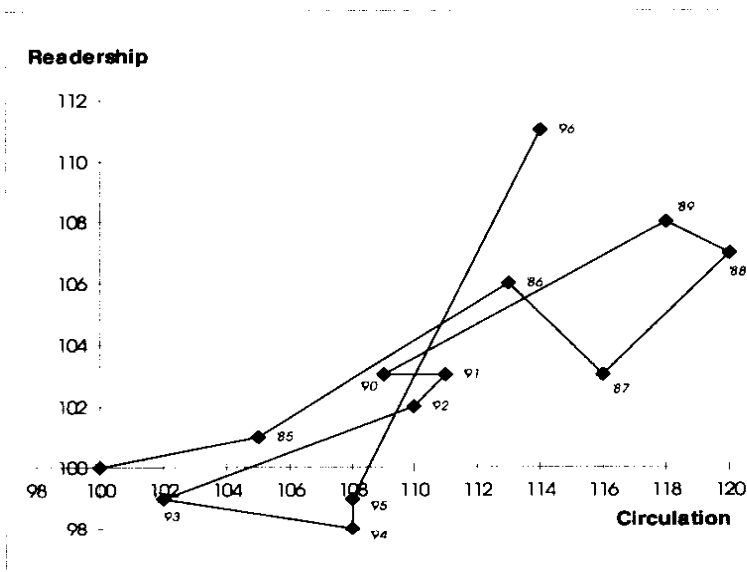
In terms of rpc it follows an expected pattern, in that rpc is going down continuously (by half in the last 10 years). This is the result of very consistent campaigns to make readers get their own subscription, delivered to their home in the morning, rather than having to get hold of a copy at work.

Evening daily 'A'



Graph 5. DYNAMAP™ of evening daily 'A'. Readership and circulation indexed: 1984 = 100.

Evening daily 'B'

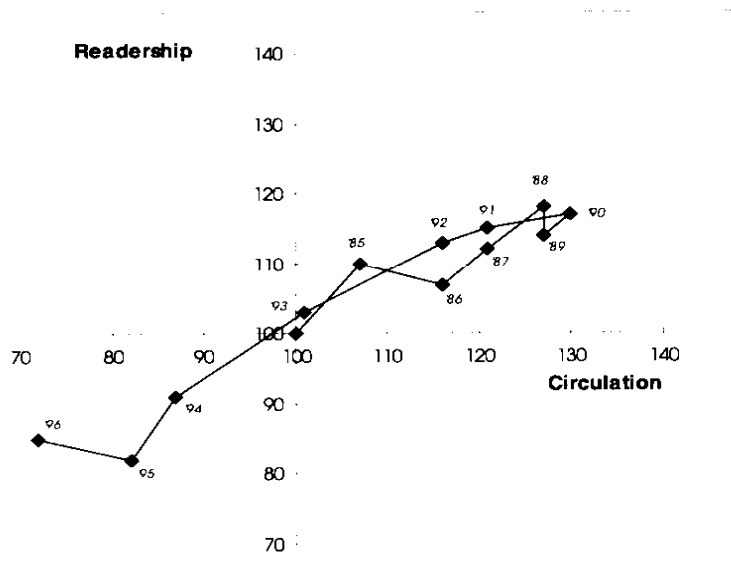


Graph 6. DYNAMAP™ of evening daily 'B'. Readership and circulation indexed: 1984 = 100.

The evening paper market was booming up till around 1988/89. 'A', being the larger of the two publications, had good years and maintained its position, while 'B' was striving for improvement. 'B' succeeded in increasing circulation quite well (taking advantage of the good times), while rpc was falling, in that readership did not increase at the same rate as circulation (well-known relation).

Then the situation for both started to deteriorate. Circulation fell by something like 20% in a few years (when the economic situation in general got worse), and readership a little less. Title 'A' has continued to fall, while 'B' initiated strong efforts for editorial renewal in 1992/93. These efforts paid off with increasing circulation in 1994, followed by a subsequent raise in readership. Finally, in 1996 'B' managed to become no 1 in circulation and in (a sky-rocketing) readership. Note that scales are different in the two cases. This example shows how seemingly similar publications may develop very differently, incl rpc. It is another clear indication to do analysis of this kind on a title by title basis, rather on a 'groups of similar titles' basis, which undoubtedly evens out any individual differences.

Cartoon



Graph 7. DYNAMAP™ of a cartoon. Readership and circulation indexed: 1984 = 100.

This publication also prospered until around 1990, with circulation and readership increasing at the same rate. Then it seems to have met with a dramatic turning event, causing both readership and circulation to fall back and beyond the starting point in 1984. It seems plausible to believe, that the reason for this rather dramatic change is to be found in the appearance of video- games, and that they have become a substitute for reading cartoons.

6) Continuous follow-up instead of forecasting

In private companies there are now tendencies to abandon the traditional financial budget work, in favour of more dynamic follow-up procedures. By analysing moving 12 months average figures, on a month by month basis, deviations from 'normal' are quickly identifiable. Closer analysis can be concentrated to where the problems are and subsequent actions be taken.

Maybe a similar philosophy can be applied to the analysis of readership and circulation. The real world does not seem to want to be modelled the way researchers (and others active in this area) would like it to be. Models are probably mostly poor in their ability to foresee the non-foreseeable. Perhaps it is therefore better to try and make frequent check-ups to see where the route goes - and why.

Compared to the 'conventional' graph 2 above, showing indexed readership and circulation changes, the Dynamap is a much more dramatic and illustrative record of the route a title takes over time. This pragmatic presentation technique allows us to give the publisher a 'motion picture' of what is happening. Only the publisher has the key to why, and the position and motivation to try and come up with a remedy.

The researcher/consultant has a big role and sometimes a tough job to do here, to make the publisher's editorial people understand and accept the Dynamap picture, as a true reflection of the real world, rather than just offending those who feel responsible for the way things are. In order to understand the situation it is also necessary to thoroughly analyse and compare with competitors' developments.

In close co-operation with the publisher it is in most cases also possible to set up internal short-term, non-official, circulation charts, which can be matched with short-term readership developments, say, three times per year. This is a simple but very useful 'early-warning system', the importance of which can be dramatic.

Consistent patterns by low margin of error?

Other authors have often mentioned the margin of sampling error of readership figures as a problem in regression analysis. Looking at the predominantly stable directions the data show in most of our 48 Dynamap graphs, it may be worthwhile to point out here:

Using the frequency scale for assessing AIR gives approximately only half the margin of sampling error, as compared to the case of Recent Reading.

This is of course because of the far greater number of interviews used in the calculation of the AIR values (the margin or error naturally has to be calculated individually for every publication, since it depends upon the variance within each scale position).

A low sampling error should help in giving an even safer Dynamap™ graphs of the readership/circulation relationship development.

Publications and their characteristics: order by R(2) values for C(t)/R(t)

Table 2.

No.	Title	Characteristics	Iss/yr	R(2) values:	R(2) values:	R(2) values:
				C(t)/R(t-1/2)	C(t)/R(t)	C(t)/R(t+1/2)
17	Expressen	Evening daily	300	.97	.98	.98
27	Knasen	Cartoons	26	.97	.94	.93
44	Vi	Family weekly	39	.90	.94	.95
26	Kalle Anka & Co	Donald Duck	52	.86	.94	.88
32	Må Bra	Health care mag	10	.89	.93	.95
11	Bilspport	Car sports mag	24	.86	.93	.93
45	Vi Bilägare	Drivers mag	20	.94	.92	.93
14	Dagens Industri	'Financial times'	300	.93	.92	.95
33	91:an Karlsson	Cartoons	26	.92	.91	.86
29	Land	Family weekly	47	.91	.91	.89
21	Hemmets Journal	Family weekly	50	.86	.90	.88
31	Motor	Drivers' org. mag	10	.90	.89	.89
25	Jaktm. & Fiskev.	Fishing&Hunting	11	.83	.89	.82
46	Vi Föräldrar	Child care	12	.92	.88	.91
5	Allas Veckotidning	Womens weekly	52	.91	.86	.78
24	ICA-kuriren	Family weekly	46	.86	.83	.83
48	Året Runt	Family weekly	52	.83	.83	.79
42	Veckans Affärer	Business weekly	41	.85	.82	.75
38	Svensk Damtidning	Womens weekly	52	.79	.76	.73
43	Vecko Revyn	Womens weekly	52	.80	.74	.60
47	Vår bostad	Condominion owners	11	.68	.72	.71
37	Svenska Dagbladet	Morning daily	300	.78	.71	.66
8	Allt om mat & vin	Food & wine mag	16	.75	.69	.33
22	Hemmets Veckot.	Family weekly	52	.71	.68	.61
19	Göteborgs-Posten	Morning daily	300	.72	.60	.64
15	Dagens Nyheter	Morning daily	300	.70	.59	.56
23	Hälsa	Health care mag	11	.61	.59	.58
9	Allt om MC	Motor bikes mag	6	.49	.59	.40
3	Aftonbladet	Evening daily	300	.44	.58	.40
40	Teknikens Värld	Car news/sports	26	.66	.53	.43
41	Trailer	Truck sports	11	.44	.52	.45
18	Femina	Womens mag	12	.64	.49	.25
7	Jakt & Vapen	Arms & Hunting	11	.42	.45	.58
4	Aktiespararen	Shareholders mag	11	.30	.43	.43
10	Antik & Auktion	Antiques mag	11	.41	.39	.46
13	Damernas Värld	Womens mag	12	.28	.34	.34
20	Hundsport	Dog sports mag	10	.66	.32	.67
30	Mitt Livs Novell	Young girls mag	25	.53	.32	.46
2	Affärsvärlden	Business weekly	42	.35	.19	.05
6	Allers familjejournal	Family weekly	52	.22	.19	.15
12	Båtnytt	Motor boats/yachts	12	.12	.15	.25
1	4 Wheel Drive	Car sports mag	6	.08	.14	.09
16	Det Bästa	Reader's Digest	12	.09	.08	.24
39	Svensk Jakt	Hunters org. mag	11	.05	.04	.01
36	På Kryss & Till Rors	Yachting	10	.08	.01	.01
28	Kommunalarb.	Municipal workers org	22	.02	.01	.00
35	Privata Affärer	Private economy	10	.00	.01	.04
34	Ny Teknik	New technology	43	.00	.00	.01

C(t)/R(t-1/2) = Circulation figures for calendar year vs Readership figures from the 12 months period starting 6 months before the calendar year

C(t)/R(t) = Circulation figures for calendar year vs Readership figures from the same calendar year.

C(t)/R(t+1/2) = Circulation figures for calendar year vs Readership figures from the 12 months period covering the 6 last months of the calendar year and the 6 first months of the next calendar year.

Period: 1984 -1996

