RECENT READING – AN EXERCISE IN DOUBLE-COUNTING ?

Ron Carpenter, Consultant to Trinity Mirror With addendum by Brian Allt

1. Introduction

- 1.1 Newspaper and magazine 'Average Issue Readership' (AIR) measurements in Great Britain have continuously used the 'Recent Reading' technique for some fifty years or more. A likely explanation for such a remarkably long and infrequently questioned sequence is partly due to the fact that Great Britain has always had a very large number of national newspapers and consumer magazines requiring measurement. Thus, for a **joint** newspaper and magazine survey, 'Through-the-Book' measures have always been considered impractical in terms of sample size and cost. Additionally, not all users have understood the definition of 'Readership'; and users are anyway generally resistant to change in the 'trading currency'.
- 1.2 It can also be hypothesized that many years ago, among some British researchers, there was an unwritten and even unspoken mutual acceptance that, with the 'Recent Reading' method, 'Replicated' and 'Telescoping' overclaims did in fact exceed 'Parallel Readership' underclaims, **but**, that this was then possibly counterbalanced by the very nature of the 'Recent Reading' method itself which did not take into account repeat reading of the same issue. Today the press media sales (and thus research) scene is more competitive and traditional practices are being reviewed hence the title of this paper.
- 1.3 Finally, throughout the remainder of this paper, it is worth noting that whether one is discussing 'Model-Bias', 'Replication' and 'Parallel Readership', 'Telescoping' or 'Double-Counting', then one is doing so at the **individual respondent level.** The final total sample readership levels will, of course, aggregate the possible **net effects** of these factors, which may vary by publication group and, almost by definition, will most certainly vary between highly regular readers of a specific title and less frequent readers of that title.

2. Content

The more detailed discussion and analyses on which this paper is based now cover the following topics:

Section Three	A reminder of the definitions of AIR and the 'Recent Reading' model.
Section Four	'Model-Bias' and 'Telescoping' – what are they?
Section Five	Contains the key analyses in this paper and summarises four quite separate approaches to the subject of Model-bias and Telescoping.
Section Six	The application of 'Reading Day' type measures to 'Recent Reading' AIR
Section Seven	Summary
Section Eight	Conclusions
Section Nine	Addendum by Brian Allt
Section Ten	References
Section Eleven	Appendices

3. Average Issue Readership and the Recent Reading Model

3.1 **Average Issue Readership (AIR)** is the universally accepted definition of 'readership' and survey estimates should represent "how many people read or look at an average issue of a publication at least once". One could discuss whether or not this definition is really what we want and should be attempting to measure - and I suspect that other Symposium papers may well tackle this more macro issue.

The purpose of my own contribution (assuming that we will continue to attempt to measure AIR as currently defined) is to examine the most commonly adopted technique for trying to achieve this objective, namely the 'Recent Reading' method.

3.2 **'Recent Reading'** is just one of several ways of attempting to measure AIR and other approaches include 'First Read Yesterday' (FRY), FRIPI (First Read in Publishing Interval), 'Specific Issue Readership' and 'Through the Book'. All of these methods have theoretical and practical advantages and disadvantages.

Relevant to Great Britain, the 'Recent Reading' technique does have some significant advantages in that: it can incorporate a large number of publications (particularly when using an 'Extended Media List/EML' multiple titles per prompt card approach); it does not overstrain respondents' memories in terms of 'first reading occasions' and it does not require as large a sample size as some of the other methods. However, the 'Recent Reading' approach assumes that AIR is **represented** by the model (or the answers to the paraphrased question): "when did you last read or look at <u>any</u> issue of Publication X?" – and from which, one **then** defines a reader as "anyone who has read or looked at **any** issue of Publication X within its publication frequency period"

Unfortunately, this Recent Reading 'representation' of AIR does suffer from both Model-Bias underclaims and overclaims which, when aggregated across the total sample, will only produce the desired AIR objective when these under and overclaims balance each other out. Like some of the other alternative techniques, the 'Recent Reading' method also suffers from respondents' memory imperfections.

3.3 The purpose of this paper is, therefore, firstly to discuss these under and overclaiming factors and then to examine their **net** effect on the readership estimates produced. In other words, do these 'Recent Reading' AIR estimates appear to make sense and do the under and overclaiming factors tend to balance out or not?

4. 'Model bias' and 'Telescoping'

4.1 The 'Recent Reading' model, together with respondents' memory factors can produce AIR over-claim and under-claim effects of the following types:

	OVERCLAIMS	UNDERCLAIMS
Model Bias:	Replicated Re Readership	Parallel Readership
Respondent error:		
Memory and Personality factors	Telescoping	Reverse Telescoping
	Bravado/Status Overclaims	Forgetting
	Title Confusion	Title Confusion

Looking at aggregated data across a group of similar publications, then 'Title Confusion' should balance out. It can also be assumed that the 'Bravado/Status' factor of trying to impress the interviewer is relatively small and can be ignored. In any case, it can hardly be allowed for in a scientific way.

'Forgetting' is self-explanatory and is more likely to occur with more fleeting (and thus less important) reading occasions. Attempts are made to reduce this factor to a minimum – with the use of masthead prompt-aids and a 'Read in the Last Year' (RPY) filter. Nevertheless, there is some evidence that, particularly in a waiting room, doctor's surgery and other similar out-of-home reading situations 'forgetting' is relatively more likely to occur and correction factors have been built in to a relevant analysis in Section 5.2.3

'Reverse Telescoping', namely 'thinking that a reading event (or any event) was **longer ago** than it really was', is thought to be of minimal effect. Indeed, the only example that I can think of (and which some of you may recognise) is about one and a half weeks after I have returned to work from a holiday – and then, it seems about a month ago !

This leaves the important 'Model-Bias' effects of 'Replicated' and 'Parallel' readership, together with the 'Telescoping' memory factor. These three require further discussion before proceeding to the readership analysis findings.

4.2 **'Replicated Readership'** is thought to be a principal component of double-counting in the 'Recent Reading' AIR estimate. It very simply and demonstrably exists when, for example, a respondent reads a specific issue of a weekly publication and then **re-reads** it exactly a week later (and, without re-reading **another issue** of the same publication within the period of the first reading event and seven days after the second reading event). This respondent effectively has double the chance of being included in the sample and thus will contribute to 'double-counting' in the final aggregated AIR estimate. Unfortunately however, authors on this particular subject sometimes fail to point out the possibly more common effect of 'Partial Replication' or 'Partial Double-Counting' as Brian Allt (reference: Ref.1) clearly described in a paper distributed at the 1995 Symposium. In similar circumstances to the above, but where the respondent (say) buys and reads a weekly magazine on a Wednesday and then reads it again (say four days later on a Sunday), it is perfectly obvious that this respondent is also a 'replicated' individual, but with a **partial** double-counting factor of 1.57 (ie one and four sevenths). Significant overclaims are thought to arise from both replicated and partially replicated readership within the 'Recent Reading' models.

Note: Replicated readership overclaims are far more likely to occur among irregular readers of a publication than highly regular readers. For example, replication overclaiming potential would be zero for a person who, over the year, buys and reads all 52 issues of a weekly publication on its day of issue.

4.3 'Telescoping' takes place when an actual event took place longer ago than a respondent thinks it did. I am sure that all readers and delegates aged 45 and over will most certainly immediately recognise what I mean and, from personal experience, I can assure everyone that this phenomenon worsens and worsens as one gets older! In the 'Recent Reading' interview situation for example, I wonder how many respondents 'think' that they last read a particular monthly magazine about (say) three weeks ago when, in reality, it was perhaps over four, five or even six or eight or more weeks ago? Considerable potential for double-counting (or probably more frequently <u>partial</u> double-counting) arises in such circumstances – particularly since most reading occasions are **not** highly important events in a respondent's life. If, for example, the actual last reading event for a monthly magazine took place six weeks ago, but a respondent 'claims' that it was within the last four weeks, then this respondent would contribute an over-claiming factor of 1.5 to the sample's final aggregated total readership figure.

This is because this respondent would presumably, "claim to have read in the last 4 weeks" at any point from one day after the actual reading event to **six** weeks later. This respondent's chance of being both sampled and being a 'Recent Reading' AIR reader is, therefore, one and a half times that of a respondent with a perfect memory – who would only have four qualifying weeks in which to be both sampled and make a qualifying readership claim.

A similar partial double-counting situation can arise for a weekly publication where, for example, the recency of a reading event is collapsed from (say) nine or ten days ago to 'read in the last week'.

Note: As with replicated readership, overclaims due to telescoping are far more likely to occur among irregular readers than among highly regular readers.

- 4.4 **'Parallel Readership'** leads to **underclaims** of AIR within the 'Recent Reading' model. 'Parallel Readership' occurs when a respondent claims 'to have read any issue of the specified publication within its publication frequency period', and is counted only once when, in fact, this respondent read two or more issues of this specific publication within its latest publication interval.
- 4.5 Within the current National Readership Survey Recent Reading method, it is not possible to isolate the effects of 'Replication' and 'Parallel Readership'. Other specifically designed surveys or experimental studies are required for this purpose. However, the **net effect** of these 'Model-Bias' and 'Telescoping' factors can be examined by comparing the AIR estimates produced by the 'Recent Reading' method with other data such as ABC audited circulation sales figures, FRY (First Read Yesterday) AIR results and Reading Days data. Such comparisons are shown in Sections 5.2 to 5.4 that follow, thus allowing the reader to make his or her own judgement on whether the 'Recent Reading Model-Bias' and 'Telescoping' factors balance each other out, or not.

5. Investigations into the net effects of recent Reading Model Bias and Telescoping Factors

Four separate ways of studying this subject follow in Sections 5.1, 5.2, 5.3 and 5.4

5.1 **Data in published reports**

It is very easy to look at published reports, choose to look at very specialist monthly magazines and then come to the conclusion that the published figures do not make sense. Apart from the following such example below (5.1.1), the remainder of Section 5 will concentrate on **major** and broader interest women's weekly and monthly titles. There are several main reasons for this:

Firstly, in one of the approaches (Section 5.2) one wishes to compare 'Recent Reading' AIR Estimates with audited (ABC) circulation sales figures. This can be done with more precision for women's titles since, on average in Great Britain, there is roughly only one adult female per household and, of more relevance, an average of only approximately 1.15 women if one excludes 'male-only' households.

Secondly, one does not wish to analyse particularly small samples and even with large magazines, FRY (First Read Yesterday) sample sizes can be pitifully small per title. Section 5.3)

Thirdly, one wishes to incorporate research other than the NRS, which attempts to measure Reading Days (Section 5.4).

Fourthly, limiting the analyses in Sections 5.2 to 5.4 to some 30 major titles, dramatically reduced the workload – as in addition to newspapers and some of their Sections, the NRS currently attempts to measure some 230 consumer magazine titles !

5.1.1 Reduction Ad Absurdum – A Specialist Monthly Magazine (an extreme example)

Magazine X: Publication interval: Monthly Readers per copy: 19.3 (Source NRS Jan-Dec 1997) Number of Reading Days (in issue period): 4.6 (Source: QRS 1998)

Translated, these figures mean that an average issue of this publication accumulates 89 gross reading days. This figure may not sound quite so astonishing, until one realises that this is equivalent to each issue having 3 readers per copy **each day** over 30 days ! Perhaps these readers read it in groups, rather like poetry readings, so that they still have time left to pursue their specialist hobby.

To compound the problem, 52% of these 19.3 readers per copy claim to be 'Primary' (household copy) readers, giving at face value, **10* 'Primary household readers per household copy****!

Note:* Michael Brown (ref 3) has demonstrated that the NRS Recent Reading model necessarily produces overproportions of 'Primary' readers and Brian Allt (ref 1), shows how the figure should be reduced by about 15 to 20% to 'adjust' the data from "individuals" to "households". However, it hardly seems worth bothering to adjust for the degree of this over-estimation above, since it would not make the figures any more believable. Note:** The average adult household size in Great Britain is very roughly 2.0.

5.1.2 A Major Women's Weekly and A Major Women's Monthly Magazine

	WOMAN'S OWN (weekly)	IDEAL HOME (monthly)
Women Readers (GB 000's)	2,968	1,266
Primary (household copy) readers (%)	54%*	53%*
Primary Women Readers (000's)	1,603**	671**
ABC (UK) Circulation	741	204
Estimated Primary Readers per copy	2.2	3.3
		Sources: NRS (GB) and ABC (UK)

<u>BUT</u>, THERE ARE ONLY APPROXIMATELY 1.2 WOMEN PER HOUSEHOLD IN 'HOUSEHOLDS CONTAINING WOMEN' IN GREAT BRITAIN ! (and over 85% of these households contain only one adult female)

Notes:

1)

* Without allowing for a Michael Brown adjustment factor (see Section 5.2, Ref 3).

- 2) ** Not allowing for oversampling of larger households and assuming that all of the women living in a primary household read the title.
- **Comment:** At face value, even assuming that **all** of the women in primary households read the publication, it is clear that we finish up with 'Primary Readers Per Copy' indices roughly two to three times what is logically possible. So what causes these inconsistencies ? Certainly, no one has ever heard of a publisher **understating** his annual sales ! Perhaps, therefore, the Recent Reading method together with the Primary readership claims, need to be investigated further.

5.1.3 Reading Frequency Probabilities

The NRS reading frequency scale is the same for all publication groups regardless of their publishing frequency. The theoretical and approximate observed probabilities for three of these publication groups are shown below:

FREQUENCY CLAIM	THEORETICAL PROBABILITY	SUNDAY NEWSPAPERS	WOMEN'S WEEKLY MAGAZINES*	WOMEN'S MONTHLY MAGAZINES*
Almost Always (3 or 4 out of 4)	Over 0.875	0.95	0.90	0.95
Quite Often (1 or 2 out of 4)	0.375	0.40	0.35	0.75/0.80**
Only Occasionally (less than 1 out of 4)	Significantly below 0.25 (say) 0.125	0.10	0.10	0.35**

* Indicates, Base = Women

** Indicates inexplicable results

Comment: Why should the results for less regular readers of women's monthly magazines be so inconsistent and massively above 'theoretical' expectation ? (a similar phenomenon arises for general monthly magazines among All Adults)

Could it be that respondents suddenly cannot grasp the concept of the frequency scale **in the middle*** of the media section of the NRS interview and in relation to monthly magazines ? Or is it that Replication and Telescoping are far more likely to occur for monthly magazines. Much more likely is the Replication and Telescoping hypothesis, because after all, many of these magazines may be kept for reference purposes. Interestingly, monthly magazines have proportionately, a much higher profile of infrequent readers than do daily and weekly publication groups. If, due to Replication and Telescoping, these probabilities for the 'Quite Often' and 'Only Occasional' readers are much too high, then their effects on AIR overclaims will be relatively even more pronounced.

Note:* In the British NRS, magazines are always sandwiched between the Daily and Sunday newspapers.

5.2 Comparing "Estimates of circulation sales from readership research" with "ABC-audited figures"

- 5.2.1 At the 1993 Symposium, Neil Shepherd-Smith tackled this exercise (Ref 1). As he noted at the time, his estimates of net AIR overclaims were necessarily **understated** since he had to assume that **all adults** in a Primary household were readers. Brian Allt's paper circulated at the 1995 Symposium (Ref 2), later pointed out that Shepherd-Smith's overclaiming estimates should have been even higher because the average sizes of 'Primary Reading Households' were , in fact, smaller than had been used in the calculations. Representing the NRS, Michael Brown had entered the discussion (Ref 3) by attempting to demolish Shepherd-Smith's argument both theoretically and by using the **same** publication as Shepherd-Smith used as an example. Unfortunately for Shepherd-Smith and fortunately for Michael Brown:
 - a) Neil Shepherd-Smith had to assume that **all adults** in a "Primary reader's household" were readers (which made it more difficult to prove that AIR overclaiming existed), and
 - *b)* Shepherd-Smith chose to use the Radio Times as an example to demonstrate his point. *Note: for a magazine, this publication has an exceptionally high proportion of Primary readers and, as we shall see later, it appears to have relatively little or no Model-Bias and Telescoping overclaims based on a FRY analysis (Section 5.3).*
- 5.2.2 In an attempt to continue the overall discussion about the Recent Reading model and Telescoping, together with an attempt to compare circulation estimates derived from readership figures against audited ABC sales, for reasons described earlier in Section 5.1.1, I shall now confine my attentions to 31 major **women's** weekly and **women's** monthly magazines. The particular selection of titles was governed by: omitting **very** specialist titles; omitting very young women's magazines (where many readers or primary readers were likely to be aged under 15); and by omitting titles which did **not** achieve a Recent Reading AIR sample size of approximately 100 or more in the QRS 1998 study. The full list of titles is shown in the Appendix. Summary results (for the remainder of Section 5) are shown for women's weekly magazines **in total** and monthly magazines **in total** together with the monthly magazines split by QRS defined sub totals. The results for each publication separately are shown in the Appendix (Section 11.3). However, these figures should be treated with caution since, even on a large sample, seemingly large differences between individual titles may be within sampling variation tolerances. (*Note: incidentally, and as a small though important digression, I note that another Symposium session may be devoted to the subject of "why readership and*"

Source: NRS (Jan-Dec 1997)

circulation <u>trends</u> do not always agree". From my own experience, and probably the research contractor's experience too (when answering queries), I suspect that in the vast majority of cases such differences can be explained by sampling variation. Further, one must always remember that even if there are **no trends whatsoever** in the circulation data, **by definition**, one in twenty of the hundreds or even thousands of comparisons we can make, will still 'appear' to be statistically significantly different at the 95% confidence limits.

5.2.3 Minimum Estimates derived from a Recent Reading Readership Survey (Method A) – incorporating 'Primary' readership claims.

Before summarising the results for the selection of 31 magazines, let us work through just one magazine as an example of the Method A calculation process:

		Woman's Own (Weekly)
		(000's)
a)	Total Women AIR readership estimate	2968
b)	Women Readers claiming 'Primary' (i.e. Household Copy) readership (54%)	1603
c)	Estimated number of 'Primary Households'	1344 (#1)
d)	Estimated number of 'Primary Households' adjusted downwards by the Michael	1021 (#2)
	Brown factor	
e)	Adjustment of the above to include 'Male only' households - thus giving	1027 (#3)
	Method A's circulation estimate derived solely from the NRS	
f)	ABC audited CIRCULATION over same period	689 (#4)
g)	OVERSTATEMENT INDEX = NRS derived circulation estimate over ABC	149

Source: NRS and ABC (January - December 1997)

Notes: (#1) This estimate of 1,344,000 is derived by using Brian Allt's method of calculating **average** 'Primary' **Household** Size' separately for each title (Ref 2). Brian Allt pointed out that claimed 'Primary' readers from larger size households are more likely to be interviewed in the NRS sample than claimed 'Primary' readers from smaller size households (because the NRS is a sample of individuals rather than a sample of households). The circulation estimates incorporating the Brian Allt method are thus based on more realistic and lower estimates of the average number of women in a 'Primary Copy household' compared with Neil Shepherd-Smith's earlier work (and, compared with the same Radio Times example in Michael Brown's Ref.3 reply) – and thus Method A gives higher 'derived' circulation estimates than Neil Shepherd-Smith. We should also remember that even with these larger Method A estimates, the figures are still based on the assumption that <u>all</u> women in a Primary Household* read the publication - thus, the readership survey derived circulation estimate is <u>still</u> a minimum or understated figure. Note:* the average number of adult women 'per household' reading, is more realistic than previous work based on 'on **all-adults** in the Primary household' having read the publication .

(#2) In his robust rebuttal of Neil Shepherd-Smith's initial work, Michael Brown (Ref 3) demonstrated that by using the Recent Reading technique, the Primary readership percentage (as published in the main NRS report) is likely to be significantly overstated. Applying Michael Brown's suggested correction factors to both readers of claimed 'Households Copies' and another up-grading factor for possible underclaims of readers of claimed 'Non-household' copies (title by title) produces, in the case of Woman's Own, a reduction in the estimated number of 'Primary Households' from 1,344,000 to 1,021,000.

(#3) Not surprisingly, this 'male-only households' adjustment has negligible effect in the case of Woman's Own.. Such an addition can however add a small amount (2% to 3%) to the circulation estimate for some of the more 'Home' oriented women's monthly magazines.

(#4) Unlike the NRS, the ABC audited circulation figures include sales in Northern Ireland. However, this amounts to roughly only 1.5% of sale for the average magazine. The ABC figures do also include unknown though probably small percentages of copies to places other than private households. If such corrections were to be made, this would have the effect of <u>increasing</u> the overstatement indices derived above and in sections 5.2.4.

5.2.4 Method B – which uses 'Primary Reader' Claims on whether the publication was for "Self only" or "Both self and someone else".

Method B is **not** independent of Method A since it starts off once again with Primary Readership claims. However it does **not** assume that **all** women in a 'Primary Household' read the publication - as it is based on 'Self/Both' claims cross analysed by the number of women in the household in order to derive approximate 'Household Copy' estimates.

Note 1: As with Method A, Method B will produce only approximate circulation estimates for each title. However one must bear in mid that we are only looking for <u>broad</u> estimates of overstatement (bearing in mind the sampling variation inherent in a readership survey together with the 'rounding up or down' that occurs in some of the percentages and figures) and thus the estimated Overstatement Index of 149 for Woman's Own previously should only be interpreted as "very roughly, about one and a half times as large"

Note 2: Since the claim "both" could also include another person <u>not</u> in the respondent's household, some allowance has been made for this situation.

5.2.5 Summary of Circulation Estimate Comparisons for 31 Women's Weekly and Monthly Magazines (Methods A and B)

	WOMEN'S MAGAZINES		WOMEN'S MONTHLY SUB GROUPS			PS
	WEEKLY	MONTHLY	'LIFESTYLE'	'GENERAL'	'HOME'	'CUSTOMER
	MAGS.	MAGS.	MAGS.	MAGS.	MAGS.	MAGS.
	(10)	(21)	(7)	(7)	(6)	(1)
ABC CIRCULATION 000'S *	6,026	5,676	1,909	2,370	1,005	392
WOMEN AIR READERS	20,228	23,390	7,517	8,666	5,056	2,151
MINIMUM CIRCULATION						
HOME COPY ESTIMATES**						
BEFORE MB ADJUSTMENT:						
METHOD A (000'S)	9,998	11,055	2,939	4,485	2,396	1,235
METHOD B (000'S)	10,484	12,053	3,563	4,721	2,475	1,294
AFTER MB ADJUSTMENT						
METHOD A (000'S)	7,780	8,555	2,241	3,510	1,803	1,000
METHOD B (000'S)	8,160	9,326	2,720	3,694	1,864	1,048
ESTIMATED						
OVERSTATEMENT						
INDICES						
METHOD A	129	151	117	148	179	255
METHOD B	135	164	142	156	185	267

Source: NRS and ABC (January – December 1997)

Footnotes:

ABC Figures plus estimates for two non-ABC publications

** Including a small number of 'male-only' households

MB = Michael Brown

The above analysis indicates that the Recent Reading method produces readership overclaims of very approximately:

- one-third for Women's Weekly magazines, and

- one-half or more for Women's Monthly magazines, with considerable variation between the different types of monthlies

Title by title comparisons are shown in the Appendix, Section 11.3

5.3 FRY (Total Issues) First Read Yesterday

5.3.1 It is generally thought that if one wishes to estimate AIR, then the most accurate way is to collect FRY data and then multiply by "the number of days in the publication interval" (i.e. 7 or 30.4) for weekly and monthly magazines respectively. FRY Data was collected on the NRS until mid-1990 for **experimental** purposes. However, to the best of my knowledge, the data was never published or even written-up in a confidential report. Confidential tabulations were, I understand, seen in the mid-1980's by a very few senior GB researchers. I infer that Michael Brown makes use of some of this information (Ref 3) in his reply to Neil Shepherd-Smith (Ref 1).

The barrier to accessing this past data has now been lifted by NRS Ltd. The mid-1980 tabulations referred to above covered a six-month interviewing period. The **title-by-title** differences when comparing the Recent Reading AIRs and the AIR estimates derived by FRY were incomprehensible and enormous – thus, no doubt, leading to the restriction placed on access to this data. However, the reason for these massive **title-by-title** differences was, I believe, very simply and very largely **sampling variation**. Since access to this past experimental data **is** now available, the following results in this paper show (in published form for the very first time, I believe) large sample NRS Recent Reading and FRY comparisons based on the same informants.

To increase sample sizes, a two-year NRS analysis period was selected - yielding a base of over 30,000 women!

Note: And even here, with such a large sample, the number of occasions where one found women claiming to read a specific major monthly magazine "for the first time yesterday" is still extremely small in many (random, I suspect) cases. The results of this analysis are shown below and, necessarily, only make sense if shown for groups of titles.

WOMEN'S WEEKLIES (10) WOMEN'S MONTHLIES (22)	RECENT READING AIR (000's) 20,925 22 482	FRY AIR (000's) 15,372 16 203	RECENT READING AIR OVERSTATEMENT INDEX 136 139
LIFESTYLE MAGAZINES (11)	8,284	5,563	149
GENERAL MAGAZINES (7)	11,279	8,938	126
HOME MAGAZINES (4)	2,044	912	224*
RADIO TIMES	4,752	4,669	102
TV TIMES	4,797	4,634	104
READER'S DIGEST	3,103	3,374	92

Source: NRS (July 1988 - June 1990); Base: Women

Notes:

- a) Necessarily, because of changes over the last ten years, there are minor changes to the monthly magazines selected (see Appendix).
- b) *Although one might expect the 'Home Interest' magazine index to (potentially) be more likely to include a replication effect, this index is only based on four titles thus, because of sampling variation, it should be treated with caution as, indeed, should the differences in the indices between the 'Lifestyle' and 'Women's General' magazine groups.

Comment: What is so interesting here is how **relatively close** these **Overstatement Indices** above are to the often maligned circulation comparison method reported on in Section 5.25 viz. :-

5.3.2

	RECENT READING AIR OVERSTATEMENT INDICES		
	versus FRY AIR	Versus PRIMARY HOUSEHOLD COPY (CIRCULATION) METHODS A & B	
WOMEN'S WEEKLIES	136	129-135	
WOMEN'S MONTHLIES	139	151-164	

Based on two totally independent approaches, the above table convincingly indicates that it is **highly unlikely** that Model-Bias under and overclaims, together with Telescoping, cancel each other out when estimating AIRs by the Recent Reading method. Aggregated within respondent double and partial double counting, appears to lead to overstated AIR estimates of roughly 33% and 50% for major women's weeklies and women's monthlies respectively.

5.4 Gross (Any Issue) Reading Days in Publishing Interval (GRDIPI)

5.4.1 I am indebted to Brian Allt who suggested that one might possibly try to disentangle Model-Bias and Telescoping effects, by looking at the concept of Gross Reading Days and comparing the internal consistency of estimating "GRDIPI" by two different methods from just one survey which uses both the Recent Reading technique and collects 'Reading Days' data. The QRS 1998 study fitted this requirement.

By comparing 'GRDIPI' estimates derived from an AIR component multiplied by 'Reading Days in Issue Period' with a similar estimate derived from 'Read Yesterday' data, and **then**, comparing any differences arising with the AIR Overstatement Indices found in Section 5.2.5 and 5.3, one can possibly try to isolate whether a Telescoping effect exists or not – since presumably, Telescoping is minimal or zero for Read Yesterday claims.

Note: a) One should be aware that if a Telescoping effect is found, then this will be Telescoping to a **frequency** question, which is not the same thing as Telescoping to a **'when last read'** question.

Note: b) Gross Reading Days has also been suggested in the past as an alternative measurement criterion to AIR e,g, Brian Allt (Ref 4). However, although such a measure might appeal to advertising end-users, it would not satisfy press media owners' internal Marketing Department requirements unless an AIR component could be extracted from it.

5.4.2 The QRS 1998 study adopted the NRS interview protocol for the AIR readership section of the interview and the QRS study similarly recorded whether 'Last Read' was 'Yesterday' or not.

Again, like the NRS, the QRS did not collect whether reading was 'First Time Yesterday'. However, in following questions, the QRS did attempt to measure "How many reading days there were in the issue period" and "How many different copies were read on the last reading day".

One can, therefore, compare two estimates of Gross Reading Days in the Publishing Interval (GRDIPI) as follows:

- Method X = (AIR Readers) x (Number of Reading Days) x (Number of Issues Read on Last Reading Day)
- Method Y = (Read Any Issue **Yesterday**) x (Number of Issues Read Yesterday) x (7 or 30.4 depending upon whether it is a weekly or monthly publication).

GRDIPI OVERSTATEMENT INDEX = 'METHOD X' over 'METHOD Y'

- □ All of the information for Method A is contained in the QRS 1998 main report
- □ I am also indebted to Zenith (Trinity Mirror's Media and Buying Agency) for supplying a special postsurvey analysis of QRS 'Read Any Issue Yesterday' data.
- Unfortunately, the QRS computer database did not allow analysis of the number of different issues read on 'the Last Reading Day' for a base of 'Read Any Issue Yesterday' informants. One, therefore, has to assume for the purposes of this exercise that it is the same as the figure for the base of 'Last Reading Day' informants.

Note: a) Fortunately, the 'Number of Copies Read on the Last Reading Day' is generally relatively small and is **not** a highly discriminating factor, b) fortunately too, because regular/replicated readers are relatively more likely to fall in the 'Read Yesterday' base and are likely to have lower 'Parallel Readerships' (ie, rather similar to the logic behind Michael Brown's adjustments for 'Primary Readership'), by making the assumption that the two 'Number of Issues Read per Day' estimates above are the same, then this possibly has the effect of marginally **underestimating** the **Overstatement** (but probably by not more than roughly 5%)

	Method X AIR x Reading Days 000's	Method Y Yesterday x 7/30	INDEX (X over Y)
Woman's Weekly Magazines (10)	42,658	17,781	240
Woman's Monthly Magazines (21)	65,668	25,996	253
- Lifestyle Magazine (7)	20,663	5,908	350
- General Magazines (7)	26,079	12,928	202
- Home Magazines (6)	13,911	4,816	289
- Customer Magazines (1)	5,014	2,344	214

5.4.3 Gross Reading Days in Publishing Interval (GRDIPI) Comparison

Note: All the "Reading Days in Issue Period" and "Number of different issues on the last reading day" figures were extracted from the QRS main report. These figures are based on Adult readers rather than women. However, as the 31 magazines have relatively low profiles of men readers – this should not affect the Overstatement Index very much. If Reading Days and 'Number of issues per day' data for women had been used, it is likely that the OverstatementI indices in the table above would have been slightly higher.

Comments:

- a) Indices for the 'Recent Reading AIR' Method X are much higher than for the 'Yesterday' Method Y.
- b) The fact that these Method X over Method Y indices are also much higher than the AIR Overstatement Indices found in sections 5.25 and 5.3, implies that the Reading Days component itself is likely to be overstated, and, overstated by a larger amount than the Recent Reading AIR overstatement factor. This would seem to indicate that Reading Days in Issue Period is a very difficult concept to measure, and, that the question may not be correctly understood by many respondents and/or that Telescoping exists and has a particularly powerful effect in relation to a 'frequency in issue period' type of question.
- c) The results also tempt one to speculate that perhaps the AIR overstatements demonstrated in Sections 5.25 and 5.3 earlier may perhaps be due solely to Telescoping and thus the replicated and parallel readership AIR Model Biases might, after all, cancel one another out?
- d)* Finally, since the differences in the Method X over Method Y indices above for the sub-groups of women's monthly magazines are so difficult to explain (particularly for the women's Lifestyle magazines), it was decided to take the QRS 'Number of different reading days in issue period' scores and apply them to two years NRS "Read any_issue Yesterday" data. The results are summarized below and go to show just how dramatically things can change once one reduces the amount of sampling variation.

5	4	ι.	4
\mathcal{I}	. 7	r.,	т.

(4,000)	(40,000)	
240	262	
253	279	
350	226	
202	299	
289	362	
	(4,000) 240 253 350 202 289	(4,000)(40,000)240262253279350226202299289362

Comment:

The high Women's Weeklies and Women's Monthlies indices are reasonably similar and the NRS data confirms concerns about substantial overclaiming in the 'Reading Days' measure - a priori, one would probably have anticipated Telescoping to be more significant for monthlies than weeklies. Because of sampling variation, the data also indicates how dangerous it would be to look at 'Yesterday' data title by title.

6. The application of Quantitative Reading Days Type Measures to Recent Reading AIR

The principal discriminating component of QRS-PEX and MPX-type measures is Reading Days. We have already seen in Sections 5.4.3 and 5.4.4 that it appears that informants' "Reading Days in Issue Period" claims (as normally measured by Method X) are highly inflated (most probably due to within-informant double-counting, or partial double-counting, arising from Telescoping).

The **fusion** or **multiplication** of such inflated claims to an already inflated Recent Reading AIR estimate therefore **compounds** the overstatement problem.

The most likely end result of such fusion or multiplication, if we think in terms of a final publication score or cost ranking value, is that we arrive at publication scores involving **treble or partial-treble counting**!

7. Summary

- 1. Recent Reading AIR estimates have been questioned for many years on theoretical Model-Bias and Telescoping grounds and because, on visual inspection, readers-per-copy figures often appear to be implausibly high. Examples quoted in the literature usually relate to the more Special-Interest types of monthly magazines.
- 2. This paper demonstrates that overstatement of Recent Reading AIR estimates for **broad-interest major** women's weekly and monthly magazines also occurs with the weeklies appearing to be overclaimed by a factor of about one third and the monthlies by about a half.
- 3. Circulation estimates derived from Recent Reading AIR figures adjusted by Primary (household) Readership and "who is the copy for" factors, significantly exceed audited ABC circulation figures (section 5.25).
- 4. **Recent Reading** AIR estimates similarly well exceed AIR estimates based on **FRY** (section 5.3) and these overclaims are very similar to the estimated overclaims when comparing Readership Survey based estimates of copy sales with audited ABC figures
- 5. It is hypothesised that this over-claiming arises from the aggregated net effect of double and partial double counting within informants due to Model-Bias together with Telescoping.
- 6. Reading Days in Issue Period appears to suffer from considerable overclaiming due to Telescoping. The Gross (Any Issue) Reading Days in Publishing Interval (GRDIPI) measure therefore, often becomes massively overclaimed when based on an inflated Recent Reading AIR estimate multiplied by an inflated Reading Days' claim.
- 7. The application (fusing or multiplication) of Recent Reading AIR estimates by **quantitative** measures based on an important 'Reading Days' component, (e.g. QRS-PEX or MPX) simply compounds the situation by multiplying together **two inflated estimates** (both of which contain double or partial-double counting.)
- 8. Recent Reading An Exercise in Double (or Partial Double)-Counting ? I rest my case.

8. Conclusions

- 1. Model-Bias and Telescoping significantly inflate Recent Reading AIR estimates for selected groups of major women's weekly and women's monthly magazines.
- 2. Assuming that in Great Britain we **do** wish to continue to measure AIR, **and**, that we wish to do so on a large and **joint** Newspaper plus Magazines survey, we need to decide "Where do we go from here?"
- 3. FRY data might seem to be an ideal theoretical solution. However, having analysed an extremely large sample (over 30,000 women from two years' NRS data) it is quite obvious that sampling variation (and thus cost) makes it totally out of the question to consider publishing FRY results for individual titles. Perhaps correction factors by publication group could be considered and though this might work for Cost Ranking analyses, much debate would surely arise on how to correct Coverage estimates. One suggests that the NRS might re-introduce the FRY question for experimental purposes (particularly as the NRS continues to collect Any Issue Read Yesterday data).

Note: FRY data is also not really viable for measuring Weekend Newspapers and their Sections – since a demographically balanced and massive loading would need to be given to the proportion of both Sunday and Monday interviews.

- 4. Through-the-Book methodology might be another ideal way of measuring AIR. However the logistics, fusion and cost of very many large sub-samples makes it impractical for a joint Newspaper and Magazine survey covering a large number of titles.
- 5. FRIPI (First Read in Publishing Interval) has been tried for example, in Holland and South Africa. Perhaps further investigation and refinement of this methodology relevant to Great Britain should be pursued. Costa Tchlaoussaglou describes experiences with the Dutch NRS in his 1997 Symposium paper (Ref. 7.)
- 6. Will an AIR measurement system based on 'Frequency of Reading' rather than 'When last Read' (an idea very recently floated by Michael Brown Ref. 6) provide a viable solution? It is still early days to judge. Perhaps such types of 'Frequency' approach will feature in the Symposium programme papers or discussion.
- 7. A **new** 'Readership' definition based on Gross Reading Days does not seem likely or 'acceptable' as not only does the measure itself appear to contain substantial over-claiming, but it would also involve introducing a completely new and different 'trading currency'.
- 8. So where do we go from here? We have in Great Britain an NRS Recent Reading AIR methodology which appears to give inflated AIR estimates for some groups of publications, BUT, we have lived quite happily with this for most of the last fifty years. Our Recent Reading AIR model does not contain any 'added value' for repeat reading of the same issue of a publication. However, with the exception of some atypical titles (e.g. the TV Programme Listings magazines and Reader's Digest), it is hypothesised that those publications with relatively greater AIR overclaims are those publications with a greater number of claimed (and inflated) of Reading Days. Perhaps there is a roughly equitable 'balancing out' after all ! Until we investigate and learn a lot more about our readership models (and repeat exposure advertising effectiveness), perhaps therefore we should continue with our 'cosy' and pragmatic joint Newspaper and Magazine Recent Reading double or partial-double counting system, providing that, we do not deceive ourselves into thinking that we achieve greater accuracy by introducing (i.e. multiplying by) inflated Reading Days factors. Such a move could lead to treble or partial-treble counting, might debase the 'trading currency' and possibly lead to 'trading anarchy'.

9. Addendum (by Brian Allt)

Since contributing to the first three Symposia in New Orleans, Montreal and Salzburg, I have continued to be concerned with readership research and the thinking exhibited at subsequent symposia.

During all that period I have tried to decide how to position them.

Are they, at one extreme, convocations of modern research scientists or, at the other, a free-masonry dedicated to the preservation of arcane and archaic beliefs and concepts?

Both when I used to be able to attend, and since, we often seem like a gathering of mediaeval alchemists. We pick up a lot of useful knowledge, but somehow never put it together in such a way as to make us realise that we should totally re-cast the definition of our quest.

Just as the alchemists were obsessed by the hope that they could turn base metal into gold, so readership measurement is still rooted in the quest for the grail of a 'true' direct measure of Average-Issue-Readership coupled, in recent years, with the design of ornamental bells and whistles to hang on it when caught. (I include first-reading as 'direct' in this sense).

In 1981 I floated the idea that we should think of our 'measurements' first as 'Press Negotiation Indices', and as 'survey research' only second - a reflection of the true status of readership data then as now.

A currency survives only if it is trusted. For decades, partly because it was easy for non-technicians to understand, and so long as no-one rocked the boat, the obvious non-comparability of AIR estimates for dailies, weeklies and monthlies was tacitly accepted and ignored.

Over the years, great efforts were put into devising ever more complex and expensive survey designs which it was hoped might get rid of the main problem - replicated readership. At the same time there was a steady flow of demonstrations of the large and differential effects which replication causes e.g. Neil Shepherd-Smith's recurring papers and my own distributed in Berlin, where I showed that he was probably considerably under-estimating the problem. Ron Carpenter has just underlined these problems.

The strange paradox is that, concurrently with the widening recognition of the likely inflation of magazine estimates, the world has been flooded with 'add-on', scores and weights (affectionately dubbed 'Quality' measures) which further inflate and compound the existing distortion.

Because, to some extent, space-buyers were confused into believing that these quantitative factors were being added on to help them plan and buy more efficiently, they do not seem to have adequately questioned them.

Perhaps on the 'add an egg' principle, they are so tickled with the extra numbers to play with, they mistakenly see such data as somehow corroborating the base metal which had, originally, j seemed to have been their burning concern! 'Quality' measures do no such thing, they simply magnify the inequalities! In my view, if we put together the great deal we know about all the different techniques applied to the measurement of reading behaviour over the decades, link that to the ever-increasing number of titles demanding measurement, and to the elaboration and cost of virtually all the proposed hoped-for solutions to the problems of differential estimation, a simple choice faces us, EITHER we settle for the long-standing cheap (!) and cheerful (?) recent reading model which we have used for ever, (**banning** the use of spurious 'quality' differential multipliers) accepting its inequalities and distortions, OR we seriously try to find out what people can honestly (i.e. with a minimum of survey-design-induced distortion) tell us about their reading behaviour that passes reasonable tests of:-

- a) Its potential as the basis of a new currency
- b) Its reliability
- c) Its credibility (people will answer any questions we put)

As long ago as 1975, in an attempt to encourage wider debate about all these problems, I published a paper in ADMAP called, perhaps unwisely, "To Confuse the Issue". For argument, I posited a survey where the basic unit of currency would be the 'Reading Day'. Frequency and AIR would be judgementally estimated via appropriate questions, simply to build alternative models of coverage and frequency recognising, of course, that our current frequency and AIR data are not 'fact', but simply particular by-products of the crude conceptual model we employ.

Is anyone out there listening yet?

".....eye of newt, toe of frog....."

10. References

Ref 1:	N. Shepherd-Smith	"Validating Average Issue Readership Levels by Circulation and Source of Copy Data". Readership Research Symposium 1993
Ref 2:	B. Allt	"Not Perfection, but Adequacy for the Purpose, or" Paper distributed at the 1995 Readership Research Symposium.
Ref 3:	M. Brown	"Lies, damned lies and replication" Admap April 1994
Ref 4:	B. Allt	"To Confuse the Issue" Admap, November 1975
Ref 5:	QRS	. The GB Quality of Reading Survey 1998. Sponsored by the PPA, IPA and ISBA.
Ref 6:	M. Brown	"Never mind when - tell me how often". EMRO Conference. May 1999
Ref 7:	C. Tchaoussoglou	"From FRY to FRIPI". 1997 Readership Research Symposium

11. Appendix

11.1

Michael Brown's combined 'PRIMARY' and 'OUT-OF-HOME' ADJUSTMENT FACTOR (Section 5.2 and Ref. 3)

NRS Report	M. Brown
Primary Readership (%)	Adjustment Factor
85	0.91
80	0.88
75	0.85
70	0.83
65	0.81
60	0.78
55	0.76
50	0.74
45	0.73
40	0.71
35	0.69
30	0.68
25	0.66

	The ABC Circulation Comparison (Section 5.2) and The Gross	The Recent Reading versus FRY Analysis (Section 5.3)				
	Reading Days Analysis					
	(Section 5.4) (1007)	(1088 1000)				
Waman's Waakligs:	(1997)	(1988-1990)				
Woman's Own	VES	VES				
Woman	VFS	VES				
Woman's Weekly	VES	VES				
Woman's Realm	VFS	VES				
My Weekly	VFS	VES				
Best	VFS	VES				
Bella	VFS	VES				
Chat	VES	VES				
Hello	VFS	VFS				
Take a Break	VES	NO				
People's Friend	NO	VES				
r copie s mend	NO	125				
Women's Monthlies:						
Woman and Home (G)	YES	YES				
Family Circle (G)	YES	YES				
Prima (G)	YES	YES				
Essentials (G)	YES	YES				
BBC Good Food (G)	YES	NO				
Good House Keeping (G)	YES	YES				
Ideal Home (G)	YES	YES				
She (L)	YES	YES				
Vogue (L)	YES	YES				
Marie Claire (L)	YES	YES				
New Woman (L)	YES	YES				
Cosmopolitan (L)	YES	YES				
Elle (L)	YES	YES				
Company (L)	YES	YES				
House Beautiful (H)	YES	NO				
Homes and Ideas (H)	YES	NO				
Homes and Gardens (H)	YES	YES				
House and Garden (H)	YES	YES				
Country Living (H)	YES	YES				
Country Homes and Interiors (H)	YES	YES				
Sainsbury's Magazine (C)	YES	NO				
Living (G)	NO	YES				
Annabel (L)	NO	YES				
Harpers and Queen (L)	NO	YES				
Woman's Journal (L)	NO	YES				
Options (L)	NO	YES				

11.2 The List of 31 Women's Magazine Titles Analysed In Section 5

G = Women's General

L = Lifestyle H = Home Interest

C = Customer Magazines

Note: The two lists are not identical as new magazines are launched and others merge and die.

11.3 (Appendix) Estimating MINIMUM Circulation levels from Readership data (Section 5.2)

	NRS Jan-Dec 1997 (Base: 23,743 women)							mbadj.fact H/Holds with	mbadj.fact Women 15+	mbadj.fact (Adults)	mbadj.fact (Adults)	OVERSTATEMENT INDEX	
		ABC Circn.	NRS Women	Primary women	Minimum H/Holds Method A	Minimum H/Holds Method B	Michael Brown adiustment	Minimum Circn. Method A	Minimum Circn. Method B	Minimum Circn. Method A	Minimum Circn. Method B	(Adults) MB adjstd	(Adults) MB adjstd
		(000's)	(000's)	(%)	(000's)	(000's)	factor	(000's)	(000's)	(000's)	(000's)	Method A	Method B
Womans Realm	ww	231	886	46	362	377	0.73	264	275	264	276	1.14	1.19
Womans Weekly	ww	565	1876	53	873	913	0.75	655	685	656	687	1.16	1.22
Woman	ww	721	2326	60	1179	1261	0.78	920	984	924	987	1.28	1.37
Best	ww	505	1806	58	893	938	0.77	688	722	690	724	1.37	1.43
Womans Own	ww	689	2968	54	1344	1431	0.76	1021	1088	1027	1094	1.49	1.59
Bella	ww	662*e	2442	57	1183	1235	0.77	911	951	912	952	1.38	1.44
Chat	ww	491	1472	64	795	839	0.81	644	680	649	684	1.32	1.39
Hello	ww	482	1766	46	678	703	0.73	495	513	503	521	1.04	1.08
Take a Break	ww	1310*e	3498	70	2086	2192	0.83	1731	1819	1754	1839	1.34	1.40
My Weekly	ww	370	1188	50	541	532	0.74	400	394	402	394	1.09	1.06
Total W.Weeklies (10)	Ww	6026	20228		9934	10421		7729	9 8110	7781	8158	1.29	1.35
Woman & Home	G	325	1294	47	544	573	0.73	397	418	399	421	1.23	1.30
Family Circle	G	290	956	61	518	552	0.78	404	431	407	434	1.40	1.50
Prima	G	504	1470	67	825	873	0.82	677	716	677	717	1.34	1.42
Essentials	Ğ	287	846	65	466	496	0.81	377	402	377	402	1.31	1.40
BBC Good Food	Ğ	327	1093	69	669	688	0.83	555	571	579	594	1.77	1.82
Good H/keeping	Ğ	428	1741	53	815	865	0.75	611	649	617	654	1 44	1 53
Ideal Home	G	209	1266	53	594	617	0.75	446	463	455	472	2.18	2.26
Sainsbury Mag	Cust	392	2151	64	1201	1256	0.81	973	1017	1000	1048	2 55	2.67
She	IF	230	907	46	327	362	0.73	239	264	239	264	1.04	1 15
Voque	L F	156	1290	39	359	450	0.71	255	320	265	327	1 70	2 10
Marie Claire	L F	398	1400	57	591	694	0.77	455	534	457	536	1 15	1.35
New Woman	L F	251	642	60	305	358	0.78	238	279	238	279	0.95	1 11
Cosmonolitan	LF	416	1707	57	705	859	0.77	543	661	549	666	1.32	1.60
Elle		186	885	49	296	375	0.74	210	278	223	280	1.02	1.00
Company		272	686	70	324	439	0.83	269	364	271	367	1.20	1.35
House Beautiful		294	896	66	533	559	0.81	432	453	444	467	1 51	1 59
Homes & Ideas	н	217	1128	53	523	553	0.75	392	415	404	428	1.86	1.00
Homes & Gardens	н	146	1120	44	441	449	0.73	322	328	344	340	2 36	2 30
Country Living	н	158	606	50	283	281	0.73	209	208	217	216	1 37	1 37
Country Homes/Int	н	83	110	46	183	101	0.74	13/	130	130	111	1.57	1.57
House & Garden		107	953	43	337	342	0.73	243	246	256	260	2 30	2 / 3
Total W.Monthly (21)		5676	23390	45	10839	11832	0.72	8389	9 9156	8557	9325	2.39	2.43 1.64
SubT-Lifestyle (7):	L	1909	7517		2907	3537	,	2217	7 2701	2747	2719	1 17	1.42
SubT- 'General' (7)	Ğ	2370	8666		4431	4664	L	3467	7 3649	3511	3694	1 48	1.56
SubT-Homes (6)	н	1005	5056		2300	2375		1732	2040	1804	1864	1 80	1.85
SubT- Gen+Hom(13)	 G+H	3375	13722		6731	7030	-	5100	5438	5315	5558	1 57	1 65
SubTCust	Cuet	392	2151		1201	1256		973	3 1017	1000	1048	2 55	2 67
*e = estimate.	0000					.200	•	010				2.00	2.07

Worldwide Readership Research Symposium 1999