

Roger Beeson
Lonsdale Advertising
London, UK

Dick Dodson
Foote, Cone & Belding
London, UK

4.1

UK DEVELOPMENT WORK ON DIARY PANELS

INTRODUCTION

JICNARS in the UK is constantly reviewing its techniques and looking for opportunities for improvement. 1984 was no exception and in the Autumn of that year we piloted three diary techniques. This paper describes some of the findings and conclusions from this investigation.

However, before that, it is relevant to relate some of the background which led to this experiment, in order to put it into perspective.

BACKGROUND - THE INITIAL STIMULUS

It seems appropriate to start at the New Orleans Symposium in February 1981. At that meeting the document by Pym Cornish and Michael Brown entitled 'Readership research reviewed' was discussed. It will probably be remembered that this paper was commissioned as a review of readership research problems and opportunities following the unsuccessful 'Cumberland Lodge Experiment' which looked at new techniques.

That summarised the potential for panels in two ways:

"All diary methods share a number of theoretical advantages compared with Recent Reading and TTB methods!"

"... that low response rate and high costs per informant would almost certainly rule out diary panels as an acceptable method for the NRS"

Later in 1981 JICNARS organised a seminar for all its subscribers in order to identify their future requirements. Several issues raised at

that seminar could, potentially, be solved by panels.

These and other events in 1981 led the advertising agencies (IPA) to put forward a proposal for a six-month test of a 500 person panel. The main reasons for the proposal were four areas of dissatisfaction:

- Accuracy of estimates for monthly magazines
- Validity of cumulative net coverage over multiple insertions in publications
- The total number of titles covered by the survey
- The speed of reporting of dynamic changes and new publications.

The IPA felt so strongly about the need for this test that it even volunteered to raise some of the necessary funding itself.

FIRST STAGE INVESTIGATIONS

This proposal was presented to JICNARS and subsequently discussed by the Technical Sub-Committee and the Development Working Party. However, at that time, (Spring and Summer 1982), various initiatives were under way.

In order to obtain additional input to their deliberations JICNARS invited comment from several research companies on the problems which were faced. Some of the resultant replies again addressed themselves to panel techniques.

At this point, the JICNARS development structure was changed to cope with the various areas of investigation. Separate 'Technical Study Groups' were established for each area; in particular a Diary Panel TSG.

This study group had the following brief:

- (1) To review and document previous diary panel experience, in the UK and elsewhere.
- (2) To consider the (previously mentioned) IPA proposal to JICNARS and to recommend on its acceptance or otherwise.
- (3) To consider whether a major test of a readership panel should be conducted to find a technique that could be applied to the NRS on an on-going basis, or conducted solely for the purpose of collecting additional data relevant to currently-encountered measurement problems.
- (4) If relevant, to prepare a proposal for a panel test.
- (5) Subject to acceptance, to commission and oversee the test and make recommendations based on its findings.

The review of diary experience was a lengthy task, which included discussion with various delegates at the Montreal Symposium. The written report was finally published in September 1983 (and is available from JICNARS).

Unfortunately the review did not shed much light on the potential diary problem areas we had defined. The main reason for this was that none of the studies reviewed had been designed to test the technique. We did, therefore, broaden our knowledge on the size and nature of the potential problems, but we did not discover whether these problems could be avoided. We were also unable to evaluate the size or complexity of these problems. The areas we examined are stated in Appendix I of this paper.

During the period of our review the TSG had also been discussing possible diary designs for testing. We had reached one major conclusion. This was that, if the technique was to be viewed as an alternative to the then current NRS technique, a fully write-in diary would be the only option. This view was based on two premises. Firstly, that a prompted diary covering some 200 titles would almost certainly create more title confusion and order effects than already existed. Secondly, that a split technique diary, prompting for some titles and not for others, would be unacceptable.

However, we knew that a write-in diary would need to be proven to be workable before a large-scale quantified study could be undertaken. Thus the TSG moved away from its original intention to propose a large-scale quantified test and began work on a more limited proposal, to test the feasibility of various diary designs.

A specification was sent to various research companies in the summer of 1984 and from their replies three were chosen to test three different diary designs.

The three techniques chosen were:

- (1) A fully write-in diary, with a reference card showing the list of NRS titles.
- (2) A fully prompted diary, covering 50 titles, which differed for male and female respondents.
- (3) A viewdata (electronic) diary which covered the full NRS list of titles.

The first and last diaries could potentially replace the present JICNARS readership survey while the prompted diary could only gather extra or alternative data for a limited number of publications on the present JICNARS survey.

In addition, panellists were to be interviewed using the standard NRS at the end of the test period and questioned in depth to identify reasons for discrepancies.

Our particular aims were firstly to identify whether the fully write-in diary would produce the same level of reading events as the other techniques, and secondly to see whether the diaries in general performed acceptably.

The method adopted was to place the diaries for one month, each with a panel of 20 adults. Quotas were set deliberately to give extra representations in the areas where it was thought difficulty was likely to occur (ABs and DEs, and the 15-24 and 65+ age groups). As soon as possible, after their four-week diary was completed (a suitable run-in period was used which was not less than one week) a standard JICNARS, EML recent recall interview was conducted with the panellists by a second research company. They went on to conduct an intensive re-interview where discrepancies were found between diary and recent recall claims to establish the reason for the differences. They also questioned the panels on attitudes to recruitment and diary keeping, method of keeping the diary, out-of-home readership, magazines in the house and the extent to which they had been read and the time they had been kept, opinions on the accuracy of the two readership methods and whether the diary had changed their readership behaviour.

The research agencies involved were:

Write-in diary	- Research Bureau Ltd
Prompt list diary	- Communications Research
Electronic diary	- AGB Cable & Viewdata Ltd
Intensive re-interviews	- Donald Osborne Research Ltd

FIELDWORK

To keep down cost and to ease recontact problems all respondents were recruited in either Lewisham in the South of England or Bromborough in the North.

Panellists for the paper diaries were freshly recruited but those for the electronic diary were recruited from an existing AGB Cable & Viewdata panel. In some cases other members of the electronic diary household were recruited to match the set quotas.

Fieldwork periods were as follows (these dates are for the effective diary periods and do not include run-in periods).

Half of the paper diaries	4 weeks	14/10-10/11/84
Half of the paper diaries	4 weeks	28/10-24/11/84
Electronic diaries	4 weeks	21/10-17/11/84
Re-interviews	4 weeks	15/11- 7/12/84

The staggering of start and finish dates was designed to ease the re-interview workload.

FINDINGS

(a) Categorisation of title readership claims

A table showing the full results of trying to reconcile diary/interview discrepancies is shown in Appendix II together with full comments on it from the report of the Director of JICNARS (James Rothman).

An important point to note is that the reduced number of total claims for the prompted diary are in line with the reduced number of titles covered.

Around a half to two-thirds of diary title claims were found to be correct at the re-interview stage, with the electronic and prompted diaries performing similarly and the write-in diary performing less well.

The method of analysis tends to exaggerate the inaccuracy of the diaries. Since the recording of infrequently read publications tended to be more inaccurate than for regularly read publications and the analysis is for number of titles claimed, it can be assumed that an analysis by total reading occasions would show a higher proportion of accurate records.

Both diary and interview overclaims were few. Underclaims against the interview and diaries were larger and the write-in diary produced more underclaims than the other two diary methods.

While the prompted and electronic diaries produced underclaims at a similar level for all four-week title claims the prompted diary produced more underclaims than the interview on an AIR basis.

Some discrepancies could not be clearly identified as either diary or interview discrepancies.

Overall the test suggests that the write-in diary would produce readership estimates that are too low and the prompted diary might also. The results for the electronic diary were fairly evenly balanced.

(b) Record keeping errors

People had difficulty with the concept of 'first-time reading' (the electronic diary tended to produce the most demonstrable). Although it was not serious, there was evidence of individual reading events being missed. Circumstantial evidence was found that people were failing to record reading events for short periods for all three diary methods.

An experiment in trying to identify multiple pick-ups of publications in the course of a day showed surprising variation which suggested the concept may not have been fully understood.

The write-in diary produced the most days with no reading events recorded. The prompted diary showed very few blank days.

(c) Reasons for errors in diaries

The main explanations given for identified underclaims were:

- out of home reading not claimed
- only glanced at the publication
- memory failure
- unintentional reading
- the computer failed to record their reading claim.

There were fewer cases of diary overclaims identified but those that were found were explained by writing the entry on the wrong line of the diary, writing the wrong title in the write-in diary and in slightly more cases, claiming to have miskeyed for the electronic diary.

(d) Re-interview errors

The most common reason for both under and overclaiming was poor estimation of the time period. This came out marginally in favour of over-estimating the time period rather than of telescoping.

Memory failure was the next most frequent reason given for errors. These underclaims were largely for one-off reading (eg a change because of a strike), glancing, old copies, out-of-home, etc.

Overclaims, which were few, were due to two claims of less than two minutes reading, one interview error, one failed interviewer correction and one case of forgetting a change due to a strike. Telescoping has already been mentioned above.

(e) Preference for diary or interview

Preference for diary or interview for

giving a truer picture of reading was overall at a similar level for both diary and interview. But, although it is at variance with the findings, respondents tended to favour the write-in diary and favoured the interview against the prompt or electronic diary.

Preference for type of diary for accurately recording readership

	<i>Diary</i>	<i>Inter- view</i>	<i>No opinion</i>
Prompt	5	11	2
Write-in	15	3	1
Electronic	<u>7</u>	<u>10</u>	<u>3</u>
Total	27	24	6

(f) Other points about diary keeping

Recruitment did not pose many problems and there were few complaints about keeping the diary. There was a slight preference for paper diary keeping over electronic.

Respondents gave differing views on how long they thought a reasonable time for them to be asked to keep a diary. Their median answers are shown below:

	<i>Median period</i>
Write-in	2-3 months
Prompted	1-2 months
Electronic	1 month approx.

Most respondents had no problems keeping their diary for four weeks and claimed they were up-to-date and kept their diary regularly. Most of them thought the diaries gave a complete picture of their reading.

Each method had claims that respondents had spotted errors in diary keeping but had corrected them.

(g) Conditioning

In the main people denied that diary keeping changed their reading in any way, but a minority did give specific examples of conditioning.

An analysis of total entries by week, including any run-in weeks does show a small tendency to record more events in the early weeks.

(h) Richness of data

Examination of individual's readership patterns shows interesting differences between individuals that would not be apparent from interview surveys.

MAIN CONCLUSIONS

(1) Panel members are more likely to forget to enter reading events than to enter events which should not have been recorded.

(2) The events forgotten tend to be the ones which are more incidental, eg out-of-home reading.

(3) The write-in diary missed the most reading events and the electronic diary the least.

(4) Errors were made in recording whether or not a publication had been seen before. The electronic diary produced the most errors of this type.

(5) Even after correcting for this the electronic diary produced more claims for reading an issue on two or more different days than did the other two methods.

(6) When asked whether they believed their diaries gave a true picture of their readership, write-in diary keepers made the highest numbers of this claim and electronic diary keepers, the least.

(7) A running-in period may be needed since diary panel entries in the first

week or so sometimes showed a different pattern from those in later weeks.

(8) Examples quoted in the report suggest that diary panels can produce valuable data on reading behaviour not available by other means.

(9) The report contains additional information about out-of-home reading, magazines kept in the home and lessons for future research.

RECOMMENDATIONS

In the light of the above, the TSG considered that whilst a readership diary panel could produce data about reading behaviour which would be valuable in media planning, whatever method is used the data is likely to be incomplete:

A write-in panel will risk missing reading events which should have been recorded. Suggestions have been received from the research company for reducing this risk but the TSG thinks it unlikely that further testing would show that it could be sufficiently reduced.

A prompt list diary is also likely to miss reading events. The loss is likely to be less than with the write-in diary but the number of publications that can be covered with a prompt list diary will be limited. Suggestions have been received from the research company using the method, for reducing the level of omissions and for testing a longer prompt list of publications. However, the TSG does not consider it likely that a prompt list diary could be used to collect data on all publications at present covered in the NRS in a uniform fashion.

An electronic diary is less likely to miss reading events and could cover the full range of publications but there may be problems with recording errors and respondent fatigue. Suggestions have been received from the electronic diary research company for reducing

these problems. The main difficulty, though, particularly with an electronic diary, is that the cost of establishing a panel which is sufficiently large to enable minority publications to be recorded accurately would be very high.

It must be appreciated that one of the advantages which the press has over television is that advertising can be more closely targeted on particular audiences. A sample in the 10,000+ range is needed to pinpoint these audiences accurately.

In view of the above the TSG recommended that further work on a diary panel should concentrate on its role as an adjunct to the NRS rather than as a substitute for it.

FUTURE WORK

The Diary/Panels Technical Study Group in the UK are now considering what would be the most useful additional readership information that could be gathered by diary/panels as an adjunct to the NRS.

When we have decided on the use of the panel, we will consider the type of diary/panel that will be most useful in collecting that data.

We then expect to carry out field tests with a larger panel over a longer period using some of the lessons learnt from our small-scale experiment. This could test one or more forms of diary/panels.

Finally, we aim to recommend to JICNARS the form of ad hoc or continuous diary/panel survey, fully costed, that will add most to our present knowledge of readership activity and thus allow subscribers to gain better insights into the efficient use of press media as an advertising vehicle. We would also expect such a survey to yield valuable marketing information to publishers to help them in their long term and tactical marketing decisions.

APPENDIX I

Promptness of recording

Areas of diary panel design examined in the first report of the TSG

Precision, bias and validation

Sample representativeness: sign-up and drop-out

Aspects of diary design

Accuracy of reporting: in-home/out-of-home readership

Length of publication list

Accuracy of reporting: 'home'/'business' titles

Issue-specific reporting

Accuracy of reporting: miscellaneous

Multiple reading occasions recording

Measurement of monthly magazines' readership

Field practice

Panel effects: variation in reporting levels over time

Respondent/interviewer contact

Validation of diary data via circulation.

Frequency of reporting

'How diaries work'

Understanding of readership diary task

Diary record cue mechanism

APPENDIX II

Categorisation of title readership claims

	<i>All four week claims</i>			<i>AIR claims</i>		
	<i>Prompted</i>	<i>Write-in</i>	<i>Electronic</i>	<i>Prompted</i>	<i>Write-in</i>	<i>Electronic</i>
Total claims per respondent	12.8	15.2	16.8	6.9	8.7	9.7
<i>Percentage claims which are:</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>	<i>%</i>
Correct	59	48	57	68	55	67
Interview overclaims	1	4	2	1	2	2
Diary overclaims	1	1	2	1	-	2
Interview underclaims	11	5	8	7	5	5
Diary underclaims	12	23	8	16	25	6
Either interview overclaims or diary underclaims	13	17	10	4	7	8
Either interview underclaims or diary overclaims	2	2	13	1	6	9
Total interview overclaims or diary underclaims	27	44	20	22	34	16
Total interview underclaims/ diary overclaims	14	8	23	10	10	16

COMMENTS BY JAMES ROTHMAN

The following points can be noted:

(a) The table should be read with caution. The bases are in any case small, and in addition a few respondents produced large numbers of claims in particular categories.

(b) Generally speaking, between a half and two-thirds of all claims are accurate. As might be expected, the position is better for AIR claims than for all four-week claims.

(c) The prompt list diary produces fewer claims. This is simply because the estimates are based on the publications on the list whereas the other two methods covered all publications in the survey.*

(d) The prompted and write-in diaries produced a negligible number of overclaims. The electronic diary produced a few more due to some mis-keying. The re-interview produced a few more overclaims but the number here was also satisfactorily small.

(e) Both the diary and the interview tended to produce underclaims rather than overclaims. The write-in diary tended to produce the most underclaims.

(f) The last two lines of the table confirm that as the methods stand at present the write-in diary would tend to produce readership estimates which were too low and the prompt list diary

might also do so. The electronic diary could in gross terms produce comparable estimates with the interview technique and on an average basis the gross error rate was not worse than for the re-interview.

(g) It should be remembered that respondents, having kept a diary, were in a more favourable position than respondents normally interviewed in the NRS. However it was interesting that respondents did not tend to mention their diaries during the initial interview stage and there is no reason to suppose that keeping a diary would cause respondents to under-rather than overclaim at the re-interview. It could be argued that some reading events might have been missed because respondents both failed to record them in their diaries and failed to mention them at the re-interview. Whilst it was not possible to investigate this area directly, indirect questions about magazines kept in the home and out-of-home reading did not generate additional claims.

Overall behaviour of individuals

I have also examined data for each individual to see whether taking their reading behaviour as a whole the different methods produce any biases. With the small sample sizes involved (18 usable prompt list respondents and 19 for each of the other two methods) there are few significant differences. Nonetheless the results tend to support the pattern shown in the preceding table:

(a) With all three methods there were more respondents with higher gross AIRs from the interview method than from the diary method. The difference was greatest for the write-in technique although this was not significant.

(b) Where errors could be identified the write-in technique tended to produce negative net errors for the diary relative to the interview method. This difference was significant at the 5-10% level.

* The table does not reconcile with total interview or diary AIR claims because where, as was generally the case, a diary ended before the re-interview, the AIR estimate was formed by calculating from the end of the diary period. Where respondents claimed to have read during the interval between the re-interview and the end of the diary and their diary record suggested that it was probable that they had indeed done so, they were given the benefit of the doubt.

(c) Where differences could not be identified, the bias was almost always in the direction of reducing diary four-week claim relative to interview four-week claims.

(d) If attention is turned to Average

Issue Readership claims, however, the electronic diary tended to be evenly balanced and in the case of uncategorised errors actually produce more cases which were either diary over-estimates or interview underestimates. The prompt list diary showed the opposite pattern.

APPENDIX III

Diary panel surveys examined for the first report of the TSG

J Walter Thompson (1966)
 Arbeitsgemeinschaft Leseranalyse (1966)
 IPC Magazines (1969)
 London Weekend Television (1973)
 London Weekend Television (1976)
 British Telecom (1976-79)
 Centre d'Etudes des Supports de Publicite (1979)

Modern Medicine (1979)
 La-Med (1982)
 Audits of Great Britain (1982) two studies
 Mass-Observation (1982)
 Roy Morgan Research Centre (ongoing)
 Attwood Panels BV (ongoing)