James Rothman
JICNARS, UK

9.2 COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS

INTRODUCTION

Personal or telephone interviewing, postal questionnaires and electronic interviewing are all examples of researcher/respondent communication channels. These comprise both the channels through which the researcher indicates to the respondent the information he wishes to obtain and the facilities provided for the respondent to record that information and communicate it back. They are important factors in the research process, but they do not comprise the whole of the data collection technique. We should not let the conspicuousness of communication channels delude us into considering them in isolation, but neither should we ignore them.

Generally, the same channel is used for all contacts with the respondents, but this need not be the case. For example, an interviewer can leave a questionnaire with a respondent and ask for it to be posted back, or supervised self-completion can be used at the end of a personal interview, ie respondents are asked to fill in a questionnaire in the presence of an interviewer who is able to help in case of difficulty. Computer interviewing is another example of a similar approach.

In addition to the out, back and during channels of communication, the respondent can also be provided with different means for recording the required information. Here, the principal distinction is between that of asking the respondent to record on the spot, directly from memory and that of asking the respondent to maintain a diary.

Different types of communication vehicle can have different types of cost/accuracy profiles, and these in turn can influence other aspects of the research design. For example, a capital cost is entailed if respondents are to be equipped for electronic communication with the research agency, eg through a Viewdata set. This means that this electronic means of communication is normally more suited to a panel operation, where a considerable amount of information for each individual respondent is communicated over a period than for a survey, where the amount of information communicated per respondent is less but the sample size tends to be greater. The situation would, of course, be different if a suitable means of electronic communication is widely available. Similarly, thirty years or so ago, it would have been necessary to have provided telephones to respondents to obtain a satisfactory sample with which to communicate by telephone.

Channels of communication and recording do not represent the whole of data collection. This is a gestalt made up not only of the channel, but also of the actual method used for obtaining the information and the instrument used for recording it. For example, personal interviews both to and from the respondent may be used to apply either a Recent Reading system or a Through-the-Book one. Even within the Recent Reading methodology, different questionnaire designs can be used, and the choice made can affect the results.

Here is an example of the way in which a change in questionnaire design can affect results obtained through one communication channel, but not those obtained through a different channel. The TGI uses a self-completion questionnaire to collect readership information, whilst the NRS uses personal interview. Results weighted by conventional demographics matched reasonably closely prior to 1984. At

9.2   COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS

Table 1

The effect of a questionnaire change – weekly and monthly publications

	TGI – unedited/unweighted data:			NRS
	Oct 83 – Mar 84	April – Sept 84 (Extended recency)	Oct 84 – Mar 85 –	Jan – Dec 84
Gross AIR % for selected publications:				
Sunday newspapers	124	96	110	132
General weekly magazines	57	39	53	60
General monthly magazines	33	18	27	31
Women's weekly magazines	52	30	53	45
Women's monthly magazines	57	35	49	64
Total	323	218	292	333

that time, the NRS went over to the Extended Media List and started using an extended recency scale, and the TGI changed its questionnaire to correspond. NRS results were largely unaffected by the change, but readership levels using the TGI's self-completion system fell. They were, however, restored towards NRS levels once the questionnaire was changed back again (Table 1).

To my mind, it demonstrates the way in which the meaning of a scale to a respondent is changed if a change in the channel of commui-

cation means that all positions are seen simultaneously as opposed to being presented sequentially.

The following discussion of channels for communicating with respondents is, therefore, based upon the assumption that reasonably optimal instruments are employed for each channel rather than that the same instrument is used, regardless of the channel over which it is communicated. The discussion is based upon JICNARS' experience of alternative channels of communication. This has arisen for two main reasons:

(a) Starting in 1981, JICNARS has been investigating what role, if any, a diary panel could

play in readership research. This investigation has covered three main phases:

(i) A review of previous work.

(ii) Small sample tests of three alternative diary panel methods.

(iii) A detailed investigation into the performance of AGB Cable and Viewdata's Weekly Readership Index (WRI). This was a diary panel covering approximately 1,000 individuals spread over 500 homes, which was operated for three months in the Spring of 1986 to monitor the launch of a new newspaper (*Today*).

(b) Since the launch of new newspapers became more common in Great Britain, JICNARS has established a system for conducting telephone interviews to obtain, as rapidly as possible, readership estimates for the new publications, and for comparison purposes, for the established newspapers.

In addition to the above, JICNARS has also experimented with two varieties of self-completion technique:

(i) A conventional form, where the questionnaire is completed by the respondent after the interviewer has left, and returned by post.

(ii) Supervised self-completion, where the questionnaire is completed by the respondent in the presence of the interviewer, who is able to answer any queries, to administer the questionnaire personally to respondents having difficulties, and to check through the questionnaire and resolve any ambiguities on the spot.

These systems have, however, only been used to collect non-readership as opposed to readership information. They are not discussed in this paper, since they are by Michael Ryan and Nigel Jacklin in 9.6.

Telephone interviews

Prior to the launch of new newspapers, RSL had had experience in administering telephone interviews on readership, since they needed to monitor for publishers and advertising agencies the readership of colour supplements in periods when circulation of the parent newspaper was disrupted by industrial action. In these circumstances, the colour supplement might still be distributed, so a knowledge of its readership was required. Telephone surveys were used for this purpose, both because the sample size achievable by the National Readership Survey over the relatively short period concerned would be inadequate, and because results were needed with a minimum of delay.

Because of the need which the research was required to meet, Research Services found it best to use a specific issue readership measure rather than an average issue readership one: ie respondents were asked whether they had read or looked at 'last Sunday's copy – the one published on ....' rather than whether or not they had read or looked at any copy. Obviously, in the case of newspapers, there will be little difference between these two measures. It should be noted that, in principle at any rate, the specific issue measure should be slightly lower than the average issue one, since it excludes reading that took place after the date of interview.

JICNARS decided to build on RSL's experience, and use the specific issue reading measure for both daily and Sunday newspapers in telephone interviews. These took place from a Monday to a Wednesday (inclusive) for Sunday newspapers, and referred to last Sunday's copy in the form set out above: for daily newspapers, they took place from a Thursday to a Saturday (inclusive) and referred to yesterday's copy.

Initially, two waves of interviews were conducted: one to provide a rapid indication in the

COMMUNICATION CHANNELS BETWEEN
RESEARCHERS AND RESPONDENTS

Table 2

	1st Tel. survey specific issue readership	2nd Tel. survey specific issue readership
% of adults (15 +) in Great Britain on the telephone who read:		
Mar/April 1986 Today	9.6	4.3
Oct/Nov 1986 The Independent	5.1	3.1
May 1987 News On Sunday		2.0*
% of Adults (15 +) in London ITV Area on the telephone who read:		
Mar 87 Evening News (London)	1.7	1.2
London Daily News	5.9	5.5

* Only one survey undertaken.

first or second week after launch, and the other to provide a better idea of the settle-down figure after four to six weeks.

Results obtained for different newspaper launches are set out in Table 2.

Even at the time it was decided to publish these figures, it was recognised that the first report would be of little value to advertisers, since readership levels were unlikely to stabilise at anything like the level achieved at the launch. On the other hand, there was initially considerable interest in the launch of these, the first major new newspapers to be seen in Great Britain for several years, and media researchers were expected to be able to answer questions about their performance without delay. After

the second newspaper had been launched, people became more sanguine, and it was decided that only one figure representing the settle-down level as near as possible would be appropriate.

In order to determine how long after the launch this single survey should be conducted, data from the WRI Panel, which will be described later, were used to produce the graph in Figure 1 for the readership of *Today* and *Sunday Today* over time.

On this basis, it was decided that interviews conducted in the fourth week after launch for a daily newspaper, and after four issues for a Sunday one, would provide a reasonable indication of the likely settle-down figure.

It is, of course, to be noted that results obtained from a telephone survey can only represent the readership of the population having a telephone. Systems are available for re-weighting results from telephone surveys to those of the general population (Miller 1987; Husbands 1987), but it was decided that it would be better to state clearly that the results represented readership amongst the population that was on the telephone than to use a weighting system which was both complex and rested on assumptions which were open to argument.

However, when we analysed the results, we found that the established newspapers showed discrepancies between their readership when estimated from the telephone survey and from the NRS for the population on the telephone. Results are summarised in Table 3.

It will be seen that, compared with the telephone-owning population, there was a tendency for readership of quality newspapers on the telephone surveys to be higher and that of popular newspapers to be lower than the NRS.

COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS
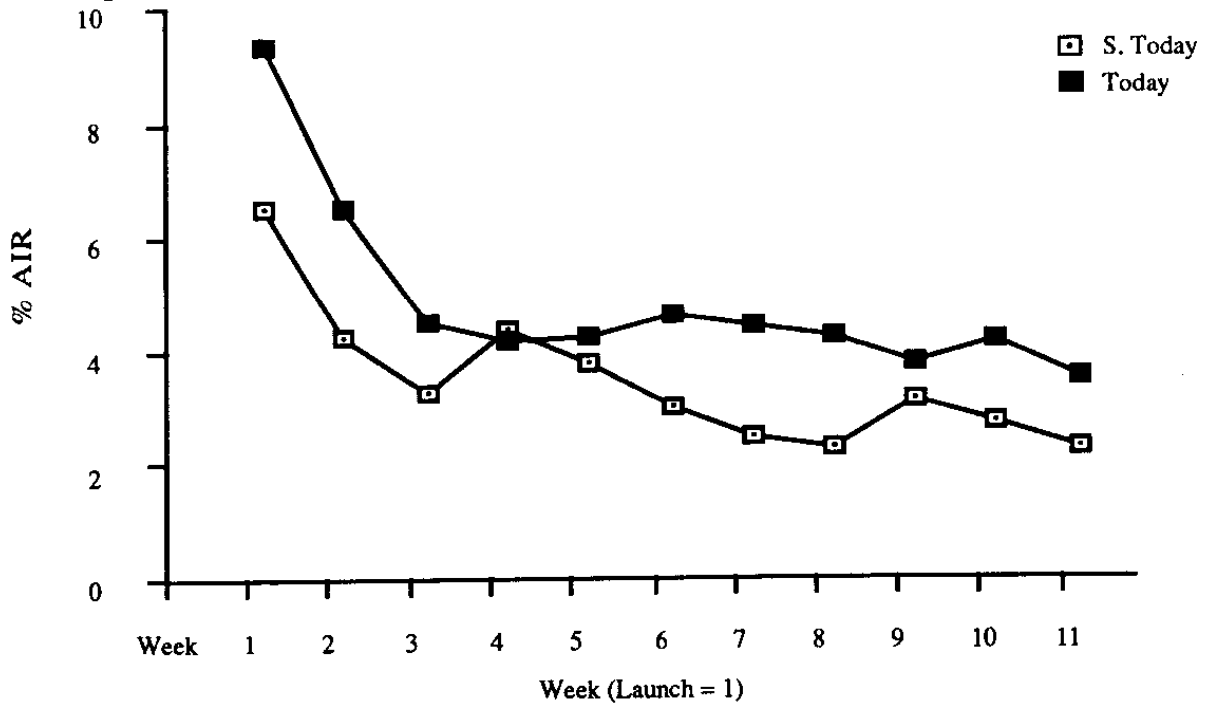
Figure 1

Readerships of Today and Sunday Today

Table 3

Gross AIR from telephone surveys indexed on gross AIR for NRS telephone owners

	Dailies	Sundays
Qualities	124	96
Mid market	101	85
Populars	88	81
Total	98	84

The following were recognised as being possible explanations for this phenomenon:

(1) Differences between the population living at addresses whose numbers are listed in the telephone directory and that claiming to have their own telephone eg:

– New telephone owners.
– Recent movers.
– Shared telephones.
– Multiple listings.

(2) Differences in response rate patterns for telephone and personal interviews.

(3) Questionnaire differences, eg use of the specific issue measure, absence of prompt cards.

(4) Difficulties in obtaining a full social grade classification on the telephone, leading to inappropriate weighting of the sample.

JICNARS is now carrying out studies in order to see the extent to which the discrepancy can be attributed to the different explanations in order to determine whether or not any action should be taken.

Diary panels

Phase 1

The diary technique is more than simply a channel of communication. It is a whole system comprising both a channel and a means of collecting and recording information.

JICNARS has carried out a programme of research to investigate the information that could be obtained from a readership diary panel.

In their paper to the Salzburg Symposium, Roger Beeson and Dick Dodson described the background to this project and the work done up to 1985. A description of this work will not be repeated. However, in the context of the present paper, it should be noted that, as part of that work, we carried out small scale tests on three different diary panel systems:

(1) A fully write-in diary, with a reference card showing the list of NRS titles, personally placed and collected.

(2) A fully prompted diary, covering 50 titles, personally placed and collected.

(3) A Viewdata (electronic) diary, covering the full NRS list of titles, recruited from an existing Viewdata panel. This diary was, of course, delivered and collected through a Viewdata system with an initial contact by telephone followed by a letter giving instructions, which were subsequently repeated on panel members' Viewdata screens.

It will be seen, therefore, that whilst we investigated three different diary systems, these employed only two different channels of communication. Interestingly enough, it was the performance of the write-in diary that was exceptional, showing fewer correct claims than either of the other two systems. It is, of course, not surprising that a write-in system should be less accurate and produce more underclaims than a prompted one, but it does demonstrate that too much emphasis should not be placed on the channel of communication considered on its own.

Nonetheless, the channel of communication itself can have an influence. For example, 72% of the electronic diary panel respondents missed recording on individual days, whilst for the prompted paper diary, only 28% had days on which they said they read nothing, and no days were left blank. Although this could be due to the system employed rather than the channel, it does suggest that an electronic channel makes it easier to use a system where recording will be missed.

Another aspect of the electronic system was that some respondents disliked it because it was impossible to modify or explain their answers. Even though few respondents using the paper-and-pencil methods took advantage of the facilities for doing this, it could be that people see the electronic diary as more of a strait-jacket. Electronic diary keepers were rather more likely to find the process tedious than those using the other two systems, although, on the whole, they still had no complaints. Nonetheless, it may be of interest that the median electronic diary panel respondent only thought it reasonable to be asked to keep a diary for about a month, compared with one to two months for the prompted diary and two to three months for the write-in one. This could, however, have been due to the fact that the electronic diary panel members did other research

work for AGB Cable and Viewdata which lasted for a shorter time.

All the above remarks only relate to the particular type of electronic diary panel that was employed. Furthermore, the research found that the electronic diary had the fewest cases where there was a clear diary underclaim, so even if those who used it liked it less, it cannot be said that the electronic channel adversely affected the accuracy of their performance.

Further details of this research can be found in the report 'Diary Panels Research, A Report on Phase 1' published by JICNARS.

Phase 2

At the end of 1985, AGB Cable and Viewdata decided to use a diary panel to monitor the launch of *Today*, the first national daily newspaper to be launched in the UK for a considerable number of years. JICNARS took advantage of this opportunity to investigate a real-life readership diary panel by subscribing to the service, and arranged with AGB Cable and Viewdata to obtain data tapes and re-interview 40 panel members at the end of the diary period using a questionnaire similar to that employed during the previous pilots.

The methodology for AGB Cable and Viewdata's panel, known as the Weekly Readership Index panel, was as follows. The company operates a national panel of homes equipped with Viewdata terminals. In these homes, household members aged 15 and over were asked to maintain a weekly diary of their readership of national daily and Sunday newspapers, Sunday colour supplements and six weekly magazines. In all, 1,210 individuals, spread over 592 homes, provided diary records for one week or more. The target sample size of 1,000 individuals was achieved during each of the first ten weeks, and did not fall below 950 during the remaining three weeks of the panel.

The panel operated over the 13 weeks, from Saturday 22nd February 1986 to 23rd May 1986, and during this time, panel members were asked to keep a paper-and-pencil diary of their readership on a daily basis, and enter the records shown in this diary each Saturday or Sunday via their Viewdata terminals to the host computer.

Results were weighted to match population statistics obtained from the National Readership Survey for October to December 1985, and further weights were calculated with the intention of ensuring that readership in week one for different types of newspaper (popular, mid-market and quality within daily and Sunday newspapers) matched the NRS for this period. These weights were applied manually, row by row, after the tables were produced, and once established, were maintained across the whole diary period.

It will be noted that the system used for the WRI panel differs somewhat from that used in the electronic panel pilot conducted for JICNARS by AGB. Apart from the scale of the operation, the major difference is that, this time, panel members were only asked to report their reading once each week rather than every day. Consequently, the system was, in principle, a mixed one, with readership being recorded on a paper-and-pencil diary, but reported electronically. I understand that this change was occasioned both as a result of experience gained from the previous pilot and also by a desire to avoid over-loading the Viewdata system.

Performance of WRI in comparison with the NRS

In the first week, the WRI produced overall levels of newspaper and programme magazine readership which were similar to the NRS. This was true even before there was any re-weighting to match NRS levels. This can be seen in Table 4.

COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS

Table 4

Comparison of WRI with NRS

Week 1

Gross Average Issue Readership (GAIR)

	WRI Wk 1 Before publication weights	WRI Wk1 After publication weights	NRS February 1988
	%	%	%
10 daily newspapers	103	92	94
9 Sunday newspapers	104	112	114
6 Sunday supplements	68	72	70
2 programme publications	42	42	42
4 women's weeklies	28	28	33

It should be pointed out that February NRS results were not available at the time that the WRI reported its first week figures. This explains why they were weighted to October to December 1987 rather than to February 1988.

This overall level of agreement for newspapers was not so close when the different categories are considered (Table 5).

It will be seen that for both dailies and Sundays, the WRI panel, even after weighting to match the general population in terms of social class within sex within region, and also in terms of age and household tenure, was still biased in favour of readers of mid-market newspapers. There was also a tendency for the panel to produce a higher estimate than the NRS for daily newspapers and a lower one for Sunday newspapers. This, however, may reflect the different frequencies of the two types of publication rather than any lack of representativeness in the panel.

Table 5

Index WRI before publication weights to NRS February 1985

	Dailies	Sundays
Popular	104	90
Mid-market	116	109
Quality	131	76
Total	111	92

After the first week, reported readership levels fell. The decline was small for newspapers but was quite marked for magazines.

The greater fall for magazines than for newspapers is probably due to the fact that, as time on the panel increased, people became more likely to omit reading, particularly casual out-of-home reading. A further factor may have been that the magazines always followed the newspapers in the diaries, since the prime purpose of the panel was to measure newspaper readership. (In the paper diaries on weekdays, daily newspapers were listed before Sunday ones, whilst on Sundays, the order was reversed. Two diary versions were used, using a forward and reverse rotation of publications within each category).

In spite of conditioning shown in Table 6, we found that the direction of month on month changes in readership from both the WRI and the NRS tended to agree with circulation changes where these were not distorted by adjustment for loss of circulation. We were also interested to see whether the week on week fluctuations in the WRI reflected circulation changes.

Official weekly circulation estimates are not published in Great Britain because of technical and other problems. However, we were able to obtain in confidence weekly circulation

Table 6

Index* WRI/NRS over time

	10	9	6	2	4
	Dailies	Sundays	Sunday supplements	Programme publications	Women's weeklies
Week 1	97	98	102	100	83
Weeks 2−5	94	96	91	79	64
Weeks 6−9	92	100	90	77	71
Weeks 10−13	93	97	89	71	61

* The table shows the results in index form, but the same pattern would be shown if actual WRI gross average issue readership estimates were taken instead.

estimates for six daily newspapers, although it was emphasised to us that these were approximate and not necessarily totally representative.

A comparison of these figures with the WRI weekly readership estimates showed that, over the 13 weeks, the two sets of figures were positively correlated, with coefficients in the range 0.2 to 0.7. Although this level is quite small, suggesting that less than half the variance in WRI estimates could be accounted for by changes in circulation, not only were the circulation indices approximate, but also allowance must be made for sampling error in the WRI estimates. The correlations quoted were, in most cases, of a similar order of magnitude to the maximum one could justifiably expect, bearing in mind the effective sample size of the WRI panel.

This suggested the conclusion that, although short-term fluctuations in readership could, at a cost, be measured by means of a suitably sized readership diary panel, the size that would be required for real changes to be measured with sufficient accuracy for the exercise to be worth-

while was likely to be considerably larger than the 500 households provided by the WRI. Furthermore, examination of the fluctuations for individual newspapers suggested that readership estimates varied in an irregular fashion, so that a knowledge of weekly fluctuations would not necessarily be of assistance in media planning. Nonetheless, a panel, if it was of a sufficient size, could help in areas such as identifying successful editorial and marketing strategies, or measuring the audience achieved by a particular campaign. The Diary Panels Technical Study Group, however, noted that, if required, similar information about readership or circulation variation could be obtained from other sources which might be more economical, particularly if agreement could be reached for wholesalers or retailers to co-operate.

Over and above the differences by different types of newspaper, we also found that, in spite of the agreement in overall readership levels, the WRI differed from the NRS in the way in which it measured population sub-groups eg WRI readership was higher than the NRS

COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS

Table 7

Gross average issue readership

Gross AIR percentages for

	10 Dailies			15 Sundays **			6 Weeklies		
	WRI* (a)%	NRS+ (b)%	Index (a/b)	WRI* (a)%	NRS+ (b)%	Index (a/b)	WRI* (a)%	NRS+ (b)%	Index (a/b)
Total	106	95	112	171	163	105	54	74	73
15-44	99	103	96	152	174	87	38	81	47
45+	113	87	130	194	150	129	73	67	109
ABCI	99	95	104	184	196	94	58	82	71
C2DE	110	96	115	163	141	116	52	69	75
Male	109	109	100	169	176	96	32	46	70
Female	103	83	124	174	151	115	75	100	75
North	99	87	114	148	153	97	54	66	82
Midlands	102	110	112	195	156	125	56	73	77
South	115	110	105	187	177	106	54	84	64

* WRI figures have not been weighted to NRS levels
+ NRS March-May 1986
** Including supplements

among over 45s, C2DEs, women, and in the Midlands.

It can be seen from Table 7 that, even as far as the total sample is concerned, although the results were reasonably comparable for newspapers, the WRI figure of 106 for the gross average issue readership percentage of daily newspapers was significantly higher than the NRS value of 95. The lower levels of the WRI for magazines are, of course, due to the fact that WRI readership estimates for these publications declined so sharply over time.

One subject in which JICNARS was particularly interested was to see if a readership diary panel could be used to test different schedule

analysis models. As is well known, one of the disadvantages of the recency method of measuring readership is that a schedule containing more than one insertion in the same publication cannot be evaluated directly. Instead, the frequency and coverage achieved by the schedule have to be modelled using the claimed frequency of reading the different publications. With a panel, on the other hand, schedules can be evaluated directly. Consequently, if reliable panel results are available, these will provide a means for evaluating or calibrating alternative schedule analysis models.

JICNARS was, therefore, particularly interested to see if the WRI panel could be used in

this way. There is, of course, one type of schedule which can be evaluated directly from a recency-based survey. This is the one in which there is only one insertion in each of several publications. The results for this schedule are simply the number of people reading zero, one, two, etc. titles in the category.

The charts in Figure 2 compare the performance of the WRI and the NRS for three such schedules, one covering ten dailies, another Sunday newspapers and their supplements and the third, the six weekly magazines on the survey. Each bar represents the percentage of the sample reading a particular number of publications according to either the NRS or the WRI.

It will be seen from this comparison that the WRI could not be used directly to test schedule evaluation, since even where gross impacts were similar, the shape of the frequency distribution was different. In general, the NRS showed more variations in reading behaviour between respondents than did the WRI, which was concentrated more heaily around the average. For example, the WRI showed 62% reading exactly one newspaper on any given day, whereas, according to the NRS, this figure was only 45%. The opposite pattern was shown for claims not to read *any* papers.

These discrepancies could be due to atypical readers being less inclined to join a panel. People who do not read any newspapers might refuse to join on the grounds that they did not read, whereas those who read large numbers would have found keeping a diary more burdensome and thus might either refuse to join or have been more inclined to drop out early. A further factor might have been that the re-interviews suggested that WRI members were relying on the regularity of their reading behaviour to help them complete their diaries. Whilst they might have routinely claimed

readership on every day when the publication was delivered, even though they might not occasionally have had time to see it, they also might have been more inclined to have missed irregular or casual out-of-home reading.

We were, however, able to use the WRI to help us check the extent to which respondents varied in their interpretation of the frequency of reading scale on the NRS. This consists of questions asking respondents to identify the publications that they read:

– almost always (at least three issues out of four)

– quite often (at least one issue out of four)

– only occasionally (less than one issue out of four)

or

– not in the past year.

As part of the re-interviews with the 40 panel members, we used an abbreviated NRS questionnaire. This enabled us to check whether the answers given agreed with frequencies of reading shown in the diary panel records. Later on in the interview, respondents were questioned in an attempt to resolve any discrepancies.

It can of course be argued that, by the time of the interview, panel members might have become more aware of their reading frequencies than the general population. However, the recordings of the re-interviews suggested that respondents treated the NRS-type questioning as a separate exercise, and did not appear to be attempting to recall what they had entered in their diaries.

Results of this comparison are shown in Table 8. This table is based on all cases where a

COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS

Figure 2

NRS/WRI comparison of number of publications read

10 DAILIES

15 SUNDAYS
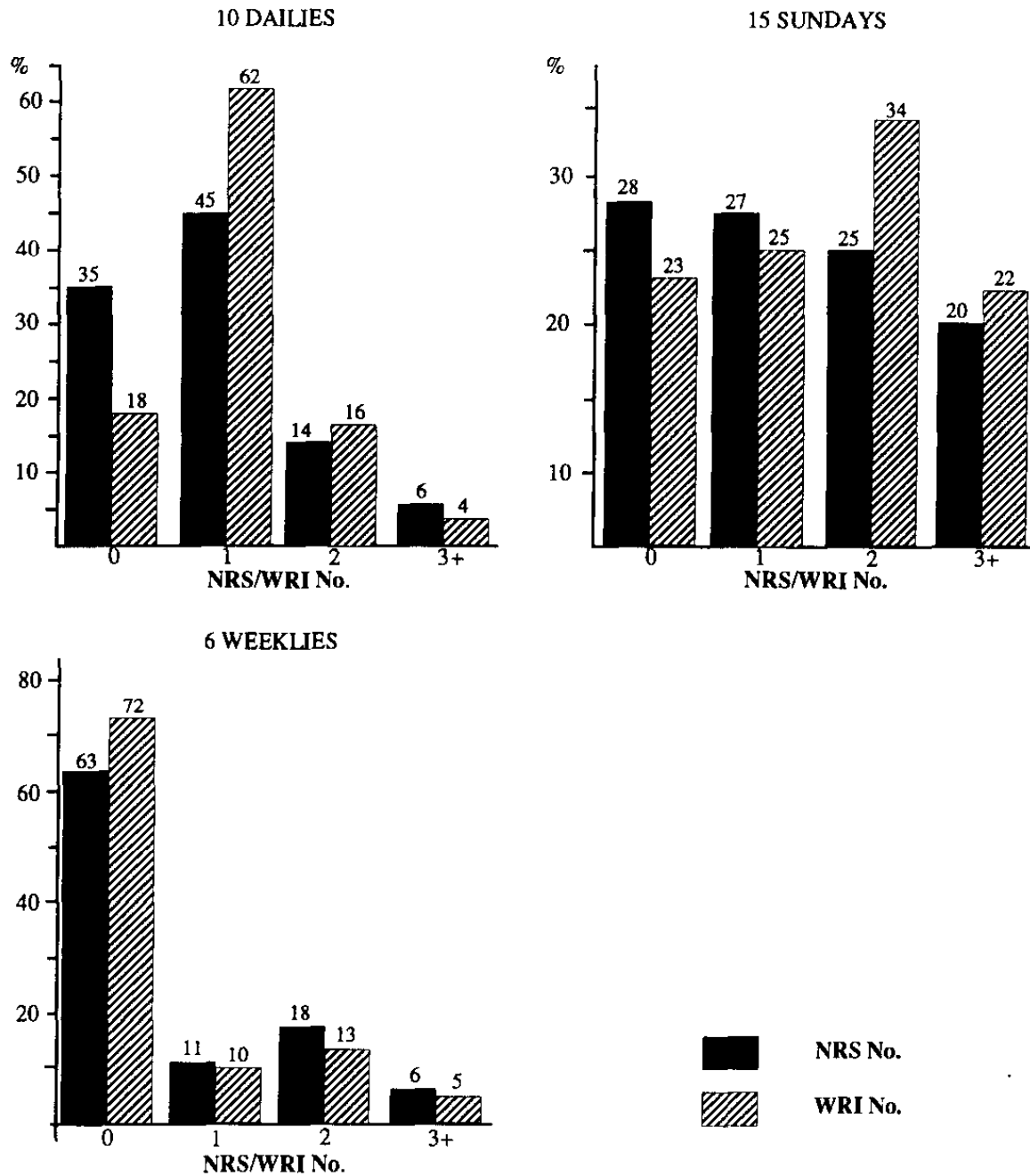
6 WEEKLIES

publication was claimed, either on the diary or the re-interview. In each case, the number of issues of a publication which had been read was expressed as a percentage of the maximum number of different issues that could have been read, bearing in mind the number of weeks over which the respondent had returned diaries. Percentages of over 75% were taken as equivalent to 'almost always', percentages in the range 25% to 75% were treated as 'quite often', whilst any other non-zero figure was interpreted as 'only occasionally'. A zero entry was treated as a discrepancy, since cases where the publication was not claimed either on the diary or at the re-interview had been excluded.

In the table, discrepancies have been categorised, according to whether explanations given at the re-interview suggested that they were due to mistakes made when answering the readership section of the questionnaire, or to mistakes made when completing or entering the diaries. The table also shows a third category of 'other discrepancies' where the source of the mistake could not be identified.

It will be seen from Table 8 that in three quarters of all cases, the interview estimates of frequency of reading agreed with those in the diary. In a further 7% the diary record was found to be in error, whilst in another 7% it was not possible to determine which system of recording produced the discrepancy. This suggests that the interview technique only produced inferior answers to a diary about 13% of the time, whilst the diary proved inferior to the interview in 11% of cases.

On the whole, both techniques produced under rather than overclaims. This arose principally because, as in Phase 1, respondents were inclined to miss isolated or rare reading events.

Table 8

Comparison of diary panel and re-interview frequency of reading claims

	Total	%
Number of respondents interviewed	40	
Number of titles claimed	505	100
No discrepancies	381	75
Discrepancies due to mistakes at:		
Re-interview stage:		
(1) overclaims	13	3
(2) underclaims	35	7
Diary stage:		
(3) overclaims	17	3
(4) underclaims	22	4
Total		
(5) overclaims	30	6
(6) underclaims	57	11
Other discrepancies:		
(7) Re-interviews more frequent than diary	20	4
(8) Diary more frequent than re-interview	17	3
Total cases in which frequency at:		
(9) Re-interview greater than diary (*)	55	11
(10) Diary greater than re-interview (**)	69	14

* Lines 1,4 and 7
** Lines 2,3 and 8

**COMMUNICATION CHANNELS BETWEEN
RESEARCHERS AND RESPONDENTS**

Accuracy of the WRI panel

There are a number of factors which can affect the accuracy of a diary panel. These are as follows:

– Representativeness of the original panel.

– Drop-outs and panel turnover.

– Completion errors.

– Sample variation.

The question of the representativeness of the original panel has already been discussed. It will be recalled that there is evidence to suppose that, even after reweighting the panel to match population demographics, it was still not representative in terms of reading behaviour. This is always a problem with diary panels, since the proportion of those initially approached who actually return diaries is so small, perhaps as little as 25%. In the case of the WRI, of course, this panel was recruited from AGB Cable and Viewdata's national electronic panel, and this, in turn, had been set up using quota sampling to match different demographic characteristics. Although, for obvious reasons, exact estimates are not available, the overall response rate is certainly less than the 70% or so achieved using a random survey such as the NRS.

Once the panel had been signed up, the response rate only fell slowly over time, from a high of 90% in week two after the panel was fully established down to a low of 77% by week 13, with an overall average of 83%. As often happens with panels, the less regular reporters were the ones who were most likely to drop out. Consequently, the average number of weeks that a respondent reported for rose from 11.4 in week two to 12.2 in week 13. This meant that people reporting in later weeks tended to be atypical in the sense that they were

more reliable. Furthermore, the older groups were more inclined to stay on the panel than the younger ones. This could, of course, be offset by the demographic weighting.

With the WRI, there were two possible sources of completion error. In the first place, as with any paper-and-pencil diary, respondents may not make accurate entries. Secondly, even if entries were accurate, they might not be entered correctly into the computer. In Phase 1 of the research, we found that some of those re-interviewed from the electronic panel adopted the ever-popular stratagem of blaming the computer to excuse the errors they admitted in their diary records.

Although diary and NRS estimates of frequency tended to support each other, inconsistencies in the details recorded in the diaries about each reading event suggested that respondents were not always completing them correctly. This could have been due to a failure to understand the definitions, since the re-interviews showed that panel members often had not understood certain terms used in the instructions on their Viewdata screens. For example, out of the 40 respondents re-interviewed, at least 12 misunderstood the term 'first time reading' in a variety of ways.

This meant that it was not surprising to find that respondents made errors, such as claiming to be reading the current issue of a daily newspaper and also to have read it on a previous day, or for a Sunday newspaper, to be reading the current issue on a Sunday, and to have read it on a previous day. Also, there were a few cases where panel members claimed to read the current issue of a daily paper on a Sunday, or to be reading for the first time an issue of a Sunday newspaper or supplement for which there was an earlier entry in their diaries. 17% of claims not to be reading a newspaper for the first time were demonstrably false, and even though the finite life of the diary meant that not all claims

of first time reading could be checked, about 2% of such claims were also false.

Figures for weekly magazines were even higher, but this could have been due to panel members being able to obtain magazines even further in advance of the official publication day than we had allowed.

AGB Cable and Viewdata helped us investigate the accuracy with which panel members entered their reading into the computer by comparing the entries in the paper-and-pencil diaries returned by 27 respondents with computer print-outs of their Viewdata returns. 14 of these 27 were chosen from the 40 who were re-interviewed.

The results of this comparison showed that only about 3% of entries for newspapers and 4% for magazines were incorrect. Positive and negative errors were equally common. However, false positive entries may be equally likely to benefit all publications, while false negative entries may be proportional to the number of claims made for each publication. If so, small publications could gain more than large ones from inaccurate data entry (Table 9).

It should, however, be noted that it is possible that some of the discrepancies we found were not due to mis-keying, but reflected cases where panel members had recalled an additional reading event when they were entering data into the computer, and had keyed it in without correcting their diaries.

With regard to the veracity of people who blame computers, it is interesting to note that 14 of the 27 sets of diaries we checked came from panel members we re-interviewed. Of these, six in 13 different cases attributed discrepancies between the diary record and the re-interviews to computer error or mis-keying. In seven of these 13 cases, there were indeed discrepancies between the paper diary and the

Table 9

	Large publication	Small publication
Actual reading events	1,000	100
False positive claims	10	10
False negative claims (say 1% of actual claims)	10	1
Recorded reading	1,000	109

computer record. These seven cases came from two of the six respondents who blamed the computer. It would seem, therefore, that mis-keying or computer error may only be guilty half the time that it gets the blame.

The final factor which can affect the accuracy of a diary panel is sample variation. This is not simply a matter of sample size, but also of the way in which a panel is designed. The key figure is, therefore, the effective sample size, ie the actual sample size divided by the design effect.

We went to some trouble to estimate the design effects and effective sample sizes of the WRI panel by examining different factors in turn. These were as follows:

Variation in weights. The more variation in the weights which have to be applied to a sample to bring it into line with the population, the greater will be its sampling error. An analysis of the frequency distribution of the weights applied to the WRI panel showed that these varied quite considerably, ranging from less than 0.2 to more than 3.0. This meant that, when making absolute estimates, the effective sample size of the WRI panel was likely to be reduced on these grounds by 30%, ie a design effect from this cause of 1.4.

Clustering within the household. The WRI panel contained an average of about two

individuals per household. This had advantages both of economy and of enabling the reading behaviour of different members of the same household to be compared. On the other hand, the reading behaviour of husbands and wives is highly correlated. We found that, in general, if one member of a household claimed to have read a newspaper, the other was more likely than not to have also read it, so the correlation co-efficient between the reading of different household members for newspapers and programme magazines was 0.8 or more. This means that, in terms of sample size, the panel is better thought of as one of 500 households than of 1,000 individuals.

Correlations over time. It is, of course, well known that a panel is a very sensitive indicator of changes over time because behaviour in one period is highly correlated with that in another. The greater is this correlation, the less is the variance of the estimate of difference between the two periods. This can be thought of as increasing the effective sample size compared with two random samples.

We found that over 60% of those reading a daily newspaper on one day had read it on the corresponding day of the preceding week, and for Sunday newspapers, the corresponding proportion was generally in excess of 70%. Programme magazines scored over 80%, but women's weekly magazines only 30-50%. This meant that the average correlation between successive weeks for the number of days on which a specific daily newspaper had been read was 0.89. Where only dichotomous answers were possible, the correlation was 0.73 for Sundays, for programme magazines 0.83, and for women's weeklies 0.65.

Using these results, it is possible to estimate the design effects and effective sample sizes for different types of estimate. The calculations are shown in Table 10.

This demonstrates that, with one exception, the WRI is markedly less efficient than a simple random sample, or for that matter, than the NRS, when making absolute estimates. On the other hand, the table confirms that a panel is a more sensitive way of measuring changes in readership than two random samples. Even so, the gain is perhaps rather less than one might have hoped. Even in the most favourable case, one would need a panel of about 2,400 (allowing for the NRS design effect) to achieve an accuracy equivalent to estimates based on the six months NRS samples – the minimum period over which JICNARS normally permits results to be reported.

Attitudes to the panel

Re-interviews with the panel members provided information about how respondents reacted to the channel of communication which the WRI provided. 13 of the 40 panel members re-interviewed found that it was boring and repetitive to be on the diary panel, the main problem being that the keying in of the information took longer than necessary because of delays in the computer response. This, of course, is likely to be less of a problem as computer facilities improve. Attitudes were, however, not entirely negative; five of the 40 found the technical nature of the Viewdata equipment interesting, and six were pleased to have the opportunity of contributing to this type of research.

It is also possible that the system of asking people to keep a diary and then enter the data themselves at the end of the week may have meant that diaries were only completely written up when it was time to make the computer entries. Answers given by respondents suggested that they may have relied unduly on the regularity of their reading behaviour. This could have caused them to miss the exceptional case where they read a publication which they

9.2 COMMUNICATION CHANNELS BETWEEN RESEARCHERS AND RESPONDENTS

Table 10

Estimation of design effects and equivalent sample sizes

		Estimates for:			
		Dailies	Sundays magazines	Programme weeklies	Women's
A	DEFF due to weighting	1.4	1.4	1.4	1.4
B	Correlation between household members	0.82	0.82	0.87	0.20
C*	DEFF due to household clustering (1 + 1.04B)	1.85	1.85	1.90	1.21
D	Correlation between reading in successive weeks	0.89	0.73	0.83	0.65
E	Proportion reporting in both weeks	0.83	0.83	0.83	0.83
F	Corrected correlation (D x E)	0.74	0.61	0.69	0.54
G	DEFF for short period absolute measures (A x C)	2.6	2.6	2.7	1.7
H	Minimum DEFF for long period absolute measures (F x G)	1.9	1.6	1.9	0.9
I	Panel DEFF for changes excluding cluster effect (1 – F)	0.26	0.39	0.31	0.46
J	Panel DEFF if all members of a household change at the same rate (C x I)	0.48	0.72	0.59	0.56

Simple random samples equivalent to a panel of 1,000 using the WRI design for:

	Dailies	Sundays magazines	Programme weeklies	Women's
Short period absolute measures	385	385	370	590
Long period absolute measures	525	625	525	1,110
Changes maximum	3,845	2,565	3,225	2,175
Changes minimum	2,085	1,390	1,695	1,785

* 1,210 individuals spread over 592 households give an average household 'cluster' size of 2.04. The formula for the design effect is (1 + (b – 1)p) where b is the average cluster size and p is the correlation between members of the same cluster.

did not normally see or did not read one which they read regularly. However, the panel members themselves thought that their paper-and-pencil diaries were an accurate record of their reading, even though they admitted that they could have made keying-in errors. As we have seen, panel members' performances in this respect may have been better than they claimed.

Another aspect of using a Viewdata system was that, in the majority of cases, members of a household did not necessarily key in their own data. In half the cases, one member of the household keyed in the data for the others, and in other cases, some respondents either took turns, or could not describe the system they used. This meant that only 15 out of the 40 panel members studied claimed that each member of the household entered their own data. Obviously, as researchers, we would prefer a system in which most people declare their reading in private and we do not run the risk of the results being distorted because of the opinion or promptings of others.

From the point of view of the present paper, a particularly interesting aspect of these answers is that it shows the way in which the features of a system can cause respondents to depart from their instructions in ways which can affect the research results. This is why, in my view, it is necessary to consider research as a total system, and not consider channels of communication in isolation.

Conclusions

The examples discussed in this paper:

(1) Show that the channel of communication is a factor in the research design whose influence interacts with the research methodology and

consequently cannot be considered separately from it.

(2) The relative costs of establishing channels of communication and passing information through them vary greatly. This not only has an impact on the research design but can also influence the nature of the different types of readership measure which are collected.

(3) At first sight this suggests that because of the synergy that exists between them, additional information could be obtained by using different research designs in conjunction with dfferent channels of communication. However, the combination of results from different research approaches is not straight forward. Readership estimates obtained from one approach are not necessarily comparable with those obtained from another and even when they are comparable on an average basis the shape of their frequency distribution across the population can vary. It is possible that this problem could be overcome by using a thorough system of controls but demographic weighting on its own is unlikely to be sufficient.

(4) Where a subsidiary survey is conducted using a different channel of communication which necessitates both a change in the sample design and the use of different types of readership measure, it is extremely difficult to interpret the reasons for any discrepancies between the subsidiary survey and the main research.

(5) Diary panel data suggest that short term fluctuations in readership vary in an irregular fashion so a knowledge of them will not necessarily be of assistance in media planning. A panel if it was large enough could measure these with sufficient accuracy to help identify successful editorial or marketing strategies or to measure

the audience that had been achieved by a specific campaign. However a readership diary panel is not necessarily the most economical way of obtaining this information.

In conclusion it can be said new technology will increase the range of channels of communication available to the researcher and will make it easier to obtain certain types of information. There will however be a problem in ensuring comparability between research produced through different channels and it is certainly not the case that new technology necessarily implies greater accuracy.

References

Allt, B (1985). Interviewing and readership – some thinking behind the Extended Media List design. Salzburg Proceedings.

Beeson, R and Dodson, D (1985). UK development work on diary panels. Salzburg Proceedings.

Husbands, C T (1987). The telephone study of voting intentions in the June 1987 General Election. *Journal of the Market Research Society*, 29, 4, October.

Miller, W L (1987). The British voter and the telephone at the 1983 Election. *Journal of the Market Research Society*, 29, 1, January.