NRS RESPONSE RATES: EXPERIMENTS AND INVESTIGATIONS

Erhard Meier, RSL - Research Services Ltd

Synopsis

This paper is in three parts. First it describes variations in response rates within the British National Readership Survey. These variations illustrate that response rates are not a uniform affair across all types of area, nor are they the same by interviewer. The second part describes various experiments and investigations conducted over recent years, which were designed to help us understand response rate problems better and to show possible ways of improving them. And finally, one particular experiment, a test of the mixed method approach to data collection, is described in detail. While this test did not produce higher levels of response than the current NRS does, it showed promise. However, only if mixed method data collection of readership data were acceptable, would this method be a candidate method for the NRS.

Introduction

The response rates of the British National Readership Survey are still of concern. They currently stand at around 62%. As I had reported in my 1995 Berlin Symposium paper on the same subject, NRS response rates had temporarily been as low as 54%, following first the introduction of PAF (Postcode Address File) as the sampling frame in January 1992 and then of CAPI (Computer Assisted Personal Interviewing) in July 1992.

We regard the CAPI effect on response rates as a temporary one due to the initial unfamiliarity of interviewers with the new computerised task and with handling the new situation of introducing it to informants. The PAF effect however was real - PAF is a more comprehensive frame and hence less response rate friendly than the previously employed Electoral Registers; it requires a two-stage sampling procedure to be undertaken by the interviewer; and the subsequent response rate calculation is more stringent.

In this connection it is illuminating to note that the second stage response rate with PAF, which was 69.4% in NRS 1996, compares fairly well with the reported overall NRS response rate achieved on the basis of the Electoral Registers which was 66.4% in 1991, the last year that that sampling frame was used for the NRS. With the Electoral Registers named residents were sampled in the office and issued to interviewers in the field, and the response rate then reflected the proportions which were successfully interviewed, except for a small part of the sample where the issued person had died or moved away and a substitute person had to be selected; the response rate for this part reflected the successful conversions of these.

With the PAF sampling method, the sampling procedure requires the interviewer to ascertain first the names and ages of residents, and then to make a selection of whom to interview from these. The response rate of this first sampling stage, which was 88.6% in NRS 1996, is a reasonably high figure. However, this figure needs to be taken into account for the calculation of the overall response rate. Thus the multiplication of first stage response (e.g 88.6%) with the second stage response (e.g 69.4%) produces the overall reported figure which was 61.4% in NRS 1996.

This little discourse on the effects of two different sampling frames on the sampling procedures in the field and hence on reported response rates, may demonstrate how difficult it is to compare the response rates reported by different readership surveys, particularly if they are from different countries. We may say, as a general rule, that the less "filtered" the issued sample is which the interviewer needs to work with, the lower tends to be the reported response rate; the more "filtered" it is, the higher the reported response rate, if everything else is equal.

The filtering-out process of the sample before it is issued to interviewers can include intentional and unintentional items and can involve correct and erroneous deletions. It all depends on how it is done in practice.

When we had used the British Electoral Registers, we had clearly filtered out any addresses for which no resident had registered his or herself as eligible to vote. With PAF, none of these were filtered out. Likewise, sampling frames which, as in many countries, are constructed by the research practitioners themselves, in the absence of reliable or suitable general sources, tend to be particularly prone to the filtering process. Here the clues lie in the universe definitions used and how they are applied on the ground. If, for instance, the construction of the sample frame involves the enumeration of addresses on the ground, through inspection by fieldworkers, and the task is defined, say, as that of enumerating only those addresses which are private and contain residents of the nationality of the country concerned, then many addresses are filtered out by design. My guess is that an additional number of addresses may also be filtered out in error; as it seems inevitable that classification errors occur in this type of operation.

This all adds up to saying that the NRS response rates after the introduction of PAF as the sampling frame, while reported to be lower than those based on Electoral Register sampling, were in fact not indicating a worse representation by the achieved sample of the sampled population. On the contrary, representation was probably better through the use of the more comprehensive sample frame.

We also believe that the probability method of sample selection is the only viable one for the National Readership Survey, even if the response rate is of concern. The alternative - quota sampling - would not be able to produce the same representation of the population as does probability sampling. I am thinking here for example of individual ages within give age groups, of individual household sizes within, say, small, medium and large households, of individual occupations within the professions or any other group of occupation, and of individual degrees of media usage within, say, light, medium and heavy usage of any given medium.

The reason why this is so lies in the fact that, with the quota approach, controls are by nature limited in number. And those groups which are controlled by quota still will tend to be deficient by individual sub-groups within the main group, because no consideration can be given by the interviewer to these sub-groups when fulfilling the given quota, except in a general intention of spreading the type of person within that main group. A striking illustration for this can be the age groups. Both the lower and the upper end of ages are most likely the ones where we can demonstrate such deficiencies in quota sampling: if the quota is by age group 15-24, which is often the case, then the outcome by individual ages is likely to be a heavy skew towards those aged 20-24 at the expense of the 15-19 year olds. And if the quota is age group 65+, then the outcome by individual ages tends to be a heavy skew towards those under 75 at the expense of those over 75. (This illustration is based on an analysis of JICRAR data of 1991.)

With probability sampling, differential sample profiles also occur, due to differential response rates. This is one of the main concerns about low response rates. However, we believe that these types of bias tend to be of less severity with probability sampling, and they tend to be open rather than hidden.

In the following, we show some examples of differential response rates within the National Readership Survey, and then proceed with a description of recent experiments and investigations.

Variations in Response Rates Within the National Readership Survey

There are great variations in the average response rates of the NRS, for instance by region, type of area, and sex of interviewer. Both the existence of these variations and that they tend to widen, are of concern.

First, response by region. The table below compares the reported response rates by survey region for NRS 1984 and NRS 1996:

Response Rates by Region

	NRS 1984 %	NRS 1996 %	Difference %
Survey Region			
London	65.0	48.0	- 17.0
South East & Anglia	67.6	56.6	- 11.0
South West and Wales	68.0	64.0	- 4.0
Midlands	73.9	59.2	- 14.7
North West	74.0	61.6	- 12.4
North East & North	75.1	70.4	- 4.7
Scotland	74.5	73.1	- 1.4
Total Great Britain	71.0	61.4	- 9.6

This table, after making allowance for the fact that the 1984 figures are based on Electoral Register sampling and the 1996 figures on PAF sampling, and therefore not strictly comparable, seems to indicate the following:

a)   that the response rate in London which always was the worse of any region, has declined to a greater degree than other regions in the last twelve years;

b)   that the declines in Midlands (includes Birmingham) and North West (includes Manchester and Liverpool) are also relatively greater than the average;

c)   that three survey regions have declined relatively little: South West & Wales, North East and North, and Scotland.

d)   that South East & Anglia have declined close to the national average.

The response rates by survey region seem to give a broad indication of the different response rates due to type of area: urban areas versus rural areas. It is not surprising that London, the large metropolis, shows the worse decline, while Scotland, which is in spite of the two conurbations Glasgow and Edinburgh a predominantly rural region, showed together with South West & Wales and North East & North the least decline. However, there are surprises, as the following analyses may show.

A particular type of response analysis regularly undertaken by us is that by ACORN, the geo-demographic classification system of small areas. Each sampling point on the NRS is classified to one of 54 different ACORN types, and the following table shows the six types with the highest, and the six types with the lowest average response rates achieved during 1996, together they represent about 13% of the total population:

Response Rates by ACORN Type
NRS 1996

Six Types with Highest Average Response

Type No.	Description	No. of points during year	Response Rate %
51	Council Flats, Greater Hardship, Many Lone Parents	20	76.1
50	Council Areas, High Unemployment, Lone Parents	40	72.3
49	Council Flats, Very High Unemployment, Singles	16	71.5
6	Agricultural Villages, Home Based Workers	38	70.6
40	Council Areas, Older People, Health Problems	61	69.7
46	Council Areas, Residents with Health Problems	44	69.5

Six Types with Lowest Average Response

Type No.	Description		
16	Well-off Town and City Areas	26	49.0
53	Multi-Ethnic, Severe Unemployment, Lone Parents	19	48.4
47	Estates with High Unemployment	26	47.9
38	Multi-Ethnic Areas, White Collar Workers	19	47.9
20	Gentrified Multi-Ethnic Areas	22	45.5
18	Furnished Flats & Bedsits, Younger Single People	18	38.7
National Average		2,644	61.4

At first glance, the surprising finding is that shown in the top half of the table, the six area types with the highest response rates. We would have expected to see rural or semi-rural areas at the top of the table. Instead, only one of the six areas was of such a kind, namely area type 6 described as "agricultural villages, home based workers". The other area types at the top of the table are typically urban: council flats and council areas!

The bottom half of the table shows findings as expected. Especially the area type with the lowest average response rate during 1996 does not at all surprise: "furnished flats & bedsits, younger single people".

So, what do the area types at the top of the table have in common, in contrast to those at the bottom of the table, that produces their relatively high response rates?

The answer seems to be the level of availability for interview at the time of the interviewer's call. The area types at the top of the table are urban indeed, but they may be peopled to a much greater degree with those who are homebound or likely less away from home than those at the bottom of the table. In this respect, our findings are little surprising and entirely plausible. "High unemployment" can be seen at both the top and the bottom, indicating that areas with high unemployment cannot be easily categorised as good or bad for response rates. It depends on other factors.

Apart from area type and region, there are also variations in response rates by type of interviewer, notably by sex and by experience. The proportions of male and female NRS interviewers was 26% male and 74% female in 1996. However, the male proportion rose to 47% in London. It is difficult to divide the area effect from the gender effect when looking at response rates by sex of interviewer as there is a strong correlation. Male interviewers tend to be assigned to the tougher areas, and it is not clear to what degree the lower response rates achieved by male interviewers shown in the table below, are a reflection of the area type they worked, or a reflection of their accomplishments:

Response Rates by Sex of Interviewer - NRS 1995/6

	Male Interviewers	Female Interviewers
Total GB		
Number of interviewers	75	219
Number of points	611	2046
Response Rate	55.6%	61.2%
London		
Number of interviewers	22	25
Number of points	130	157
Response Rate	38.5%	51.9%

(Based on NRS July 1995 to June 1996. Response rates of this table are calculated on original returns, that is prior to re-issues which are carried out by different interviewers to the original ones.)

With the factor experience, the findings seem more conclusive. Interviewers with considerable experience tend to turn out better response rates than interviewers with lesser experience. This may be seen from the table below, which shows an analysis of response by length of experience of working NRS assignments:

Response Rates by Experience - NRS 1995/6

	New NRS Interviewers (who started in last 12 months)	1-2 years NRS Experience	2+ years NRS Experience
Number of interviewers	59	44	187
Number of points	244	306	2092
Response Rate	49.8%	51.6%	60.9%

(Based on NRS July 1995 to June 1996. Response rates of this table are calculated on original returns, that is prior to re-issues which are carried out by different interviewers to the original ones.)

As may be seen from the above, interviewers with two or more years' experience of working NRS assignments are working the bulk of NRS points in the course of the year (79% of points during the period July 1995 to June 1996). But there is a turnover of interviewers. There are always those with less than a year's experience and those with less than two years' experience. These tend to achieve markedly lower response rates than those with longer experience, with very little difference between those of less than one and those of one to two years' experience.

These findings by interviewer experience are very much the same across the country and by all types of area.

An increased turnover rate of interviewers cannot be blamed for falling NRS response rates. The turnover has been fairly stable over the years, with, in 1995, 26% of interviewers starting their NRS career and 28% completing their last NRS assignment that year, out of all interviewers working NRS assignments that year.

Summary of Recent Experiments and Investigations

A number of additional measures, experiments and investigations were undertaken in the last two years with the aim of finding ways of improving the NRS response rates. They were conducted on the basis of the acknowledgement that there are different levels of co-operation, which are often depending on the situation in which the interviewer finds the informant, on the type of area and on the experience of the interviewer. These included:

a)   the addition of names to the addresses issued from PAF. For about 70% of PAF addresses we are able to merge elector names from the Electoral Registers with PAF. This helps the interviewer in the first stage contacting process.

b)   the requirement of selecting two persons for interview in large households. This reduces the weights due to household size, but it also helps the interviewer by providing two possibly mutually reinforcing opportunities for interview.

c)   an investigation into the practice of sampling residents in institutions which form part of the NRS universe. In practice, only 0.2% of NRS informants are classified as living in an institutional address, when the figure should be in the region of 1%. This led us to conduct a study of the difficulties of interviewing in student halls of residence, which were indeed found to be great. The main problems were those of getting access to the premises and of conducting the sampling task. Once a student was selected, co-operation levels were found to be unproblematic.

d)   an experiment of lengthening the time in which an NRS assignment must be completed. This experiment, conducted as a split sample test using NRS fieldwork of January to May 1995, was not conclusive. There was no improvement with the test design, which prescribed that re-issues were always to take place four weeks after the original start of the assignment as opposed to between one and three weeks after the original assignment date. We thought that our test design was too rigid and therefore not successful.

e)   a test to establish whether re-issues were more effective if prior contact is made with potential informants by telephone. We used February and March 1996 NRS fieldwork for this test. No notable improvement was seen. One probable reason was that advance telephone calls, similar to advance letters, have a mixed effect. While it works for some, others become alerted and hardened against granting the interview.

f)   a "doorstep investigation" conducted in 1995 and expanded in 1996 in the form of observation studies into aspects of how the NRS interview was introduced, the different situations encountered, the different apparent reasons for co-operation and non-co-operation, including observations on body language and mood as well as external circumstances such as the weather. This led to the conclusion that greater weight should be placed on training of interviewers in "doorstep techniques", and a programme of such training, given by professionals in sales training, is now in force.

g)   a study of a mixed method approach to sampling and interviewing. This 1995 study was funded and conducted privately by RSL. We describe it in greater detail below.

h)   a new study to test the effect of lengthening the assignment. A proposal for this study is at present under discussion. It would involve the lengthening of an assignment from the current basic seven days to three months in which period the field management procedures would include constant dialogue with interviewers, improved call patterns depending on the circumstances, possibilities of advance letters or phone calls, appointment making and similar.

It should be noted that there are no incentives at present for the NRS interview, except a pen as a token of appreciation. Substantial monetary incentives are ruled out for cost reasons but also because they would irrevocably change the research culture in Britain. However, proposals for tests involving prize draws have been made, and of giving magazines as incentives. These are still under discussion.

A Test of The Mixed Method Approach

The mixed method approach is one where different media of data collection - mail, telephone, personal interview - combine to maximise response. The test which RSL privately conducted in September 1995 was inspired by two papers given at previous Symposia. The first was by Richard Lysaker of Audits and Surveys at the Hong Kong Symposium of 1991 entitled "Using Multiple Media to maximise Response Rates". The second was by Julian Baim of MRI at the San Francisco Symposium of 1993 entitled "The Impact of Incentives and 'Multi-Media Approaches' on the Levels of Response Rates and The Quality of Response".

Richard Lysaker showed that the use of mail contacts which were followed by telephone contacts and then followed by personal contacts produced exceptionally high response rates of well over 90% in a series of surveys which were designed as verification checks on behalf of the American ABC. In these surveys, a short household related questionnaire could be answered by either the head of household or the housewife. The sequence of attempts, first by mail, then by telephone and then personal contact was chosen for economic reasons. For the first large bulk of interviews there was the cheaper mail method, for the more hard-to-get informants there was the more expensive telephone method, and for the final but relatively small hard-core there was the most expensive, the personal interview method.

Julian Baim's report showed that the MRI test had no successful outcome, despite high expectations. In our definition, the MRI test was not a full mixed method test as defined above, as the readership interview could only be conducted by the personal method. It was an advance telephone/mail contact test, as follows:

In the MRI test design, households first received a telephone call to screen the respondent and to make an appointment for the main (personal) interview. This was then followed by a card sent by mail to confirm the appointment, followed by a remainder telephone call just prior to the appointment. Finally, the main personal interview was conducted. A US $10 bill was offered as an incentive. (The control group received, as is standard for MRI, a generic advance mail card to the household, and no phone calls, and the incentive was a Swiss Army Knife.) In this test, the experimental design did not perform better than the control design in terms of response rates. In fact, the test results were considerably lower with 49.1% against the 65.1% of the control sample.

Julian Baim offered two explanations for this apparent failure. One was that there had been logistic problems. The more serious one was that the advance phone calls had a negative effect on many potential informants because they had been given an easy opportunity to refuse on the phone, thus halting the process of interviewing them later on.

RSL's experience turned out to be in many ways similar to MRI's. Our test was designed as a full mixed method test and conducted in September 1995 with the purpose of gaining experience with such an approach.

There was the mixed method of data collection even of the readership data - by mail, by telephone (CATI) and by personal interview (CAPI). In addition, we also changed the sampling approach, designed to maximise response. All adult household members at a sampled address were selected for interview. This meant that only one household member, the contact informant, needed to provide all the necessary household data, while the other household members only needed to provide the readership and other personal data.

In our design, the sequence of attempts to interview was as follows:

MAIN TEST

Stage 1. All issued addresses

● Conduct a CAPI (personal) interview with a contact person - for household data.

● Conduct a CAPI readership and personal data interview with at least one household member aged 15+. This could be the contact person or another person. Priority was set for young male members of the family if available.

● Make a placement of self-completion readership and personal data questionnaires for all other household members aged 15+.

Stage 2. Non-productive addresses of stage 1

● Conduct a CATI (telephone) interview with one household member - for household data.

● Post self-completion readership and personal data questionnaire to all household members aged 15+.

Stage 3. All issued addresses

● Collect in person the placed (at stage 1) or posted (at stage 2) self-completion questionnaires. At this stage, incentives (vouchers of a department store) were handed over. For non-completed self-completion questionnaires, return envelopes were placed for mailing them back to the office.

● Make further attempts to interview unproductive households of stages 1 and 2.

For this test, we sampled 110 random addresses in five postcode sectors spread over the country. A minimum of three calls was prescribed at each of the three stages above. However, these calls were to be conducted within three days in relation to stage 1, and within two days at each of the following two stages. This proved, with hindsight, to be a very constraining time-frame. Our desire was not to extend the time-frame beyond what is currently the fieldwork timing of an NRS assignment. The fieldwork dates were 14 to 25 September 1995.

A further feature of our pilot was the design of a parallel test of 110 addresses for which the sequence of attempts was changed. In this parallel test, we started with telephone interviews, followed by mail and then by personal attempts, as follows:

PARALLEL TEST

Stage 1. All issued addresses

● Conduct a CATI (telephone) interview with a contact person - for household data.

● Post self-completion readership and personal data questionnaires to all household members aged 15+.

Stage 2. All issued addresses

● Collect in person completed self-completion questionnaires. Hand out incentives (voucher of a department store). Place return envelopes for mailing back questionnaires, if necessary.

● Make further personal calls on unproductive addresses of stage 1 using CAPI and/or self-completion.

The outcome, in response rate terms, was at first sight disappointing. For the main test, the final response rate was calculated to be 62%, virtually the same as the current average national response rate of the NRS. For the parallel test, it was a disappointing 42%. In summary, the response rates were calculated as follows:

496

	Main Test	Parallel Test
A) Productive CAPI or CATI contact interview at any stage	71%	57%
B) Productive CAPI or self-completion readership and personal data interview at any stage	87%	74%
C) Overall response rate (product A x B)	62%	42%

At a second glance, however, we realised that the rather disappointing result came mainly from the contact stages, as may be seen from the table above.

The reason was obvious to us - we had not allowed enough fieldwork time for fully exploiting the contacting stage. The three days we had prescribed were very restrictive. If we had allowed more time, say six days for stage 1, and if this would have resulted in a contact interview response rate in the region of, say 80% at the main test, and if the second stage response rate would have been the same as that achieved in the test, then the overall response rate for the main test could have been expected to be in the region of 70%. This would have been a marked improvement on the current NRS response rate.

So, our conclusions were: (a) the mixed method approach shows promise, (b) we would need to allow ample time for all fieldwork stages, and (c) the sequence of attempts which puts the personal contact first seems to have greater effect than the sequence with telephone contacts first.

The mixed method approach to readership interviewing relies on the assumption that the modes of interview employed will be acceptable; that is to say that readership interviews whether by personal interview or by self-completion - as employed in our test - will be regarded as interchangeable. This is by no means certain, people will need to be convinced of this. From the outset however, we did not regard the telephone method as a CAPI compatible form of readership data collection because of the absence of visual stimuli.

For the main test, about two thirds of readership interviews were conducted by self-completion, one third by CAPI. Of course, sample sizes were very small. But even so, it was heartening to see that the general magnitude of the overall readership results was found not too dissimilar between the two methods, though the self-completion interviews tended to produce results below those of the CAPI interviews. This may be seen from the following table:

(Base)	Main Test		For comparison: NRS 1995
	CAPI interviews (65) %	Self-completion interviews (105) %	(38,102) %
All titles gross RPY	1,471	1,254	1,259
All titles gross AIR	561	422	545

There are over 260 titles on the NRS, and the above findings are expressed as their gross percentage penetrations.

Conclusion

None of the experiments which were briefly described in this paper brought the NRS response rates potentially nearer the desired levels of, say, 70% and more - except that we believe that the mixed method approach to sampling and data collection holds some promise. We also have hopes that the planned test of extending the length of assignment will point to potentially marked improvements. These experiments and investigations have however helped us understand the problems better, and have shown where amendments to our procedures could be fruitful. Maintaining the current levels of response is nearly as much a challenge as improving them.