# NRPS REVIEW

## Outline of Proposed Design Recommendations

**Prepared for:**

**Passenger Focus**
**Fleetbank House**
**2-6 Salisbury Square**
**London**
**EC4Y 8JX**


**Prepared by:**

**Roberts-Miller Associates**
**PO Box 104**
**Dorking**
**Surrey**
**RH5 6YN**
**Tel: 01306-741368**
**E-mail: mail@roberts-miller.co.uk**


**Richard Roberts-Miller, FMRS**
**Gary Bennett, FRSS, MMRS**

***2015***

Contents Page

# EXECUTIVE SUMMARY

1. Passenger Focus commissioned RMA to carry out a technical review of the methodology of the NRPS. The project team consisted of Richard Roberts-Miller FMRS, and Gary Bennett FRSS, MMRS.

2. The GB rail network is complex, and it is a challenging goal to design and operate a survey that is reliably representative of all passenger journeys.

3. There were 1.59 billion rail passenger journeys in 2013-14 departing at the rate of 30.58 million per week, carried by 23 TOCs, from the 2,535 stations in the network.

4. It is quite ambitious for a survey of only 1560 fieldwork shifts at only 526 stations (producing a total of 29,524 interviews) to aim to produce a representative, detailed picture of the whole network.

5. The NRPS does so up to a point, i.e. to the extent required by the objectives, which are (in brief) to measure, twice a year, passengers' satisfaction with their journeys (on 35 separate parameters), for individual TOCs and for the GB network as a whole, and to detect changes over time.

6. However, many stakeholders have for some years been seeking to use the data to look at smaller (sub-TOC) segments of the survey.

7. The adoption of the sub-TOC building block (BB) approach, which arose mainly to facilitate year-over-year comparisons when TOC boundaries change, has further encouraged sub-TOC level analysis, but has also increased the complexity of the sampling and weighting.

8. NRPS sampling and weighting now involves an extremely complicated shift allocation process. It aims to balance, within a sample of stations (some of which appear in several BBs), the conflicting goals of achieving: (i) a representative sample (one for each BB), (ii) required TOC sample size targets, (iii) a representative sample both for each TOC and (iv) nationally, (v) without using excessive weighting.

9. The only part of this process that is straightforward is the initial "Probability Proportional to Size" (PPS) sample of stations. The next stage, shift allocation, is effectively a compromise that requires a significant input of judgment and experience, and is far more complicated than is described in the Detailed Technical Overview.

10. In fact, the PPS sample of stations appears to be the starting point for a circular process of shift allocation, based on previous returns, that tries to balance the above competing goals. The process is, we believe, less than ideal.

11. While we are impressed with the ingenuity with which the shift allocation processes have been developed and are applied, and with the apparent stability of the customer satisfaction data that is produced, we are convinced that a more transparent, more efficient and more objective approach is available, and should now be adopted.

## IMPROVED SAMPLE DESIGN

12. Whilst we understand how the conflicting objectives have pushed the sampling in the direction it has taken, we believe that not dealing with it now could potentially lead to greater problems in the future.

    We conclude that the sampling process needs to be simplified and the sample design improved, and we have recommended a solution that we believe achieves both goals.

13. We believe the NRPS now needs two elements (and if it were starting today we believe this is how it would be done):

    **A main NRPS survey** of about 20,000 interviews: based on a simple but high quality, quasi-random, PPS survey of stations and shifts using ORR and NRTS data, using at-station recruitment, and designed to produce a stable, representative sample of all GB passengers that is also representative of all parts of the network.

    **Booster surveys** (one for each BB) totalling about 10,000 interviews: based on high quality, quasi-random, PPS surveys of stations and shifts as above (but using on-train as well as at-station fieldwork, as appropriate), designed to produce a stable, representative sample of passengers at each BB (and thus, when aggregated, at each TOC).

    Fieldwork for both the main and booster surveys could be conducted at the same time. The results would be added together to produce an NRPS report (with approximately the same 30,000 sample size) that meets all the key goals in a much simpler and more objective way than at present.

    There will obviously be concerns about comparability that must be addressed, but we see no *a priori* reason to believe there will be discontinuities in the data (as effectively the new design aims to build the same survey as before, but from two logically separate parts).

14. We conclude that there would be fewer design compromises than at present; weighting would be simpler and lighter-touch, so effective sample size and precision would probably be improved, as would stability. Other advantages would be:

a) It would be much more easily documentable.
b) It would be more transparent.
c) It would be easier to execute.
d) It would be easier to monitor its effectiveness.
e) The extra cost of each booster could be identified.

15. We believe that (after any one-off set up or pilot costs) ongoing costs could possibly be lower as (a) the method requires far less manual intervention and (b) the use of PPS sampling at shift level is expected to improve response rates throughout the bottom 50% of stations by passenger volume (and also improve the representativeness of the sample overall).

## RESPONSE RATE

16. Another conclusion of this review is that the questionnaire is now **too long** and the response rate is now **too low**:

| NRPS Wave | Number of Questions | Response Rate |
|---|---|---|
| Wave 10 (Spring 2004) | 49 | 41.2% |
| Wave 30 (Spring 2014) | 67 | 29.2% * |

We believe that it is essential that the main NRPS questionnaire is reduced in length until it is closer to Wave 10 dimensions, and until the decline in response is reversed.

Also, no blocks of extra questions should be added in future, as these are known to damage the overall response rate. Instead all such occasional or minority topics should be covered by follow-up surveys (on-line or by phone etc.) or by completely separate surveys.

17. If the response rate could be raised again to 41.2%, the NRPS sample size would increase from about 30,000 to about 40,000 (or alternatively the cost could be reduced).

* The figure of 30.7% quoted in the Wave 30 Overview report is not comparable with Wave 10 data as it includes 5,077 questionnaires all or most of which were distributed on routes not covered in 2004 (Heathrow Express, London Overground etc.) and/or those employing non-standard sampling and fieldwork methods. See footnote on page 15.

## OTHER ISSUES

18. We also make recommendations regarding monitoring of sampling fractions, and recruitment distribution by station, building block, and time of day. We believe that the quality of fieldwork will be improved by these measures.

Other recommendations are that the weighting process should be split into design weighting and non-response weighting. The new sample design, and the use of detailed universe data for rail journeys plus better information on recruitment patterns, will facilitate this.

Among the other conclusions: we regard face-to-face recruitment at stations, followed by a self-completion questionnaire, as the most cost-effective approach for the main NRS sample (as at present).

Also, face-to-face recruitment on trains, followed by a self-completion questionnaire (but without on-train collection), is an acceptable approach for non-franchised TOCs or as a cost-effective booster method for rural routes and smaller stations. Both the at-station and on-train approaches would benefit from the detailed changes to fieldwork procedure and documentation that we have recommended.

We have also recommended detailed improvements to the NRPS technical documentation re: sampling, fieldwork, response rates etc.

## PRIORITIES

19. In terms of priorities (i.e. importance) it is essential to shorten the questionnaire to improve the response rate. A better response rate would obviously increase the sample size (or potentially lower the survey cost). It would also:

- reduce the need for corrective weighting
- thus increase the effective sample size
- thus reduce margins of error
- and increase confidence in the NRPS results

This process could start immediately.

20. In our opinion it is equally important to migrate the NRPS sample away from the present design to an improved (but simpler) design, but this cannot be done overnight.

Instead, we recommend that the proposed new "main" survey design is piloted on a small scale alongside the main NRPS wave, with a view to implementing the new design in the following wave or the wave after that.

21. Meanwhile, all or most of the recommended improvements and changes to the fieldwork record-keeping, procedures, instructions and documentation, and response rate reporting etc., could be implemented, as none are dependent on the new sampling process - in fact the reverse is the case:

    We believe these improvements should not be delayed, as the fieldwork record-keeping and response reporting data will be very helpful in comparing the detailed results of the new design main pilot with the previous method (and in comparing in more detail the results of the at-station and on-train interviews), all of which will be important input into the detailed planning of sampling and weighting for the first full NRPS wave using the new, improved design.

22. For 15 years, the NRPS has been important in contributing to improving the performance of rail operations and thus increasing passenger satisfaction. We believe that the enhancements we have suggested will help the NRPS maintain that role in the future, and we thank Passenger Focus for the opportunity to support its objectives.

# 1. BACKGROUND

The National Rail Passenger Survey (NRPS) was introduced by the Strategic Rail Authority in 1999[1]. The survey was designed to provide reliable data on trends in satisfaction with each train operating company (TOC) that provided franchised services to the SRA, and for the GB heavy rail network as a whole.

Passenger Focus, an independent public body set up by the Government to protect the interests of Britain's rail passengers[2], became responsible for the survey in 2005. Since then, Passenger Focus has initiated various projects to identify if improvements can be made to the survey's methodology, so that it could be used more effectively to drive improvements to train services for the benefit of passengers, while bearing in mind the value of maintaining the continuity of the data set. The survey has been enhanced gradually as a result, but it has also become more complex.

Over the same period, use of the survey by stakeholders has grown more intensive. Also, NRPS-based performance targets increasingly form part of TOC franchising contracts. The NRPS results are now classed as Official Statistics. Passenger Focus wishes to ensure that the NRPS remains fit for purpose both today and into the future, and has commissioned Roberts-Miller Associates (RMA) to conduct this review.

The review work was carried out by Richard Roberts-Miller and Gary Bennett. Both are experienced members of the Market Research Society, and both have previously worked successfully on a number of similar technical reviews in this field. There are further details of their backgrounds in Appendix A.

The approach taken was to meet with key personnel at Passenger Focus and at the current NRPS contractors, BDRC Continental; to collect documents, information and data (in greater detail than is published in the NRPS Detailed Technical Survey Overview volumes) re: the critical elements of NRPS sampling, weighting, fieldwork and analysis. There were numerous further discussions with the contractors on points of detail, and we also collected further documentation and data from elsewhere. As well as examining this information as thoroughly as was possible in the time allowed, further diagnostic cross-analysis of the NRPS and other databases was carried out.

This document includes an outline of our recommendations re: aspects of the future design and operation of the NRPS, plus some of the key background reasoning.

We would like to thank all those individuals whose help made it possible for us to complete this project.

[1] Until autumn 2013 it was known as the National Passenger Survey (NPS).
[2] Now also bus, coach and tram passengers.

## 2. OBJECTIVES OF THE NRPS

The first key objective of the NRPS is:

**A: to measure, on a consistent basis, passengers' satisfaction with their rail journeys, so that the performance of individual, franchised train operating companies (TOCs) can be compared over time.**

The second key design objective is that:

**B: NRPS data for the individual TOCs can be amalgamated so as to be able to measure rail passengers' satisfaction over time for GB overall.**

Objective **B** of the NRPS requires that the survey design for each individual TOC is suitable for such amalgamation on a like-for-like basis. The simplest way to achieve this would be to use an identical survey design for each TOC.

However, objective **A** requires that the sample size for each TOC is sufficiently large that its performance changes can be detected. Some TOCs are large and some are small; some TOCs have simple networks are some have complex ones. This means that the required sample sizes will inevitably vary between TOCs. However, any planned differences in their sample sizes can be adjusted at the weighting stage to produce correctly representative national data.[3]

It is important to note that the NRPS was never intended to provide data at the level of Government Office Regions, or PTEs, or routes, or individual stations. The design priorities remain the production of representative data for individual TOCs, and for the national network as a whole.[4]

Also, it is important to remember that the NRPS intentionally produces a representative sample of all rail journeys[5]. This means, for example, that passengers who travel five times a week will be five times more likely to be sampled than will those who travel once a week.[6]

[3] Although, if the weights differ very widely, statistical problems may arise.
[4] The basic NRPS design happens to able to provide some useable data below TOC level (without e.g. booster samples) but only in a limited number of situations, and usually only with caveats on the interpretation).
[5] More precisely, it provides a sample of rail journey legs. Roughly one in six or seven rail journeys requires a change of trains, and many of those involve more than one TOC. Therefore the NRPS must measure satisfaction with a specific journey leg (i.e. departing from a specific station and travelling with a specific TOC). The NRPS does also collect information on whether there was another rail leg before or after the one examined in detail (see Q43 and Q47 of the Spring 2014 NRPS).
[6] Even so, the NRPS design does still allow the views of regular and occasional travellers to be looked at separately if that is required.

# 3. SUMMARY OF RECOMMENDATIONS

The conclusions and recommendations expressed below are based on our study of all the documentation, data and information made available to us, and include judgments and opinions evaluated in the light of our combined experience of working on the design, execution and interpretation of many large quantitative surveys (including previous reviews of the NRPS, and other rail, bus and travel surveys).

## 3.1 FIELDWORK METHODS

### 3.1.1 Standard NRPS At-Station Recruitment + Self-Completion

Given the clear objectives of the NRPS (see Section 2), we believe that no other existing or foreseeable fieldwork method (on-line, telephone, face-to-face) appears likely to meet the objectives of the NRPS as cost-effectively the present standard approach.

**We still regard the standard NRPS fieldwork method, of face-to-face at-station recruitment, followed by a postal self-completion questionnaire, as the single most cost-effective approach available for the established purposes of the NRPS.**

### 3.1.2 Alternative Survey Methods - Electronic

It has been suggested that perhaps the NRPS should use an on-line fieldwork method. However, in our view all the current alternatives are much less satisfactory. The **on-line method** may be economical but there is no way of which we are aware that a satisfactory sample of journeys could be achieved using the on-line approach alone, and even if there were, we do not believe that a questionnaire as long and complex as the NRPS could achieve a satisfactory response on-line (still less using a mobile phone); **telephone interviews** would face similar sampling problems (and would also be much more expensive); **face-to-face interviewing** (as opposed to face-to-face questionnaire distribution) would raise the cost of the NPS by an order of magnitude, and would be very difficult to carry out at peak times.[7]

[7] Areas in which we do believe such other methods could be helpful are: (a) as an optional alternative offered to NRPS respondents who are recruited in the standard way, but who would prefer to complete the survey on-line (this has been piloted and evidently only a minority choose this option, but the numbers may grow), and (b) for follow-up surveys to respondents who have already completed a standard NRPS survey. See Section 3.6.

### 3.1.3 Alternative Survey Methods - On-Train Recruitment

The one area of weakness of the standard NRPS "at-station" question distribution method is when recruiting at small stations with few passengers departing per hour, where a fieldwork shift may yield very few interviews (fortunately, these represent a small proportion of the NRPS sample). For this reason, it was decided some time ago that, on the very small Isle of Wight "Island Line", it was more cost-effective for questionnaires to be distributed on board the train. We regarded this as acceptable, as being a very minor exception from the standard procedure.

However, since then, on-train distribution has been used on a larger proportion of NRPS interviews. In this method, interviewers distribute questionnaires on one train and then distribute questionnaires on another train travelling in the opposite direction, and this round trip may be repeated more than once. Whatever the merits of this on-train approach, it raises questions about how comparable these on-train results are with the results from standard NRPS at-station recruitment interviews - and the obvious sampling and weighting issues are not addressed in the NRPS Technical Survey Overview.

Also, on-train questionnaires are now being collected by interviewers before passengers have completed their journey. We do not know how many (we understand that this number is not routinely recorded) but we assume it is probably most of the on-train respondents, as the procedure to collect them on board is included in the relevant interviewer's instructions.

The explicit intention of the NRPS is that respondents complete the questionnaire after they have completed their entire journey. As many NRPS questions (for example on Wave 30: Qs 35, 47, 51, 52 & 53) cannot be validly answered while on board the train, those on-train replies are either absent (or invalid).

Also, on-train recruitment followed by postal self-completion is less cost-effective (i.e. has a lower response rate) than the standard at-station method at medium/large and busy stations (i.e. the bulk of the NRPS sample).

In our judgement, on-train distribution is acceptable as a limited substitute for the standard at-station procedure, e.g. on rural routes where at-station distribution is demonstrably not cost-effective, provided that:

a) questionnaires are not collected on board train (unless those questionnaires do not form part of the official NRPS franchised TOC survey, in which case the issue may be less critical),

b) adequate thought is given to how the round-trip on-train samples can be validly integrated with the standard at-station sample, and

c) adequate documentation is provided to explain how the on-train sample design, weighting design and fieldwork instructions vary from the NRPS standard, and how the results were integrated.

On this basis, within the proposed new Simplified Two-Part NRPS Survey Design (see Section 3.4) the on-train method could be helpful for some of the proposed NRPS Booster Survey samples, but should not be employed anywhere on the proposed Main NRPS survey sample (see Section 3.5).

### 3.1.4 Other Fieldwork Issues

A simple change which will facilitate the alignment of the at-station and on-train samples, and provide a number of other benefits, is to alter the contact sheets (for recruitment) such that the station of boarding, time of recruitment and building block are recorded for all recruits. This will:

- Enable the calculation of the "probability of being sampled" for all station/time-of-day combination, including for on-train interviews;

- Enable better management and efficiency of the fieldwork process, by giving the agency the ability to monitor questionnaire distribution by station[8], time-of-day and building block.

- Facilitate the separation of design weighting and non-response weighting for all interviews (see Section 3.7.6).

Also, we recommend that it is advisable to:

- Review and adjust the algorithm for determining the number of questionnaires to be distributed per shift. Our recommendation is that this number is as close to some nominal maximum as possible to ensure that the sampling fraction is consistent as possible across different station/time band strata.

The setting of guidelines for and monitoring fieldwork activity is an area that we believe needs to be improved. We believe that more detailed written guidance should be provided to interviewers. Feedback from the improved contact sheets should also be used to develop tighter procedures for the distribution of questionnaires by building block for stations that appear in several building blocks. This will also provide checks on the methodology employed for on-train recruitment, ensuring that a good distribution is obtained by station, across both (or all) legs of the journey.

[8] For on-train as well as at-station interviews.

## 3.2 SAMPLE SIZE AND SURVEY FREQUENCY

Issues have often been raised in the past regarding the inability of the NRPS to meet the goals that some stakeholders desire, e.g. to deliver greater accuracy generally, or precise results at station level or at route level, or to have more interviews at more of the smaller stations, or more frequent survey waves, or (perhaps the most frequent suggestion) just to have a larger sample size overall. However, we still believe that:

**The present sample size (around 30,000 interviews per wave) is sufficient to meet the stated objectives of the NRPS, provided the fieldwork, sampling and weighting are designed and executed sufficiently well.**

If that is the case, and the objectives of the NRPS remain the same, it is very difficult to justify moving in any of the directions that are suggested, as the costs would be huge.

For example, to merely halve margins of error on NRPS results would require four times the present sample size (and costs would rise almost as much).

To increase the frequency from two to four waves per annum would require annual sample to double (and costs would rise almost as much).[9]

Producing representative data individually for an average station is even more dramatically beyond the scope of the NRPS[10]. The NRPS sample is designed so that the weighted results for each building block, or TOC, are representative of the entire passenger volume of that building block or TOC (and also in aggregate and for the GB network as a whole). This is achieved despite a sample in which many individual smaller stations may have only one interview shift (or none) in each wave.

To add just one interview shift to every station in the GB that does not currently have one, would roughly double the cost of the NRPS today, and it would produce only about 20 interviews per station (and all of them in the same 3 hour period, and on just one day of the week - which would not produce much useful information).

To be able to have 20 shifts at every station (which is roughly what would be the bare minimum necessary needed to look at individual station profiles) would multiply the NRPS cost by 20 times.

[9] Unless sample size per wave were halved. However that would make year over year comparisons less robust than they are now, so you would have to wait until the next quarterly wave anyway before being as confident as you were last year that any observed change in the result was significant. Similar issues apply to spreading the sample out across the whole year.

[10] It is feasible that for the very few very large stations (e.g. those which already have at least about 20 shifts), the NRPS design could be revised to produce slightly "better" data (i.e. more representative in profile for that station individually) but with the current NRPS design, the necessary changes would tend to clash with one of the main objectives: to produce representative data for each building block or TOC.

## 3.3 QUESTIONNAIRE AND RESPONSE RATE

Two conclusions re: the questionnaire have not changed since our earlier Reviews. We could suggest very small changes to clarify a few individual questions, but broadly: **In our view the questionnaire covers all the topics necessary for its objectives (and more). Also, the semantic scales are effective and should not be changed.**

However, we have also previously expressed concern that the NRPS questionnaire has grown very long, and we have drawn attention to the falling response rate and urged that the **questionnaire length should be limited**.

We believe the situation is now critical. As the NRPS questionnaire has grown from 49 questions to 67, the response rate has collapsed from 41.2% to 29.2%:

| NRPS Wave | Number of Questions | Response Rate |
|---|---|---|
| Wave 10 (Spring 2004) | 49 | 41.2% |
| Wave 30 (Spring 2014) | 67 | 29.2% * |

We believe now that it is essential that the NRPS questionnaire:

(a) is reduced in length until it is close to Wave 10 dimensions
(b) is also redesigned to look less intimidating, and
(c) no extra blocks of questions should be added ever in future.

The questionnaire content should be restricted to the minimum that it is essential to include in every wave for the key purposes of the NRPS (this will help the stability of the results too). Any other occasional or minority topics should be covered by follow-up surveys instead (see Section 3.6), or by completely separate surveys.

The NRPS may well have been affected to some extent by the fact that response rates have been gradually falling in UK market research generally, but going from 49 to 67 questions - roughly a 37% increase in interview length (on an already long questionnaire) is very likely to be the major factor in this case. As a side effect, a shorter questionnaire could allow interviewers to carry a larger supply, which would certainly increase potential response numbers. Also, if the response rate can be driven up once again to 41.2%, the current NRPS sample size of 30,000 per wave would rise to 40,000 (or alternatively the cost of the survey could be reduced accordingly).

If nothing is done, the response rate may well continue to fall, which would risk undermining the credibility of the whole NRPS.

* The figure of 29.2% is taken from "9070SamplingPlan300614" data supplied by BDRC Continental. The figure of 30.7% quoted in the Wave 30 Overview report is not comparable with Wave 10 data as it includes 5,077 questionnaires which were distributed on routes not covered in 2004 (Heathrow Express, London Overground etc.) and/or those distributed and collected on board train, which is not the established fieldwork method as (a) it requires a non-standard sampling design, and (b) certain questions will have invalid or missing answers. The 29.2% figure excludes these non-standard questionnaires.

## 3.4 SAMPLE DESIGN - STATIONS AND SHIFTS

The theoretical process described in the Technical Overview for the PPS (Probability Proportional to Size) sampling of fieldwork shifts for specific stations is robust. However, the documented approach does not provide full detail of the current practice. For a number of reasons (e.g. varying TOC sample size targets, and the increased use of building blocks) the process that is currently employed for time of day shift selection has evolved to become very complex and subjective. Although the process employed seems to have reasonable success in balancing the competing objectives of the design, it falls short of the PPS ideal. We have recommended an improved approach which is more consistent with the original principles of the sample design.

We recommend once again that NRTS data on weekday journey volumes by time of day should be employed to create a PPS sample of shifts for each station selected.[11]

In our view this would have various beneficial effects:

1) The sample would reflect the best available external evidence.
2) The sampling procedure would be more transparent and objective.
3) The sampling procedure would be more efficient at small stations.
4) The sampling procedure would be better understood.

The shifts sampled currently yield very low numbers of interviews at some small stations (occasionally nil). Using PPS sampling to allocate shifts by time of day would be more likely to place shifts where and when the passengers were present in larger numbers, and thus would be expected to increase the potential number of interviews achieved (hence effect #3 above).

[11] Weekends should also use the PPS method, but in the absence of NRTS weekend data, the best available source would be TOC estimates, as is currently used for the weekday:weekend split.

## 3.5 SIMPLIFIED AND IMPROVED SAMPLE DESIGN

If the NRPS were started now, it would very probably employ the same fieldwork method (face-to-face recruitment at stations, followed by a postal self-completion questionnaire[12]), but it is very unlikely that it would employ the present sample design, as a number of factors have changed since 1999 (including varying TOC sample size targets, increased use of building blocks, the demands placed upon it, and the availability of NRTS data). The overriding need for the NRPS to provide stable, credible measurements of customer satisfaction, has tended to discourage any visible changes to the design. However, we believe that the process is now so complex that some revisions to the design are essential. We recommend that the NRPS sample design is split into two parts, each of which is much simpler than the present approach, but which together would achieve the same result - and in a clearer, more objective, understandable, measurable and reproducible way:

**Stage One: National NRPS Sample**

This would be a high quality, stratified, quasi-random, PPS survey with the objective of producing a stable, representative sample of all GB heavy rail stations and shifts at the national level, and using at-station recruitment only. Within that it would be designed to be representative of all major regions, of all station size bands, and of all station types (e.g. terminus, interchange, non-interchange) and route types (e.g. long-distance, commuter, rural) and all times of day, using ORR-based PPS station selection and NRTS-based PPS shift allocation. This would have a base size of about 20,000 interviews.[13] Few changes to the design should ever be necessary (enhancing stability going forward). Little design weighting would be necessary (as there are zero design factors to correct for) so the effective sample size per 1,000 interviews will be higher (i.e. better) than that of the current NRPS.

**Stage Two: NRPS Booster Samples**

Wherever the Stage 1 sample falls short of the agreed TOC and building block targets (which will be fairly simple to estimate), a booster sample will be added to that TOC or building block, to rectify that shortfall. Each booster sample will be a high quality, quasi-random, PPS survey with the objective of producing a representative sample of that TOC or building block, of similar design to Stage 1, but using at-station or on-train fieldwork or both, as appropriate.[14]

---

[12] As is now also used by BPS and TPS.

[13] This is an estimate: the exact size can be calculated relatively easily depending on the needs of Stage 2, which in turn will depend on the extent to which the agreed TOC and building block targets depart from their natural occurrence in Stage 1. The goal should be for franchised TOC interviews at Stages 1 & 2 to add up to about the same total as the present NRPS, i.e. about 30,000 per wave (optional extra boosts to sample size beyond the agreed levels would be additional).

[14] This same Stage 2 approach could also be used for any other add-on samples considered, such as non-franchised TOCs, individual stations or areas (e.g PTEs) or route types.

Because it will be simple to estimate roughly the Stage 2 sample size requirements in advance, the fieldwork for Stages 1 and 2 could be conducted simultaneously. Only in the event of an unexpected shortfall would extra *ad hoc* boosters be required at the end of fieldwork (as occurs today).

**Merging the Results**

The complexity of the current NRPS operation has evolved from the increasing difficulty of combining and balancing all the different (and conflicting) requirements of the two above stages into one single design.

In contrast, Stage 1, and each of the Stage 2 components, are all now individually simpler designs.

Merging Stage 2 booster survey results with Main Stage 1 survey results will be relatively straightforward, as Stage 1 is designed to provide a suitably solid, unchanging foundation and framework. Stability should be increased. The results would be expected to be superior to the present design at both the national level and the boosted local level. The other advantages of this approach are that, being simpler:

a) It would be more easily documentable.
b) It would be more transparent.
c) It would be easier to execute.
d) It would be easier to monitor its effectiveness.
e) The extra cost of each booster sample could be precisely identified.

Also, any number and scale of boosters can be added (or dropped) without altering the main (Stage 1) NRPS design or execution in any way (which is not true now), in which case the new solution will also be more stable overall.

That the present situation has evolved is not surprising, because the starting point was in effect to produce a number surveys, each representative of one TOC, that could be added together to produce a grand total that was representative of the whole GB network. Since then the proliferation of building blocks (which was logical and justified) and the consequent increase in their overlaps, have made it increasingly difficult and complex to balance the different and competing objectives within one design.

By separating the two objectives (TOC samples and GB sample), and reversing their conceptual order, the above process aims to achieve the same goals more simply and effectively.

Not least, this design will greatly facilitate any weighting that is required.

## 3.6 FOLLOW-UP SURVEYS

Although, for sound reasons, we do not regard on-line as a viable way to operate the whole of the NRPS survey, we do support the idea that NRPS respondents, who have received a self-completion questionnaire in the usual way, should be offered the option of responding on-line if they wish to. This approach has been piloted by Passenger Focus, and although the numbers choosing this option are currently low, they are likely to increase over time.

More important in our view is the potential of online (and other methods such as telephone and postal) to help reduce the length of the main NRPS questionnaire:

In the past from time to time, blocks of questions have been added to NRPS that relate to minority questions, for example the opinions of people who have had reason to call upon British Transport Police, or those who are suffering from a disability. The resulting information is very important to those who are dealing with those issues, but we would argue that blocks of minority questions should not be incorporated into the main NRPS survey (making the questionnaire significantly longer and more intimidating for everyone, while asking questions that are of interest or relevance to relatively few).

Instead we suggest that minority respondents are asked just two questions: for example, whether they have experienced a security problem on the railway, or whether they are affected by any disability - followed by "would you be willing to answer some further questions on the subject?".

The follow-up could be by phone, or post, or by email (if they are willing to give their relevant contact details). The main NRPS data for each individual could be linked to their follow-up replies so the NRPS questions would not need to be asked again. This could allow a separate (and probably longer) questionnaire on the chosen topic (without prejudicing the NRPS response rate).

Adding a question to all NRPS questionnaires, inviting respondents to provide their email address, would allow Passenger Focus to continue to build a very large database of potential interviewees for future surveys that could be accessed quickly and economically by email for an on-line survey, whether aimed at minority groups, or other topics that need to be looked at only occasionally.

Given that 97% of weekday passengers have access to the internet, surely this is an opportunity to be pursued, in preference to squeezing extra questions into an already very long NRPS questionnaire, and prejudicing the main NRPS itself.

Some other groups of questions are not looking at issues affecting minority groups, but are covering more general questions, such as asking how time was spent on the train journey (reading a book? or newspaper? working? playing a game on a laptop? etc.). The answers may be of interest and value to those

who ask them, but it seems to us very unlikely that they require the precision of a 30,000 base size sample to achieve their goals (especially when its addition to the questionnaire will prejudice the response to the NRPS as a whole[15]).

Such topics too could be handled through a follow-up survey by phone, or post, or by email.

If it is ever possible to justify attaching any such block of questions to the NRPS questionnaire, we very much doubt that it would require the whole 30,000. If every tenth NRPS questionnaire contained the extra section[16], that could still yield about 2,500 completed copies (but not 3,000, because - based on experience - the response rate on the longer version NRPS questionnaires would be expected inevitably to be lower than for the rest).

[15] Anyone standing on a rush hour commuter train could well find such questions annoying.
[16] Distributed evenly throughout GB to avoid producing regional variances.

## 3.7. OTHER TOPICS

### 3.7.1 Data Sources Used to Compile Sampling Plan

The recommended approach is to use ORR data for determining station size (as at present) and NRTS data for station time of day profiles.

### 3.7.2 Spread of Shifts Across Days and Hours

This would be greatly improved by using the recommended NRTS-based PPS shift sampling.

### 3.7.3 Balance of Outward and Return Journeys

This would be greatly improved by using the recommended NRTS-based PPS shift sampling and the simplified sample design.

### 3.7.4 Sample Sizes Per TOC

The main technical problems arising from this issue are resolved by the new sample design. The problem of deciding what size Passenger Focus should provide for each TOC requires policy decisions outside our remit - but with the proposed new sample design whatever TOC sample sizes are selected, it would require simple design weights for that TOC's data, and would in no way affect the rest of the NRPS sampling process (which it does now).

### 3.7.5 Timing of Fieldwork

Changing field dates would reduce backwards comparability.

### 3.7.6 Weighting

Our main recommendation is to separate the design and non-response components of the weighting. This is currently possible for the at-station sample, and the proposed change to the contact sheets (Section 3.1.4) will facilitate this for the on-train element and thus for the whole sample. This change will more clearly demonstrate the impact of survey design decisions on weighting efficiency whilst also facilitating a much clearer picture of which variables contribute to differential response rates. We believe that this approach along with the proposed sample design (Section 3.5) will result in "lighter-touch" weighting, with the benefit of a larger effective sample size (and thus narrower margins of error).

### 3.7.7 Reporting

Our main recommendation re: reporting is that the "Detailed Technical Overview Report" should contain more detail on key aspects of NRPS survey design and execution, including: the sampling procedures, the fieldwork procedures, recruitment information, the response rates by category (i.e. what proportion of questionnaires issued are completed and returned, analysed by category), and (at least from time to time) the refusal rates (i.e. the proportion of passengers who refuse to accept a questionnaire).

## 4. PRIORITIES & TIMING

In terms of priorities (i.e. importance) we believe it is essential to shorten the questionnaire to improve the response rate. A better response rate would obviously increase the sample size (or potentially lower the survey cost as it would require less fieldwork to achieve the same sample size). It would also:

- reduce the need for corrective weighting
- thus increase the effective sample size
- thus reduce margins of error
- and increase confidence in the NRPS results

This process could and should start immediately.

At the same time, work should go ahead on trying to make the design of the NRPS questionnaire look less intimidating, as a further aid to response.

In our opinion it is equally important to migrate the NRPS sample away from the present design to an improved (and simpler) design, but this cannot be done overnight. In the interests of consistency of NRPS results, some pilot work will definitely be necessary, and it may also be advisable to phase in the changes (e.g. running the two designs in parallel for one wave, perhaps each with a target of 15,000 interviews).

We recommend that the proposed new "main" survey design is piloted on a small scale alongside the next NRPS wave, with a view to implementing the new design after that (in a single wave, or in phases over two waves).

Meanwhile, all of the proposed changes to the fieldwork recruitment record-keeping, procedures, instructions and documentation, and response rate reporting etc., could be implemented in Spring 2015, as none are dependent on the new sampling process being in place - in fact the reverse is the case:

The data from fieldwork record-keeping and response reporting will be very helpful in comparing the results of the at-station and on-train interviews, and in comparing the results of the new design main pilot with the previous method, all of which will provide important input into the detailed planning of sampling and weighting for the first full NRPS wave using the new, improved design.

# APPENDICES

## Appendix A - The Project Team

The project team comprised:

Richard Roberts-Miller, FMRS
Gary Bennett, FRSS, MMRS

The following pages set out their background and experience.

# Richard Roberts-Miller, FMRS

Richard Roberts-Miller established RMA in 1989, as a management, marketing and research consultancy specialising in travel, transport and tourism. Since then, RMA has carried out a wide range of research and consultancy projects for national tourist boards, tour operators, airline, cruise line and railway companies, hotel groups, cottage and villa companies, as well as many projects in other sectors such as finance and education. RMA has conducted reviews of the NRPS (twice), BPSS, BMTS and BPS (each carried out for Passenger Focus) and a review of the UKTS (carried out for its sponsors). Roberts-Miller's background also includes working on other major jointly-sponsored continuous surveys (BNTS and HBI/HBS in travel, and the NRS in media) plus the single-source TGI. He has extensive experience of designing and managing customer satisfaction surveys for airlines, cruise lines and tour operators (including developing the very widely imitated Thomson "CSQ" system). RMA projects have included:

- quantitative and qualitative surveys
- consumer research surveys
- business to business research surveys
- international research surveys
- customer satisfaction surveys
- research design consultancy
- technical evaluation of surveys
- branding, advertising and brochure research
- marketing development studies
- product launches and re-launches
- takeover and merger projects
- investment project evaluation studies

Richard Roberts-Miller is a BSc Economics graduate of Southampton University. He has been awarded Fellowships of the Market Research Society, the Royal Geographical Society, the Tourism Society and the Institute of Travel & Tourism. He has delivered papers to MRS, ESOMAR, the Marketing Society and ADMAP on market research topics such as: research for travel, research for publishing, and effective market research buying. He served on the Jamaican Government's Marketing Advisory Committee for Tourism and on their Tourism Advisory Council, and was given their "Blue Mountain Award" for services to Jamaican Tourism. Prior to establishing RMA, Roberts-Miller worked at the British Market Research Bureau (research executive), Times Newspapers (senior research executive), Thomson Organisation (group research manager), Thomson Travel (research and planning manager), Thomson Holidays (marketing controller), Thomson Travel (board member), Thomson Vacations Inc. (President & CEO), Thomson Travel Inc. (North American President & CEO) and International Thomson (US) Inc. (Executive Vice President).

Richard is also chairman of the trustees of The Alex Roberts-Miller Foundation, a charity that provides educational, sporting and social opportunities for disadvantaged young people in the UK.


## RMA (Roberts-Miller Associates)

**PO Box 104 • Dorking • Surrey RH5 6YN**
**e-mail: richard@roberts-miller.co.uk**
**website: http://www.roberts-miller.co.uk**
**Phone: 01306-741368 Fax: 01306-741356**

# Gary Bennett, FRSS, MMRS

Gary Bennett has been an established, independent marketing research consultant and statistician since March 2004. He provides a consultancy service to market research agencies, field & tab and data analysis companies and also works with end users of research, both directly and in conjunction with partners.

He offers a portfolio of tools and services primarily focused on adding value to research through the use of advanced multivariate statistical and modelling techniques. The work ranges from the design and analysis of choice and stimulus exercises needed to build models of behaviour from research data, to providing guidance on the statistical confidence around the resulting data and models. His clients have included major industrial and financial organisations, pharmaceutical companies as well as leading research agencies and consultancies.

Gary has 22 years experience in the UK market research industry, including eight years as a project director at MORI (now Ipsos-MORI), then RS Consulting. Prior to this he also worked in a project management role at Benchmark Research and London Transport (now TfL).

Whilst at TfL Gary spent two years managing the Greater London Bus Passenger Survey (GLBPS) – a complex tracking survey conducted on bus, used to calculate the proportion of Concessionary and Travelcard journeys. The results fed directly into revenue apportionment discussions with the London Boroughs and Bus operating companies.

Gary is a Full Member of the Market Research Society, a Fellow of the Royal Statistical Society and an Associate Member of the American Marketing Association.

Further details can be found on our web site:

http://www.logitresearch.com/

# Appendix B - The Documents Reviewed

Documents reviewed included (but were not limited to) the following:

**Examples of NRPS Reports:**
NRPS Wave 29 Detailed Technical Overview Report, PF/BDRC, 2013.
NRPS Wave 30 Detailed Technical Overview Report, PF/BDRC, 2014.
NRPS Wave 30 Fieldwork Report, BDRC, 2014.
NRPS Wave 30 Spring 2014 Main Report, PF/BDRC, 2014.
NRPS Wave 30 User Guidance Report, PF/BDRC, 2014.

**Examples of NRPS questionnaires and fieldwork documentation:**
Wave 9 Questionnaire (at-station version), PF/BDRC, 2003.
Wave 10 Questionnaire (at-station version), PF/BDRC, 2004.
Wave 30 Questionnaire (at-station version), PF/BDRC, 2014.
Wave 30 Questionnaire (on-train version), PF/BDRC, 2014.
Wave 31 Questionnaire (on-train version), PF/BDRC, 2014.
Wave 30 Attendance Sheet, BDRC, 2014
Wave 30 Distribution Instructions, BDRC, 2014
Wave 30 Instruction Sheet (Main), BDRC, 2014
Wave 30 Interviewer Shift Schedule, BDRC, 2014
Wave 30 Journey Purpose Showcard, BDRC, 2014
Wave 30 Respondent Record Form, BDRC, 2014

**Other related documents:**
Estimates of Station Usage 2012-13, Methodology & Validation, SDG for ORR, Feb.2014.
LATS, National Rail Results, An Introductory Report, SRA Statistics Team, Feb. 2005.
National Passenger Survey - Report on a Research Review, RMA for PF, November 2005.
*Published as: Findings of a Review of the National Passenger Survey, PF/RMA, 2006.*
National Passenger Survey - Report on a Research Review, RMA for PF, December 2010.
National Passenger Survey - Report on Weighting, PHA, April 2008.
*Published as: National Passenger Survey - Report on Weighting, PF/PHA, August 2008.*
NRTS Overview Report, DfT + Transport Scotland, December 2010.

**In addition, analysis was carried out on the following databases:**
NRPS Results Wave 30 (Spring 2014), BDRC, 2014.
NRPS Sampling Plan Wave 30 (Spring 2014), BDRC 2014.
NRTS Results, DfT, 2010.
ORR Estimates of Station Usage 2012-13, SDG for ORR, 2014.

**Key:**
BDRC = BDRC Continental (formerly Continental Research)
DfT = Department for Transport
LATS = London Area Travel Survey
NRPS = National Rail Passenger Survey (formerly NPS/National Passenger Survey)
NRTS = National Rail Travel Survey
ORR = Office of Rail Regulation
PF = Passenger Focus
PHA = Paul Harris Associates
RMA = Roberts-Miller Associates
SDG = Steer Davies Gleave
SRA = Strategic Rail Authority