Rail satisfaction survey: method development

Summary of what we've learned about the methodology and next steps

June 2022

transportfocus

Outline of this report

CAUTION:
Any data in this report which references passenger satisfaction with rail services is not based on a nationally representative sample of journeys, and responses were given in a trial of data collection approaches

transportfocus

Executive summary

Covid-19 forced the *National Rail Passenger Survey (NRPS)* to halt from 2020, and since then the way that train services are managed and evaluated in Britain has been changed. In light of all this, **Transport Focus has extensively reviewed and tested possible approaches for re-establishing a robust, nationally-applicable measurement of rail passenger experience.** This culminated in large scale pilots, firstly in late 2021 and then a "refinement" pilot in early 2022; outcomes from the latter are the focus of this report. This work was done when we were aiming to modernise the NRPS, but since then the industry (including Transport Focus) has undertaken to work towards there being one industry wide measure of passenger satisfaction. **So learnings from all of this work will now contribute to a new, joint survey programme, delivered in partnership with the Department for Transport, Great British Railways, Network Rail and the Rail Delivery Group; this is currently in development from 2022.**

Background: the pilot conducted in Autumn 2021 resulted in a proposed new approach:

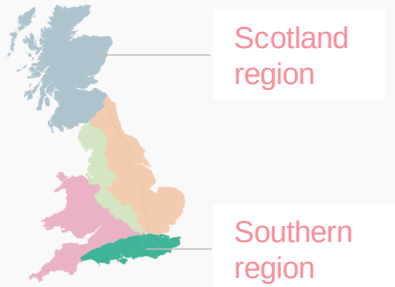
- Passengers recruited face to face at railway stations
- Most take part online, with paper forms available for those needing them
- A continuous survey, throughout the year

As we begin to shape a joint measurement tool with partners, **we have undertaken a further, "refinement" pilot, with two main objectives:**

- To estimate the likely productivity of fieldwork, to facilitate budgeting and resource planning in advance of setting up a live survey if this methodology is adopted
- To further understand some elements of the approach in more detail

This refinement pilot study took place…

4 – 27 Mar 2022



Key findings from this refinement pilot, if such an approach is adopted:

- For resourcing and efficiency, it makes sense to focus core fieldwork (i.e. the recruitment of passengers) at larger stations – likely those in Network Rail categories A-D*. This is a move away from the NRPS where we intercepted passengers at a sample of all stations, including E and F categories*.
- This should be supplemented with some fieldwork on board trains, to ensure the inclusion of passenger journeys which start at smaller stations, and the different experiences of passengers using those facilities and services.
- We would need to investigate information sources that can be used to weight resulting survey data, to ensure representation by route (in terms of journey volumes) for example, and other key experience influencers such as journey purpose. This has not been possible thus far due to uncertainty in journey patterns as affected by and now recovering since the Covid pandemic.
- This pilot also generated some further more tactical and detailed learnings about the methodology, covered within this report.

transportfocus

** See page 10 for more information on station categories.*

Introduction

Introduction

Background

Transport Focus champions the needs of transport users in Great Britain, with an emphasis on evidence-based campaigning, gathered in part via well-respected primary research. A key example has been the National Rail Passenger Survey (NRPS), measuring satisfaction with representative samples of 50,000+ rail journeys annually.

While the NRPS has been integral to measurement and accountability for the rail industry, it had some acknowledged limitations, including: 1) it provided feedback at points in time rather than year-round or more frequently; 2) findings took time to be released (due to the method, and publication process), delaying the industry's response to results. There have been comprehensive reviews of and enhancements to the survey over the years; however – partly for data continuity, and partly because the reviews have not recommended major changes – the fundamentals of the surveys have largely remained.

In 2020, after passenger numbers were severely affected by Covid-19, the NRPS was cancelled altogether after its Spring 2020 wave and throughout 2021. During this time there have also been government-led changes to the way that rail services are managed and evaluated, which may have implications for the way that passenger feedback is used in future, and therefore the way it is collected.

While all of this necessitated a break in the continuity of NRPS data, in the meantime Transport Focus has used 2020-22 to completely review and potentially update the way we could measure passenger experience. We have reviewed possible future approaches to insight collection, including pilots of some options, all leading to the intended launch of a new survey programme. Our learnings will now contribute to a jointly delivered passenger experience measurement programme, alongside the Department for Transport, Network Rail, Great British Railways and the Rail Delivery Group.

Our piloting work to date

Having trialled some ways of recruiting passengers to a survey which did not involve any face-to-face interaction, and confirming their limitations for our purposes – and as passenger footfall began to resume through 2021 – in Autumn 2021 we conducted a large scale pilot* to test:

- The (renewed) potential for intercepting passengers face-to-face and in the moment of making a journey, to invite their feedback
- The potential for encouraging most respondents to take part online, for its time, cost and flexibility (as well as likely environmental) benefits.

This approach was found to be broadly successful, and would meet our objectives for collecting and measuring passenger experience. **Passengers were willing to engage with the fieldworkers, and most were directed to an online survey** which they completed very soon afterwards. A small proportion required a paper questionnaire option, and this is likely to be needed as an alternative to ensure surveys can be inclusive.

transportfocus

** A full report on the previous trial is available separately.*

Objectives for this pilot study

Our previous work in Autumn 2021 showed that, in principle, passengers can (still) be recruited face to face at railway stations, from where most can be directed to an online survey which they will complete soon afterwards. This subsequent "refinement" pilot built on this with three core objectives:

1. **Test the potential for sampling at regional level (rather than by Train Operating Company as in the NRPS), using A-D category* stations only**

   This put into practice a proposed sampling process which had been arrived at from a separate exercise to investigate possible data sources that could be used to generate a sampling frame, and to consider some theoretical points around statistical suitability of sampling options. This separate investigation was conducted by Dave Chilvers Associates (DCA), and is available separately.

2. **Understand the productivity of fieldwork based on this sampling approach**

   To check the viability of the approach (i.e. whether it would yield sufficient number of responses for a robust survey), and begin to facilitate an estimation of budget required to run this at scale and longer term, including with consideration for how productivity varies by time of day, region, station category, etc.

   (For the Autumn 2021 study, rail passengers were recruited at stations in a number of towns and cities were chosen, and no specific systematic way of determining these locations was used in this earlier study)

3. **Test supplementary recruitment options**

   This was partly to validate the approach based on A-D stations, and partly in anticipation of its likely limitations meaning we wished to explore some alternatives*

   - *Recruiting on board trains to cover routes likely to be excluded from the at-station fieldwork*
   - *Recruiting passengers as they disembarked, to try and include a broader range of stations at which passengers had originally boarded*

We also tested a number of more tactical points of the methodology:

- Offering a take-away postcard showing a QR code, and shortening the URL – both developments on the survey access methods that we had tested previously
- Collecting email addresses within the survey itself from passengers who had not provided this at point of recruitment (for tracking purposes)
- Encouraging survey completion after the journey (we found that some passengers did so during their journey, or even pre-boarding in the Autumn 2021 study)
- Testing a way of collecting demographic information for the passenger universe, which we anticipate needing in order to inform weighting

transportfocus

A report on the passenger experience findings is available separately. This report focusses on the key learnings about the methodology itself.

** See page 10 for more information on station categories*

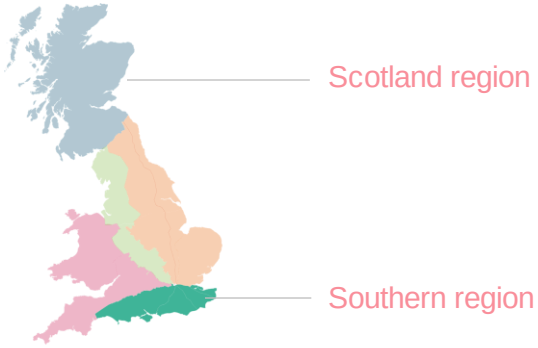
Overview of the fieldwork

4 – 27 Mar 2022

Across all days of the week

Passengers recruited between 6am and 7pm

(With additional tests on later evening fieldwork)



At the outset of this pilot, our proposed approach to a new survey was expected to measure passenger experience at the level of Network Rail regions or routes.

Two regions were covered in this pilot, providing a mix of:

- Stations in more rural, urban, suburban and metropolitan areas, and different types of routes (heavy commuter routes, long distance, etc.)
- Routes on which a large proportion of passengers flow through a small number of stations, and routes those with many smaller stations.

Passengers recruited as they waited to board a train, at railway stations.

(Additional comparative tests conducted with disembarking passengers, and on board)

Eligible passengers invited to take part online, accessed by a choice of:

- Scanning a QR code at point of recruitment
- Providing an email address or mobile number in order to receive a link by email or SMS
- Taking down a short URL
- Taking away a postcard displaying the QR code and URL

Paper self-completion questionnaires offered for those who could not take part online.

215 fieldworker (recruitment) shifts completed

6,776 passengers recruited overall*

2,113 took part in the survey

transportfocus

** Recruitment figures should be taken with some caution. We suspect that some fieldworkers may not have recorded all recruitment interactions fully (especially where a paper questionnaire was given out and there was less immediate and practical need to record the details of this). Recruitment may therefore have been a little higher than this.*

Overview of types of shift

215 fieldwork shifts were conducted in total during the pilot, split across a number of test approaches

		"Core" pilot approach	Additional test cells			
Overall	**215**	**154**	**12**	**18**	**20**	**11**
		A-D stations*	E-F stations*	"Out of hours" *(7-10pm)*	"Disembarking" *(Passengers recruited as they got off a train rather than as they waited to board)*	On board
Southern region	123	84	10	9	10	10
Scotland region	92	70	2	9	10	1

transportfocus

** See page 10 for more information on station categories.*

The sampling approach: at-station recruitment

The process for sampling railway stations in the core sample cell: at A-D category stations*

1. Establish a list of all railway stations within the Network Rail Regions covered (Scotland and Southern)
2. Identify the number of passenger journeys originating from each station. This was done in the same way as for the NRPS, taking all entries and half of the interchanges at each station, as recorded in the most recent published data
3. For the core sample, which we determined for this pilot would cover categories A-D: remove all other stations
4. Select stations from the remaining list, with probability proportional to the number of passenger journeys (PPS) for each region separately. Some (larger) stations may be selected more than once, and many (smaller) stations may not be selected at all. Each selected occurrence would thus form the location for a fieldworker shift

   The PPS selection was conducted across the whole of each region, meaning that A and B category stations were typically selected more often than C or D category stations. In practice for this pilot, some manual adjustment was made at this point to produce a set number of shifts within each category, to enable comparison of response (otherwise in a natural selection there would have been too few for analysis in some categories)
5. Assign a day and time to each selection, covering all days of the week equally, and an even split of 3-hour segments between 7am and 7pm. (Half of the 7-10am slots were then moved to 6-9am to better cover earlier morning commuting times). Days and times were given at random within each category, for similar coverage for all categories
6. Review time and day assignments in light of an analysis of when services called at the selected stations, including any planned disruption where this meant stations were closed during the fieldwork period (this affected Glasgow Central Low Level in particular). This led to some minor manual adjustments to times or days where it was clear that few or no passengers could be intercepted
7. Assign a fieldwork date to each selection.

This provided a good spread of stations in each region, by route and geography, and steps 5-7 generated a fieldwork plan.

Note: the final fieldwork varied a little from this original selection process. This was due to limited fieldworker resource, meaning some planned shifts could not go ahead in practice. This mainly affected shifts at E-F stations in Scotland.

Other at-station cells used the same overall principles:

- For fieldwork shifts at E-F stations, steps 4-7* were followed using relevant stations only.
- For "disembarking" shifts, we wanted to compare fieldwork productivity directly to the core sample cell, so a sub-set of the core fieldwork plan – including time of day and day of the week – was taken, at random but balanced by category. While these shifts then took place on the same day of the week as their paired shift in the main sample, they were assigned a different week.
- For "out of hours" shifts, steps 4-7 were followed using A-D category stations, but with all selections assigned as 7-10pm

transportfocus

** See following two pages for further detail on some of these steps, the definition of station categories, and the rationale for using A-D stations*

The sampling approach: on-board recruitment

A pragmatic approach was taken to produce the sample and fieldwork plan for the on-board test cell

1. A list was made of all stations which had not been selected in the core cell of A-D category stations. This was therefore mainly made up of D, E and F category stations*.
2. Routes were then identified across each of the two regions of interest, which had a high proportion of these stations. This was done manually.
3. Published timetables were then used to manually identify return trips that could be made within a roughly 3-hour period (the length of a typical fieldworker shift), ensuring that the resulting fieldwork plan covered all days of the week and times of day between approximately 7am-7pm. Some of these return trips did of course pass through or turn around at a larger station, but this was acceptable providing the majority of stations along the route were in the lower categories.
4. As for the at-station shifts, the resulting fieldwork plan was reviewed and manually adjusted slightly to ensure a reasonably good geographic coverage

Note: the final fieldwork varied a little from this original selection process. This was due to limited fieldworker resource, meaning some planned shifts could not go ahead in practice. This mainly affected shifts at E-F stations in Scotland.

If on-board recruitment is used in any future survey, a systematic sampling approach will be needed to make this scalable

If on-board recruitment is used to fill gaps (i.e. to allow inclusion of a wider set of smaller, especially E-F category, stations) a systematic way of defining these routes will need to be found that is consistent across regions. This could build on the approach used to set up on-board recruitment work in the NRPS.

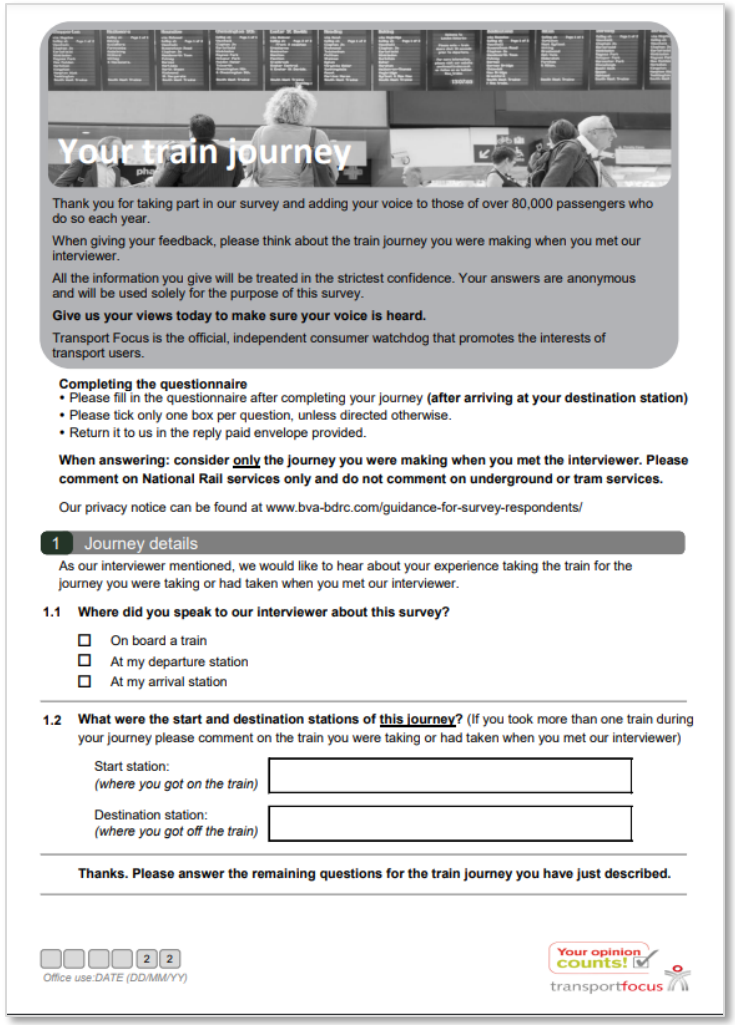
We do not see a way to reduce the manual element of identifying practical return trips within a shift. It may be possible to partly automate this using a database of all timetabled journeys, but from our experience here (and from the BPS which saw similar practical issues), we know that manual consultation of timetables is very likely to be required to work out worthwhile return trips for fieldworkers in practice, especially on lower frequency routes which make up a large part of these. This is time consuming and inevitably introduces some subjectivity, and so will be a key consideration for the viability of, and budgeting for, including on-board recruitment in the new passenger survey programme.

transportfocus

* See page 10 for more information on station categories.

The questionnaire

The core questionnaire *(page 1)*



Your train journey

Thank you for taking part in our survey and adding your voice to those of over 80,000 passengers who do so each year.

When giving your feedback, please think about the train journey you were making when you met our interviewer.

All the information you give will be treated in the strictest confidence. Your answers are anonymous and will be used solely for the purpose of this survey.

Give us your views today to make sure your voice is heard.

Transport Focus is the official, independent consumer watchdog that promotes the interests of transport users.

Completing the questionnaire

- Please fill in the questionnaire after completing your journey **(after arriving at your destination station)**
- Please tick only one box per question, unless directed otherwise.
- Return it to us in the reply paid envelope provided.

When answering: consider only the journey you were making when you met the interviewer. Please comment on National Rail services only and do not comment on underground or tram services.

Our privacy notice can be found at www.bva-bdrc.com/guidance-for-survey-respondents/

1 Journey details

As our interviewer mentioned, we would like to hear about your experience taking the train for the journey you were taking or had taken when you met our interviewer.

1.1 Where did you speak to our interviewer about this survey?

- ☐ On board a train
- ☐ At my departure station
- ☐ At my arrival station

1.2 What were the start and destination stations of this journey? (If you took more than one train during your journey please comment on the train you were taking or had taken when you met our interviewer)

Start station: *(where you got on the train)*

Destination station: *(where you got off the train)*

Thanks. Please answer the remaining questions for the train journey you have just described.

☐☐☐☐22
Office use:DATE (DD/MM/YY)

Your opinion counts!
transportfocus

The core questionnaire collected key journey details, headline experience questions (overall journey satisfaction and other measures known to be important as part of this), and passenger details such as demographics.

Shown here is the paper version of the questionnaire, which was almost identical in content to the online version.

After completing the core set, respondents were invited to continue with some additional questions. In this pilot, those willing to continue were served a longer questionnaire, which was either a "NRPS" equivalent, or an alternative long version which included most of the same questions as the NRPS, plus some others including questions on priorities for improvement.

For online respondents, the core questions took approximately 5 minutes for most people*. The longer versions took around 11 minutes for most people*.

Full copies of all questionnaire versions, including recruitment screeners, are available.

**Using median average, and excluding outliers where a respondent spent an hour or more on the survey, but does include any respondent who may have dipped in and out of the survey over the course of up to an hour*

transportfocus

The core questionnaire *(page 2-4)*

1.3 What was the approximate scheduled departure time for this journey?

Hours: ☐☐ Minutes: ☐☐ ☐ Don't know **- GO TO QUESTION 1.3.1**

Please use the 24hr clock e.g. 17:34

1.3.1 Are you able to estimate the departure time for this journey?

- ☐ Before 07:00
- ☐ 07:01 - 09:59
- ☐ 10:00 – 15:59
- ☐ 16:00 – 18:59
- ☐ 19:00 onwards
- ☐ Don't know

1.4 And which train company did you use for this journey?

- ☐ Avanti West Coast
- ☐ CrossCountry
- ☐ Gatwick Express
- ☐ GWR - Great Western Railway
- ☐ LNER – London North Eastern Railway
- ☐ London Overground
- ☐ Lumo
- ☐ ScotRail
- ☐ Southeastern
- ☐ Southern
- ☐ South Western Railway (incl. Island Line)
- ☐ Thameslink
- ☐ TransPennine Express
- ☐ None of these
- ☐ Don't know / cannot remember

1.5 What was the main purpose of this journey?

- ☐ Daily commuting to/from work
- ☐ Less regular commuting to/from work
- ☐ Daily commuting for education (to/from college/school/university)
- ☐ Less regular commuting for education (to/from college/school/university)
- ☐ On company business (or own if self-employed)
- ☐ On personal business (job interview, dentist, etc)
- ☐ Shopping trip
- ☐ Travel to/from holiday
- ☐ A day out
- ☐ Sport
- ☐ Visiting friends or relatives
- ☐ Other leisure trip

1.6 What type of ticket did you use for this journey?

- ☐ Anytime / Peak
- ☐ Off-Peak
- ☐ Other. *Please write in:*
- ☐ Season
- ☐ Advance (booked train only)

2 Your journey experience

2.1 Please tell us a bit more about this journey. We're interested in what was good and what was bad. Please also tell us anything else that you think is worth mentioning about this journey.

Please write in:

We would now like your opinion on the train company you used during this journey.

2.2 How would you rate the experience on the train of that company for the following...?

	Very good	Fairly good	Neither good nor poor	Fairly poor	Very poor	Did not use/ no opinion
Punctuality/reliability of the train (i.e. that the train arrived and departed on time)	☐	☐	☐	☐	☐	☐
Cleanliness inside the train	☐	☐	☐	☐	☐	☐
Toilet facilities on board	☐	☐	☐	☐	☐	☐
Sufficient ventilation on board the train	☐	☐	☐	☐	☐	☐
Sufficient room for all passengers to sit/stand comfortably	☐	☐	☐	☐	☐	☐

2.3 How satisfied were you with your overall journey taking into account where you boarded the train and the train you travelled on?

- ☐ Very satisfied
- ☐ Fairly satisfied
- ☐ Neither satisfied nor dissatisfied
- ☐ Fairly dissatisfied
- ☐ Very dissatisfied
- ☐ Don't know/ no opinion

2.4 How satisfied were you with the the value for money for your journey?

- ☐ Very satisfied
- ☐ Fairly satisfied
- ☐ Neither satisfied nor dissatisfied
- ☐ Fairly dissatisfied
- ☐ Very dissatisfied
- ☐ Don't know/ no opinion

3 About you

In this final section we ask for some information about you. Any information you give us here is used for research purposes only and not to identify any particular individual. You are also free to decide whether you want to give us this information or not.

We ask these questions so that we can understand how different passengers' experiences vary, so, for example, what do younger passengers think compared to those who are middle-aged or of retirement age.

3.1 How would you describe yourself?

- ☐ Female
- ☐ Male
- ☐ Prefer to self-describe

Please write in

- ☐ Prefer not to say

3.2 Please tell us your age?

☐☐

- ☐ Prefer not to say

3.3 Were you travelling with:

Please tick all that apply.

- ☐ Heavy/bulky luggage/other large items
- ☐ A pushchair/buggy/pram
- ☐ A folding bicycle
- ☐ A dog
- ☐ A helper
- ☐ A mobility scooter
- ☐ A wheelchair
- ☐ None of the above

3.4 We would like to ask if you have any disability. This is to help us better understand the needs you might have. If you do not consent to us collecting this information, please tick the last box below.

Are you affected by any physical or mental health conditions or illnesses lasting or expected to last 12 months or more?

Please tick all that apply

- ☐ No: None
- ☐ Yes: Vision (e.g. blindness or partial sight)
- ☐ Yes: Hearing (e.g. deafness or partial hearing)
- ☐ Yes: Mobility (e.g. only able to walk short distances or difficulty climbing stairs)
- ☐ Yes: Dexterity (e.g. difficulty lifting or carrying objects or using a keyboard)
- ☐ Yes: Difficulty with learning or understanding or concentrating
- ☐ Yes: Memory
- ☐ Yes: Mental health
- ☐ Yes: Stamina or breathing or fatigue
- ☐ Yes: Socially or behaviourally (e.g. associated with autism, attention deficit disorder or Asperger's)
- ☐ Something else

Please specify:

- ☐ Prefer not to say

3.5 We would like to ask your ethnic background. If you do not consent to us collecting this information, please tick the last box below.

- ☐ White
- ☐ Mixed/multiple ethnic groups
- ☐ Asian or Asian British
- ☐ Black, African/Caribbean or Black British
- ☐ Chinese
- ☐ Arab
- ☐ Other ethnic group
- ☐ Prefer not to say

THANK YOU FOR COMPLETING THIS SURVEY AND SHARING YOUR OPINIONS WITH US SO YOUR VOICE IS HEARD

transportfocus

Key findings on the methodology

A separate report covers what we heard from passengers about their rail journey experiences

Sampling and recruitment at category A-D stations, versus alternatives

Findings: 1 / 10

Fieldwork productivity varied but was lower than we might have hoped. Unsurprisingly, larger stations were more productive than smaller stations

Key findings from this pilot:

- Overall, fieldwork at these stations was less productive than we saw previously for NRPS (even if we were to compare the productivity at each station category, like for like), and this echoes our previous pilot.
- This may be partly due to reduced railway usage since Covid-19 (in which case productivity may increase a little if patronage increased again), and / or a decrease in the public's willingness to take part in research (which reflects market research generally, across many industries).
- Nevertheless, once a passenger was recruited, response rate was reasonably strong

Fieldwork productivity: core approach (category A-D stations)

	Avg point-in-time footfall*	Avg no. recruits per shift	Avg no. complete surveys per shift	Conversion rate (completes/recruits)
A	56	44	13.0	30%
B	24	36	11.9	33%
C	14	34	8.8	26%
D	9	23	6.6	29%
Total A-D	**30**	**36**	**10.6**	**30%**
*NRPS***	*NA*	*69*	*15*	*22%*

Response was stronger in the Southern region (with 34% conversion rate and an average of 11 complete surveys per shift) than in Scotland (with 25% conversion and 9 completes per shift). Relative patterns by station size were, however, similar.

Arguably, the significantly stronger productivity at larger stations could even call into question the value of including category D stations in the sample. At the very least, in developing such a future approach further, we should seek to build in steps to the sampling so that fieldwork for category D stations is focussed only at busier times. *For the bus equivalent to this survey, our proposed sampling approach uses a measure of passenger journeys taking place in each daypart within each sampling location, and PPS selections are made at "location-daypart" level rather than location level only. A similar idea could be considered here.*

Regardless, while the figures used here can inform budgeting in the short term, it will also be desirable to increase productivity further in any future survey, and closer to the achievements in the NRPS. Measures – and potentially, in the longer term, targets – to achieve this should be sought from bidding contractors if this survey approach is adopted and tendered.

transportfocus

** Fieldworkers paused for five minutes at a pre-determined time during each shift to count the number of people who might potentially be recruited at that point in time. This gives us a snapshot of footfall, building to a pattern of typical footfall across the fieldwork areas as a whole*

***Figures taken from NRPS Autumn 2019*

Findings: 2 / 10

Recruiting at E-F category stations, and disembarking passengers, was very unproductive

By constraining the sample to larger stations, we will inevitably exclude some passenger journeys, meaning the experiences of those using smaller stations – with fewer facilities, staff, and sometimes fewer opening hours – could be overlooked in survey results. This may affect some routes, journey types and operators disproportionately.

We therefore wished to test the effectiveness of on-board fieldwork and the recruitment of passengers disembarking at larger stations, as possible ways to capture the experiences of people using a broader range of origin stations. As a counterpoint to these tests, we also ran a control cell, with recruitment at E and F category stations.

Key evidence from this pilot:

- Given the lower yield from category C and D stations compared to A and B stations, it was unsurprising that even fewer passengers could be recruited from smaller category E, and un-staffed category F stations – as shown in the table below.
- Shifts targeting those getting off a train (arriving into a station) were even less productive. A low recruitment rate here was compounded by a relatively low conversion to survey completion, since these people were much less willing to stop and engage with fieldworkers, and were less of a captive audience than those waiting to board.

Fieldwork productivity: at-station recruitment tests

	Avg point-in-time footfall*	Avg. recruits / shift	Avg. complete surveys / shift	Conversion (complete / recruits)
E-F	**12**	**12**	**4.8**	**39%**
A-D (arriving)	**22**	**14**	**2.7**	**20%**
A-D (departing)	30	36	10.6	30%

We also took this further to understand what relative contribution E-F stations might make to the total sample, if all stations were sampled with probability proportional to size. As summarised below for the Southern region, E-F stations would yield less than half the number of responses than they "should" if they were to contribute proportionately (1.7% vs. 4.1%). They would therefore either need to be oversampled (taking more time and resource), or upweighted (reducing data integrity), by a factor of over 2.

For Scotland, where E-F stations are the origin point for a larger proportion of all journeys, this effect would be even more extreme.

Projected fieldwork outcome using PPS station selection: Southern

	Stat-ions	Pax. journeys	No. shifts in typical PPS sample (out of 100)	% responses based on avg. completes / shift in this pilot
A-D	**55%**	**95.9%**	**96**	**98.3%**
E-F	45%	4.1%	4	1.7%

While we would need to find a way of normalising the sample composition across different routes and regions (and operators) so that resulting satisfaction data is measured in a fair and consistent way across the country, we would also need to balance this with efficiency and ensuring the survey provides value for money. Neither conducting fieldwork at very small and unstaffed stations, nor intercepting passengers as they disembark a train, is a productive way to achieve this, so alternative options would be preferable.

transportfocus

** Fieldworkers paused during each shift to count the number of people who might potentially be recruited at that point in time, providing a snapshot of footfall across the fieldwork areas as a whole*

Findings: 3 / 10

On-board shifts were more effective overall for capturing responses about smaller stations, than those taking place at stations

Key evidence from this pilot:

As shown in the table here, on board recruitment yielded three times as many completed questionnaires per shift compared to at-station recruitment at E and F stations – a very significant difference.

Even from this small test, overall we can also say that the on-board recruitment approach showed potential to broaden the range of different E-F origin stations covered as well as the potential to collect more responses about them:

Fieldwork productivity: on board and at-stop recruitment comparison

Test cell	Avg. point-in-time footfall*	Avg. recruits per shift	Avg. complete surveys per shift	Conversion rate (completes/recruits)
On board	**25**	**43**	**14.4** *(of which 5.4 related to E-F origin stations)*	**33%**
Total E-F	12	12	4.8	39%
A-D (departing)	30	36	10.6	30%
NRPS (on board)*	*NA*	*not collected*	*26.2*	*not collected*

- Each of the on-board shifts in this pilot generated responses about passenger journeys with nine different origin stations, on average
- Of the responses captured via on-board recruitment, 36% of the passenger journeys evaluated had started at category E or F stations; this equated to an average of 5.4 category E-F stations captured per shift (and 3.8 different E-F stations per shift). This compares to just under 5 per shift at E-F stations, of which all would have been from the same station in that shift.

If the broad approach to measuring rail passenger experience as piloted here is adopted, then conducting some of the fieldwork on board trains would be a useful supplementary feature – as it was in the NRPS – to help ensure that the views of people starting their journeys at smaller stations are included in the survey findings. On-board recruitment shifts should be targeted at those routes where smaller proportions of all passenger journeys are represented by the A-D stations.

**Figures taken from NRPS Autumn 2019, across three TOCs on which on-board recruitment was conducted*

Who took part in the survey, and how they accessed it

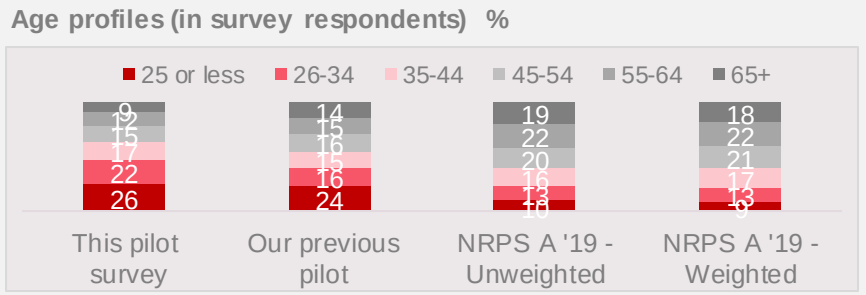
Findings: 4 / 10

Like our previous test on face to face recruitment with a steer to online completion, the approach generated response from a good mix of passengers, including groups that had been harder to engage with the NRPS

Key evidence from this pilot:

The overall age profile of rail passenger respondents in this pilot was much younger than both the weighted and unweighted profiles for the NRPS (though we had never specifically weighted by age in the NRPS). It is likely that this is partly due to the methodology itself, which emphasises online completion and a range of easy ways to access this, and partly reflecting a real change, where anecdotal evidence suggests that the population of rail users is younger since the pandemic.

Age profiles were very similar for the Southern and Scotland regions.



The journey purpose profile was quite different to our previous pilot, with a stronger contribution from commuters (48% in the Southern region, 40% in Scotland). This was very slightly lower than in the NRPS (52% and 39% respectively), which had been sampled and weighted to achieve representative profiles by journey purpose and where commuters had been harder to recruit and convert. Again, this probably reflects a real reduction in commuting by rail overall since the pandemic, but suggests that the new mainly online approach can pick up this group without any special oversampling or weighting.

There also appeared to be a small bias towards females rather than males (56% vs 41%, with the remainder preferring another term or not to answer). This reflects our previous surveys, and research generally across most sectors.

Weighting and / or over or under-sampling may be needed to correct for sample biases in any future live survey, though likely less so than in the NRPS, making for more robust data overall. The precise nature of these weights, and data to inform target universe profiles, would need to be determined with further more in-depth work as we moved towards a full live survey. At present, given changes in the rail traveller population since the pandemic (which may still be in flux and not yet settled), there is no reliable source for this. However, in the short term we tested a possible way of estimating a universe age and gender profile, as outlined on the next page.

Findings: 5 / 10

The method tested for collecting demographic data to inform weighting looks to have been successful

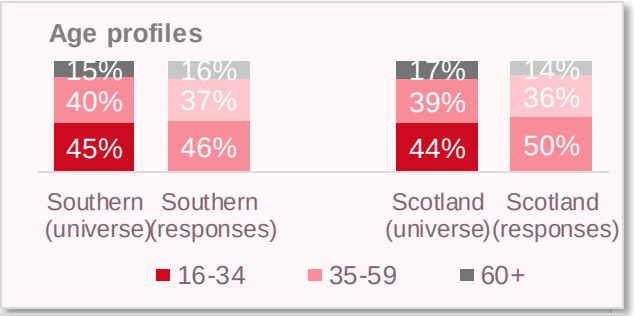
There is no nationally consistent information source which gives us the demographic profile of train passengers, against which to assess how representative a survey like this is, and therefore how we might weight it if needed (which, based on previous experience, we expect). Therefore we may need to collect this information ourselves. It is possible that a source such as the National Rail Travel Survey could be used, but this may not be frequent enough for our purposes).

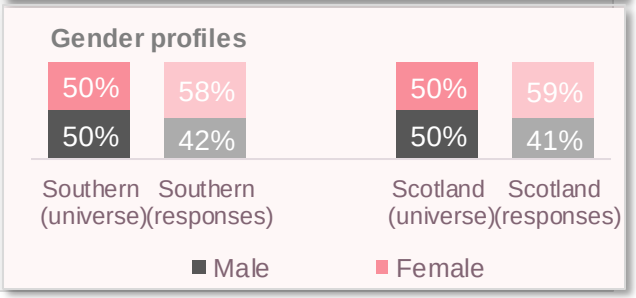
Alongside this pilot, we adapted an approach to collecting demographics that had previously been used successfully for the Bus Passenger Survey (BPS) up until 2019:

- Fieldworkers were directed to pause at a given interval during each shift, to record the number of people who might feasibly have been approached (i.e. were standing at a reachable distance, and were clearly waiting for a train) at that point in time, and their observable gender and age (in three broad age bands)
- The same principle was applied to passengers on board, and disembarking, for the test-cell shifts as relevant
- The count was made electronically, with fieldworkers pressing a counter button on their device for each observed person against each age-gender cell.

Key evidence from this pilot:

- These profiling observations look to have been conducted as directed, the data was reasonably complete for all types of shift, and there were no indications that fieldworkers had difficulty with the task. Looking at the same evidence from the bus equivalent to this study, this indicated improved quality of this task and its resulting data, compared to the same process in the BPS (where the task was more manual, on paper).
- Again, evidence from the bus version of this pilot also indicated that the passenger profiles collected were fairly similar to those in the BPS, and (with acknowledgement of some known changes in the traveller population) seemed intuitively sensible, bringing some reassurance of the credibility and accuracy of the data. We can infer the same in this rail scenario.
- The universe data is close to the profile of respondents. This suggests that some minimal corrective weighting would be required to upweight males (as expected) and possibly to downweight the youngest passengers.





We would suggest using this approach for estimating passenger profiles to inform weights in a future live survey that used the approach piloted in this project. We also acknowledge the need to investigate ways to inform any weighting by other variables known to impact satisfaction ratings, in particular by journey purpose.

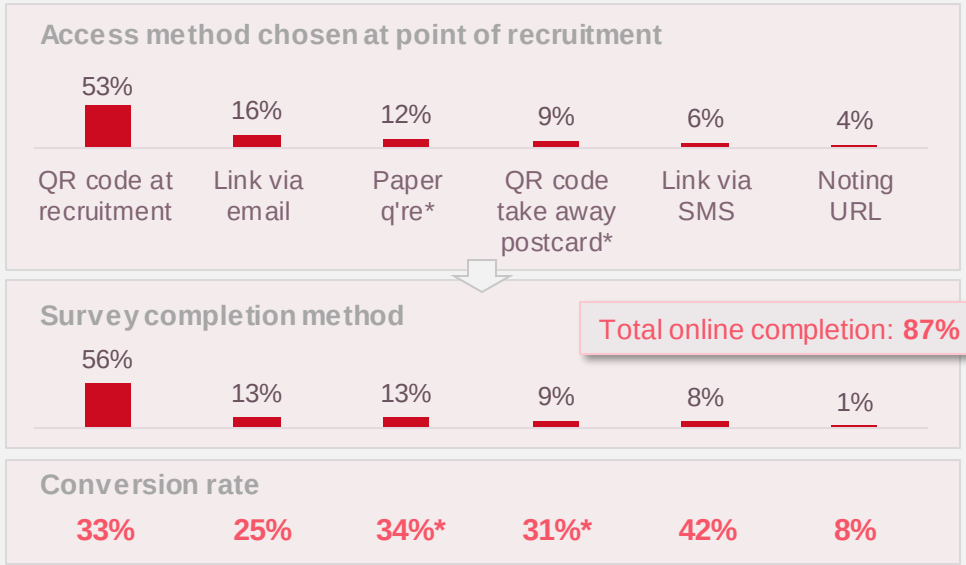
transportfocus

Findings: 6 / 10

Ways of accessing the survey very broadly reflected findings from our previous trial, with QR code and emailed links the most popular, and most responses being made online– marking a major shift away from paper questionnaires as in NRPS

Key findings from this pilot:

- The most popular access methods at point of recruitment were the QR code followed by the emailed survey link. Both appealed to a range of passengers, but the QR appealed to a slightly younger group overall.
- Paper questionnaires were preferred by a substantial minority, who were typically older people – a third of all respondents aged over 60 (and two thirds of over 70s) took part on paper, confirming that this option remains important for maintaining inclusiveness of the sample.
- The URL option added very few recruits, and these were typically in the travel-to-work cohort who were also well (and more so) recruited via other methods.
- The take-away QR code was newly introduced in this pilot, intended as an option for those who may not have time to stop and scan the code there and then. This method appealed to a small group and was important, alongside the in-situ QR code, for under 35s, though in practice it did not stand out as especially more or less appealing to any particular group by journey purpose.



Note: in the previous trial we were able to confirm that, while respondents completing the survey via different methods – and accessing it in different ways – have different response patterns for satisfaction questions, this is largely driven by demographic differences and can be controlled for with weighting

We propose that the new passenger experience measurement survey should be primarily online, with paper questionnaires used sparingly to enable participation by those not able to do so online. All of the access methods used here should be used in such an approach, except for the URL which brings little additional value, and we would continue to work with the research supplier(s) to maintain and/or improve on conversion where possible.

transportfocus

** Recruitment and conversion data for paper questionnaires and take-away postcards should be taken with some caution. We suspect that some fieldworkers may not have recorded all recruitment interactions fully, and if so this will have mainly affected data for these methods, since survey participation was possible without completing the full recruitment form.*

Findings: 7 / 10

Around half of those entering the survey via "non-contact" methods prefer to be completely anonymous

We were able to collect contact details, enabling more contact with individuals, from most of the remainder

When passengers choose to enter the online survey via an emailed or SMS link, they provide contact details in the process. Entering via the QR code or the URL ("non-contact") methods means they do not need to do this at point of recruitment, and this may be part of these methods' appeal.

Capturing an email address or mobile phone number has some key advantages, however, including the ability to send reminders if a respondent has not completed the survey, and allowing respondents to pick up where they left off if they temporarily drop out. It also enables us to fully track the efficacy of different access methods (e.g. by region, time of day and so on) to assist our ongoing monitoring and continual refinement of the methodology.

Key evidence from this pilot:

- In this pilot we asked half of the non-contact respondents (at random) if they would provide an email address so that we could send them a link to continue in case they were interrupted. Of 689 people asked, 40% agreed. 53% preferred not to share their email, and 7% said they did not have one.

  This pattern varied a little by region, with half of respondents in Scotland providing an email address, and closer to a third doing so in the South.

- We also saw that just over 11% of all entrants to the survey who were asked this question dropped out at this point. This was the very first question – and we know from previous experience that the earliest questions provoke the highest drop-out rates as people enter and change their minds – so this is to be expected to an extent and may not have been entirely caused by the question itself, but this was particularly high (with a maximum of 4% drop out for other questions, including those around the date and time of a journey, which we have previously seen to be drop-out triggers).

Initial ideas for increasing the proportion of non-contact recruits who subsequently do provide an email address – or at least reducing drop-out – might include:

- *Better understanding the barriers to providing contact details, either at recruitment or after entering the survey, and working to mitigate these*
- *Amending the wording in the way we ask for contact details, once a respondent has started the survey*

And we would like to investigate the use of unique QR codes for each individual recruit, where they choose this access method. This will enable tracking of response (if not to the individual unless they subsequently provide contact information) and the ability for the respondent to pick up where they left off if they temporarily drop out.

Providing non-contact routes to the online survey would be useful to ensure inclusion of people who may be uncomfortable with sharing their details at recruitment – clearly many feel this way – but it is possible to collect this from some people once they enter the survey itself.

We may be able to increase the proportion to agree to this however, and crucially we would need to find a way of doing this that provokes minimal resistance. We would seek proposals for these points from bidders for any future live survey using this approach (with some initial ideas given above).

How respondents engaged with the questions

Findings: 8 / 10

Some people were willing to answer additional questions about their journey, but fewer than expected

Key evidence from this pilot:

38% of online respondents answered beyond the core question set. These people were a little more likely to be in the travel-to-work cohort*, but these variations were fairly small, and there was no real difference here between Scotland and the Southern region.

Overall this compares to three quarters in our previous pilot using a similar approach.

Key differences in the two studies' questionnaires which may have influenced this included:

Half of the sample in the previous study were asked an open-ended question, immediately before the invitation to continue, about desired improvements to their train service. This was not included in the second pilot.

- Two questions before the invitation to continue: respondents in the previous study rated their overall journey experience using both a verbal satisfaction scale and a more light-hearted version – a star rating or an emoji-style rating. Only the verbal scale featured in this pilot.
- A longer estimation of likely duration for the remaining questions: in the previous study respondents were told that the additional questions would take around 5 additional minutes. In practice this was found to be an under-estimate and so in this subsequent study respondents were told it would take around 5-10 minutes longer.

We had previously concluded that structuring the new survey as a core set plus optional further questions was a useful way to maximise response among those resistant to a long questionnaire. It also brings potential for more flexibility, with the opportunity for using different question modules in different circumstances, for instance. However, more work would be needed to understand how to maximise willingness to continue with further questions, for this to be viable. Some ideas are presented here, but we would seek proposals from bidders for any live survey and continue to explore this further.

transportfocus

** For example this group was more likely to be aged 35-59, and more likely to travel daily for commuting*

Findings: 9 / 10

The accuracy of journey details given by respondents was not as high as we would have liked

Respondents were asked to give the origin and destination stations for their train journey (meaning the specific leg of the journey when they were recruited, if they were making any changes during the journey as a whole), as well as the scheduled departure time and the train operating company (TOC).

These details were not critical in this pilot, but could be important in a live survey so that each response– i.e. each journey evaluation – could be linked to a specific route, and potentially a specific TOC. (In the NRPS, each survey response was attributed to a specific TOC, and a specific sub-route within that TOC, to make the results as actionable as possible as well as ensuring that the data could be used to hold TOCs to account fairly and accurately. A fairly substantial validation step, comparing responses to timetable databases, was including in part of the data cleaning stage, to facilitate this).

In the online questionnaire in this study, some validation was programmed in so that a respondent could not choose a TOC which did not serve at least one of the two stations (unless they selected "don't know" for the TOC). No additional validation was conducted on the time of boarding for example, but some cleaning was applied to correct TOCs where the respondent did not know this, or when some journey details were infeasible but resolvable.

Key evidence from this pilot:

- Typical examples of incorrect details were where a respondent picked a similar sounding TOC name but was wrong, or entered the station name for a final destination, where we wanted to know the destination of their current leg.
- This was necessary for approximately **14%** of all responses, and this was fairly similar whether the respondent was recruited on board a train or at a station.

Since some of this cleaning task is quite manual, 14% requiring some cleaning is higher than we feel is acceptable. It also means that up to 14% of respondents may have been evaluating a slightly different journey than we might use the results against.

We would seek proposals for improving on this – whilst turning results around quickly – in any contractor tender. Some initial thoughts include:

- Pre-filling the start station based on the recruitment / shift location (this would require, among other things, unique QR codes to be generated for each location, for respondents accessing the survey this way)
- More direction from fieldworkers on what information will be asked of them in, and how to answer, the first few questions
- Tighter and more advanced programming in the script to prevent respondents entering impossible 1-leg journeys

Findings: 10 / 10

What happened when we encouraged people not to fill in the survey until after the journey?

An idiosyncrasy discovered in our previous pilot of the broad approach used here was that some online respondents had started– and sometimes completed – the questionnaire before their train had departed. This means that they are technically less able to fully evaluate that journey. The majority of people in this group were more frequent users (mainly commuters), and so we assume that they answered some questions, such as the cleanliness of the inside of the train, based on frequent experience and expectation. The majority of these people accessed the survey via the QR code, which was of course possible to do immediately.

In this "refinement" pilot, we sought to prevent this as far as possible by fieldworkers telling all passengers at the point of recruitment that they should complete the questionnaire after their journey.

(It may have been possible to programme a delay or instruction within the online questionnaire, to prevent people from continuing if they entered a train departure time in the future. This was decided against because it might put people off or they might forget to return at a later time, and it could be acceptable to people to fill in the survey while on board (indeed we know that this had always happened with paper questionnaires in the NRPS).

Key evidence from this pilot:

- Of 1,846 online responses, 1,340 (73%) were made on the same day as the passenger was recruited. 99 (5%) appear to have been completed before the scheduled departure time of the journey, as reported by the passenger (and some others which were completed soon after the departure time may well have been started before the passenger boarded). This does appear to be an improvement on our previous pilot, where this was 12%.

  As in the previous pilot, almost all of these early completers (95) accessed the survey via the QR code.

- Overall the response rate (conversion from recruitment to survey completion) was 22% in the previous pilot, and was stronger at 30% overall here. There was no discernible negative impact on response rate, therefore, by instructing respondents to wait a little.

We would expect to use this fairly light touch tactic to discourage respondents from completing the survey too early in any future survey, since it appears to provide a good balance between keeping early completion to a minimum while not hindering participation. It is likely that this will need to be monitored however to ensure that fieldworkers are consistent, and to ensure that the proportion of early completers does not creep up.

Summary

Recap on suggested key principles for a new survey approach

Our piloting and review work up to the end of 2021 allowed us to outline some proposed essentials for a future approach to measuring passenger experience:

Passengers feed back about a **single leg of a journey, made on the day of recruitment**

Pro-active recruitment of passengers, at point of usage

…. meaning **face to face** interception as passengers make journeys; this is verified and inclusive

Online as the main method for survey completion – but paper option has a place

Concise questionnaire focusing on essential metrics – with optional, modular question sets

Structured questions about overall journey experience, with core measures carried over from previous NRPS and BPS

This pilot work has sought to refine elements of this approach in further detail

transportfocus

Summary: our findings for the fundamentals of this approach

The core at-station recruitment work at larger stations

Given the lower productivity we saw here (and in our earlier pilot) than historically – either due to changing travel habits after Covid, and / or a long term, gradual reduction of public interest in taking part in surveys – **it makes sense from a resourcing perspective to focus at-station recruitment at larger stations, if this overall approach is adopted. This is a move away from the NRPS where we selected from a sample frame of all stations, including E and F categories*.**

- Taking evidence from this pilot we feel that **a realistic assumption for the average number of survey completions per (3-hour) shift** using this approach, and using A-D category stations as the core sample, **could be between 10 and 11.**
- Transport Focus has previously proposed an annual sample size of 50,000 responses, equating to c. 1,000 per week. **If this proposal was adopted, we would therefore require in the region of 4,500-5,000 fieldworker shifts per year (90-100 per week),** spread across the country.
- There will be variations within this by area, time of day, and type of station, and some of these shifts would be conducted on board rather than at stations (potentially reducing the overall number required), but this initial estimate could be posed as guidance for research agencies to consider in their bids, if and when a tender is issued.
- **An alternative might be to conduct the core fieldwork at A-C stations only.** According to our previous investigation of sampling options, this would have the potential to reach between 64% and 71% of all passenger journeys, rather than 75-84% using A-D stations – so more of the fieldwork would need to be conducted on board to ensure good representation of smaller stations. **This could work at a rate of around 12 completed surveys per shift, meaning the total shift resource for a year would be in the region of 4,100-4,200, so somewhat easier to resource.**
- **Both options could be posed to bidding agencies so that a decision can be arrived at based sample coverage implications and associated value for money**

Representing passenger journeys which originate at smaller stations and are excluded from the proposed core approach

Attempting to recruit passengers as they disembarked a train did not work. **Our proposal for a future survey would be to focus only on passengers waiting to board.**

This means **the best way to widen the reach of the survey would be to conduct some of the fieldwork on board trains.**

- On-board recruitment shifts should be targeted at those routes where smaller proportions of all passenger journeys are represented by the A-D stations, and as far as possible the selection of these journeys should be systematic.
- As far as possible the task of determining (productive) return trips for fieldworkers should be automated, using timetable databases

transportfocus

** See page 10 for more information on station categories.*

Summary: conclusions for more tactical aspects of the method

Fieldwork

- The recruitment approach worked reasonably well during the evenings (7-10pm). Providing it can be made more productive overall – and providing that safety of fieldworkers can be managed well – **we would prefer that fieldwork should take place at least between 7am (or 6am) and 10pm**, as it did for the NRPS.

- Building on the findings from our previous pilot, which tested the broad principle of recruiting face to face at stations, with most passengers taking part online, **all of the access methods used in this trial could be used, except for the URL option which brought little additional value.**

  We would continue to work with any research supplier(s) to maintain and/or improve on efficacy of each method where possible. This would include refining the collection of email addresses, where respondents are willing, from those who prefer "non-contact" access methods into the survey in the first place.

- **We would also work with future supplier(s) to maximise the value of the questionnaire itself,** including:
  - Maximising the proportion of respondents who complete a larger set of questions beyond the core, if the proposal for a modular questionnaire structure is adopted
  - **Improving the accuracy of journey details provided by passengers** so that they evaluate a specific, known journey leg for which results can be correctly attributed to a route or TOC
  - Improving compliance from respondents to complete the questionnaire once they have experienced the precise journey (leg) in question, rather than filling in some or all of it based on expectations from previous experience – but without hindering response rate.

Data preparation and analysis

- As in our previous surveys for rail passenger experience, the method trialled here does have some response bias, particularly towards females, and so **some weighting would likely be needed in a future, full scale survey using this approach.**

- **The method tested here for collecting demographic data to inform this weighting has been successful, and so we would propose to use this in the future.**

- **We would also need some further investigation of other universe profiling data, especially by journey purpose** (perhaps using ticket type as a proxy for this), to better understand how representative the generated sample is, and how to weight it if needed.

transportfocus

Next steps

Transport Focus is currently working with GBR, Network Rail and RDG, on behalf of the DfT, to develop and tender a new measurement tool for rail passenger experience, which will account for some of the different perspectives but ultimately shared overall goals of these collaborating partners.

The exploratory work we have done so far, including learnings from this pilot, will contribute to this.

Dependent on the method settled upon for the new joint survey programme, it is likely that more work will be required to determine how to weight the responses, especially by route or region, journey purpose, and possibly time of day and day of the week. This has not been possible up to this point, as passenger travel behaviour has been somewhat in flux since the beginning of the Covid-19 pandemic.

Contact Transport Focus

Any enquiries about this report should be addressed to:

David Greeno
Senior insight advisor
david.greeno@transportfocus.org.uk

Transport Focus
Albany House
86 Petty France
London
SW1H 9EA
www.transportfocus.org.uk

Transport Focus is the operating name of the Passengers' Council

Transport Focus is the independent consumer organisation representing the interests of:

- bus, coach and tram users across England outside London
- rail passengers in Great Britain
- all users of England's motorways and major 'A' roads (the Strategic Road Network).

We work to make a difference for all transport users.

transportfocus