Rail Passenger Customer Experience Survey – Final Report of Response Rate Experiments

17 November 2023

Contents

1. Introduction

In summer 2022, the rail industry commissioned a method review with the objective to identify the optimal methodological approach to deliver an ongoing customer experience survey amongst rail passengers in Britain.

Field trials were conducted from April to June 2023 with the top methodologies that emerged from the method review. A recommendation report was produced with detailed findings from the trials and a recommended approach for any future the future ongoing survey.

The field trials revealed that response rate was lower than expected. As a result of this the DfT decided to conduct further response rate experiments (RREs) to measure different approaches of how the response rate could be improved. The overall objective of the RREs is to determine the efficacy of four isolated methodological interventions on increasing the response rate of the future Rail Customer Experience.

2. Methodology

The findings from the field trials showed that the on board methodology produced a greater number of completes compared to the at station methodology. Despite performing better overall, the on board methodology had a lower response rate than was expected (17% compared to an expected 30% based on previous work done by Transport Focus). If no further experiments were carried out and the on board approach was incorporated, a greater number of shifts, and as such budget, would be required in order to reach the 5,600 completes for one rail period, which is what the DfT is hoping to achieve in future surveys.

The testing of the four interventions was conducted using an on board intercept approach offering email or QR code equally as response options (and base it on natural fallout). From the field trials this was the approach that achieved the higher number of completes per shift and as a result was more cost effective based on cost per complete.

Each intervention was tested on its own to clearly understand how it contributes towards overall response rate.

We conducted 30 shifts for each intervention which add up to a total of 120 shifts.

With regard to sampling, we have reused journeys from the sampling plan of the field trials. In some cases (24 shifts) the RRE sampling incorporated duplicates of the same shift from the field trials. This is because we focused on TOCs that made it easiest to provide permissions for working on their services, as the fieldwork needed to be completed so that possible changes can be made to the pilot survey. These duplicates were kept separate from each other by applying a different intervention type to each and not running the shifts at the exact same time and date.

The approach of reusing shifts from the field trials also allows for a direct like for like comparison between what was achieved in the field trials (serving as a benchmark) and what has been achieved in the RREs, controlling variables as much as possible. In the RRE analysis we have only compared shifts with the equivalent shifts used in the field trials.

We recruited the same fieldworker for the experiments who was used in the field trials for 68 shifts.

The shifts for the RREs had a good spread across days of the week and time of day and covered the following TOCs:

- Avanti West Coast,
- East Midlands Railway,
- Thameslink,
- Heathrow Express,

- Merseyrail,
- Northern, and
- Transport for Wales.

The TOCs also represent a good spread across TOC types (urban, commuter, long distance, etc.).

3. Intervention specifications

The following interventions were tested in the RREs:

3.1 Blue shifts: Removing instruction to complete at end of journey.

This intervention includes removing the instruction to complete the survey at the end of journey both from the fieldworkers' instructions and the survey script.

Interventions to prevent respondents from completing the survey before the end of their journey had a negative impact on completes per shift in the field trials. In addition, anecdotal feedback from fieldworkers revealed that such an instruction put potential respondents off from completing the survey (which was supported by feedback in the qualitative element of the field trials).

3.2 Green shifts: Shorter 5-6 minute survey

In this intervention the survey was shortened to roughly five minutes (even though the median turned out to be slightly longer than four minutes – see later on) developed by the DfT project team and this was communicated by the fieldworker at the time of recruitment, in the invitation emails (if email was selected) and in the introduction of the online survey.

In the field trials a notable proportion dropped out of the survey just before halfway through the survey which implies that the survey was too long. The qualitative element supported this view.

3.3 Red shifts: UX modifications

This intervention includes having some modifications applied throughout the survey, the recruitment script, the email invites and reminders, to enhance the user experience. The BVA Nudge Consulting team was involved in this intervention. They proposed some enhancements and nudges to be applied to the survey script and the invites and reminders for a more positive experience to boost response rate. Fieldworkers were also asked to wear high vis jackets when working on red shifts to add legitimacy to the survey.

3.4 Yellow shifts: Completion incentive

This intervention includes offering respondents the chance to win one of five £200 shopping vouchers. For maximum take up, the prize draw is mentioned during recruitment, in the email invitations and reminders, in the introduction text of the survey and at the end of it when we ask if people wanted to participate in the draw.

4. Outputs

Table 1: Comparisons of response rates and dropout rates between field trials and response rate experiments

Experiment	Field trials vs. RRE - Response rate											
	Removing instruction to complete at end			Short survey five min			UX enhancements			Incentive		
	Field trials	RRE		Field trials	RRE		Field trials	RRE		Field trials	RRE	
# Shifts	**30**	**30**		**30**	**30**		**30**	**30**		**30**	**30**	
Total recruited	**1480**	**1628**		**1555**	**1634**		**1373**	**1650**		**1411**	**1603**	
QR	1132	1355		1107	1295		1033	1341		1105	1287	
Email	348	273		448	339		340	309		306	316	
Total complete	**242**	**301**		**220**	**437**		**251**	**327**		**221**	**341**	
QR	175	248		161	367		184	264		167	270	
Email	67	53		59	70		67	63		54	71	
Dropout	**794**	**717**		**721**	**699**		**695**	**733**		**823**	**754**	
QR	763	684		684	629		664	699		795	723	
Email	31	33		37	70		31	34		28	31	
Total response rate	**16%**	**18%**	**2%**	**14%**	**27%**	**13%**	**18%**	**20%**	**2%**	**16%**	**21%**	**6%**
QR	15%	18%	**3%**	15%	28%	**14%**	18%	20%	**2%**	20%	21%	**1%**
Email	19%	19%	**0%**	13%	21%	**7%**	20%	20%	**1%**	23%	22%	**-1%**
Dropout rate	**54%**	**44%**	**-10%**	**46%**	**43%**	**-4%**	**51%**	**44%**	**-6%**	**58%**	**47%**	**-11%**
QR	52%	**42%**	**-10%**	44%	**38%**	**-5%**	48%	**42%**	**-6%**	56%	**45%**	**-11%**
Email	**2%**	**2%**	**0%**	**2%**	**4%**	**2%**	**2%**	**2%**	**0%**	**2%**	**2%**	**0%**
Recruits per shift	49.3	54.3	4.9	51.8	54.5	2.6	45.8	55.0	9.2	47.0	53.4	6.4
QR	37.7	45.2	7.4	36.9	43.2	6.3	34.4	44.7	10.3	36.8	42.9	6.1
Email	11.6	9.1	- 2.5	14.9	11.3	- 3.6	11.3	10.3	- 1.0	10.2	10.5	0.3

Completes per shift	8.1	10.0	2.0	7.3	14.6	7.2	8.4	10.9	2.5	7.4	11.4	4.0
QR	5.8	8.3	2.4	5.4	12.2	6.9	6.1	8.8	2.7	5.6	9.0	3.4
Email	2.2	1.8	- 0.5	2.0	2.3	0.4	2.2	2.1	- 0.1	1.8	2.4	0.6
Total complete in time	**130**	**120**		**145**	**157**		**155**	**170**		**137**	**170**	
Total completes early	**74**	**155**		**47**	**221**		**52**	**113**		**45**	**112**	
Total completes unknown (no JPT info)	*38*	*26*		*28*	*59*		*44*	*44*		*39*	*59*	
Total share of completes in time (from known arrival time)	**64%**	**44%**	**-20%**	**76%**	**42%**	**-34%**	**75%**	**60%**	**-15%**	**75%**	**60%**	**-15%**
Total verified completes in time response rate	**9%**	**7%**	**-1%**	**9%**	**10%**	**0%**	**11%**	**10%**	**-1%**	**10%**	**11%**	**1%**

xx%/xx% significantly lower/higher than the field trials

Differences for completes per shift in red font are significantly different.

Key take outs:

- Generally higher response rate in RREs; considerably more completes per shift for short questionnaire; but overall a higher proportion in RREs that complete before journey end.
- Higher completes per shift for QR are due to the higher number of recruits per shift amongst those selecting QR (not the case for email).
- Looking at this in detail by intervention, there were specifics that led to the increase in completes per shift:
    - For the intervention where instructions to complete at the end of the journey were removed, we saw a 23% uplift in completes per shift. The uplift was driven by an increase in the numbers recruited per shift and fewer dropouts. However, removing this instruction had an impact on those completing the survey being in time, with a -20% point reduction from 64% in the field trails to 44% in the RRE based on their known arrival time.
    - Answering a shorter survey led to the number of completes per shift doubling. The recruitment increased compared to the field trials, including an increase in the number of QR recruits. The QR code response rate was ultimately the key driving force behind the increase in completes overall for this intervention. This intervention had a significant impact on the share of completes in time, with 42% completing the survey in time compared to 76% recorded in the field trials.

- The UX enhancements employed at the recruitment stage and throughout the main survey led to a 30% uplift in the number of completes per shift. The recruitment was the key driver behind the increase in completes per shift, predominantly from QR code.
- Incentivising the survey led to a 54% uplift in the number of completes per shift. Recruitment increased for this intervention, however the uplift in completes per shift was ultimately led by response rate, where fewer were dropping out compared to the field trials.

Table 2: Demographic profile of respondents of field trials compared with response rate experiments

Field trials vs. RRE - Profiling												
Experiment	**Removing instruction to complete at end**			**Short survey five min**			**UX enhancements**			**Incentive**		
	Field trials	**RRE**		**Field trials**	**RRE**		**Field trials**	**RRE**		**Field trials**	**RRE**	
Base size	**242**	**301**		**220**	**437**		**251**	**327**		**221**	**341**	
Gender												
Male	54%	49%	-5%	40%	47%	7%	41%	47%	6%	54%	50%	-4%
Female	45%	48%	2%	58%	52%	-6%	56%	49%	-7%	45%	46%	1%
Other	1%	3%	2%	2%	2%	-1%	3%	4%	1%	1%	4%	3%
Age												
16-24	14%	17%	2%	14%	23%	9%	17%	21%	5%	17%	20%	3%
25-34	23%	23%	1%	25%	25%	0%	21%	27%	6%	21%	25%	4%
35-44	23%	19%	-4%	21%	19%	-2%	18%	16%	-2%	18%	**20%**	2%
45-54	12%	14%	2%	16%	14%	-2%	17%	14%	-3%	18%	12%	-6%
55-64	14%	16%	1%	16%	9%	-7%	16%	13%	-3%	13%	16%	2%
65+	12%	10%	-2%	6%	8%	2%	10%	7%	-3%	13%	7%	-6%
Prefer not to say	1%	1%	1%	1%	1%	0%	1%	1%	0%	0%	1%	0%
Disabled	**11%**	**8%**	**-3%**	**10%**	**10%**	**0%**	**13%**	**9%**	**-4%**	**12%**	**11%**	**-2%**

xx%/xx% significantly lower/higher than the field trials

Key take outs:

- Generally younger profile in RREs; mixed differences across interventions regarding gender; no major differences with regard to proportion of disabled respondents.

Table 3: Comparison of satisfaction scores and net promoter scores of field trials compared with response rate experiments

Field trials vs. RRE - Satisfaction												
Experiment	**Removing instruction to complete at end**			**Short survey five min**			**UX enhancements**			**Incentive**		
	Field trials	**RRE**		**Field trials**	**RRE**		**Field trials**	**RRE**		**Field trials**	**RRE**	
Base size	**242**	**301**		**220**	**437**		**251**	**327**		**221**	**341**	
Overall satisfaction with journey												
NET Satisfied	90%	85%	-5%	84%	82%	-2%	85%	85%	0%	88%	91%	2%
Very satisfied	49%	46%	-3%	49%	43%	-5%	51%	47%	-4%	54%	53%	-1%
NET Dissatisfied	4%	6%	2%	7%	7%	0%	6%	6%	0%	5%	3%	-3%
Very dissatisfied	2%	**2%**	0%	**3%**	**3%**	**0%**	**2%**	**1%**	**-1%**	**2%**	**1%**	**-1%**
Overall satisfaction with departure station												
NET Satisfied	86%	85%	-1%	87%	79%	-8%	86%	83%	-3%	90%	88%	-2%
Very satisfied	50%	45%	-5%	51%	40%	-11%	50%	48%	-2%	51%	49%	**-2%**
NET Dissatisfied	6%	3%	-3%	7%	6%	0%	6%	4%	-1%	4%	4%	0%
Very dissatisfied	1%	1%	0%	3%	2%	-1%	1%	2%	0%	1%	0%	-1%
NPS	24	29	4.2	39	26	-13.2	30	19	-10.6	35	40	4.9
Promoters	46%	48%	2%	56%	46%	-9%	50%	43%	-7%	53%	52%	-1%
Passives	32%	32%	-12%	27%	33%	6%	29%	32%	3%	29%	36%	6%
Detractors	22%	20%	10%	17%	21%	**4%**	20%	24%	**4%**	18%	12%	**-6%**

xx%/xx% significantly lower/higher than the field trials

Key take outs:

- Lower satisfaction with overall journey amongst those completing the short questionnaire compared to other interventions; overall satisfaction with the departure station considerably lower amongst short survey respondents compared to field trials and also other interventions; NPS variance mixed but also notably lower amongst those completing the short questionnaire.

Table 4: Satisfaction with departure station from the five minute survey, by age and gender

	Short survey - five mins - Satisfaction by age and gender							
Overall satisfaction with departure station	Female	Male	16-24	25-34	35-44	45-54	55-64	65+
NET Satisfied	83%	75%	74%	81%	78%	83%	82%	83%
Very satisfied	47%	33%	33%	45%	37%	43%	41%	47%
NET Dissatisfied	4%	9%	5%	7%	10%	6%	5%	0%
Very dissatisfied	1%	3%	2%	2%	2%	3%	0%	0%

Table 5: Satisfaction with overall journey from the five minute survey, by age and gender

	Short survey - five mins - Satisfaction by age and gender							
Overall satisfaction with journey	Female	Male	16-24	25-34	35-44	45-54	55-64	65+
NET Satisfied	82%	83%	82%	83%	78%	78%	88%	94%
Very satisfied	51%	37%	36%	45%	38%	49%	53%	53%
NET Dissatisfied	6%	7%	4%	8%	9%	8%	8%	3%
Very dissatisfied	2%	2%	1%	2%	4%	5%	3%	0%

All base sizes >30

Key take outs:

- Lower satisfaction with departure station amongst males and the youngest respondents; satisfaction with the overall journey is balanced by gender and somewhat lower amongst 35-54 year olds.

Table 6: Age and gender profile of footfall counts of field trials compared with response rate experiments

Field trials vs. RRE - Footfall count												
Experiment	**Removing instruction complete at the end**			**Short survey five min**			**UX enhancements**			**Incentive**		
	Field trials	**RRE**		**Field trials**	**RRE**		**Field trials**	**RRE**		**Field trials**	**RRE**	
Base size	**24**	**24**		**23**	**23**		**23**	**23**		**24**	**24**	
Footfall count	**1479**	**1808**		**1107**	**1595**		**1410**	**1901**		**1395**	**2133**	
Footfall per shift	**62**	**75**	**14**	**48**	**69**	**21**	**61**	**83**	**21**	**58**	**89**	**31**
Female	48%	46%	-2%	48%	47%	-1%	50%	46%	-4%	45%	45%	1%
Male	52%	54%	2%	52%	53%	1%	50%	54%	4%	55%	55%	-1%
16-34 years	32%	43%	10%	37%	38%	0%	45%	40%	-4%	42%	45%	3%
35-64 years	54%	50%	-5%	52%	53%	1%	46%	49%	3%	48%	50%	2%
65+ years	13%	8%	-5%	11%	9%	-2%	10%	11%	1%	10%	6%	-5%

Key take outs:

- Notably higher footfall per shift compared to the equivalent shifts in the field trials. There were distinctly more 16-34 year olds for the shifts where the instruction to complete at the end was removed. This did not correlate to a significant increase in the number who completed aged between 16-34 as shown earlier in the report.
- The actual profile for shorter survey respondents is quite a bit younger compared to the footfall profile. In the actual sample 48% of the short survey respondents are 16-34 years old whereas in the footfall count profile 38% fall into that age bracket. This is much more balanced for the other experiments although for the UX enhancements the actual sample is also 48% for 16-34 year olds compared to 40% in the footfall count.

Table 7: Time to complete questionnaire for different interventions

	Short survey five min	Removing instruction to complete at end	UX enhancements	Incentive
Median	04:04	13:14	13:25	14:41

Key take outs:

- The incentive route does not generate speeders. Generally, the length of the survey was shorter than in the field trials which might be, in part, due to the younger profile in the RREs.

Table 8: Comparison of costs for achieving minimum samples between field trials and response rate experiments

	Field trials	Removing instruction to complete at end	Short survey five min	UX enhancements	Incentive
Increase in completes per shift	-	+24%	+99%	+30%	54%
Number of shifts required for 5600 completes	374	301	188	287	242
Total number of shifts for four rail periods (excluding boosts)	1,494	1,204	750	1,149	970
Total shift costs (@£285/shift) - four rail periods	£425,790	£343,226	£213,869	£327,385	£276,364

Key take outs:

- When looking at the pilot survey, we originally estimated that we needed 374 shifts to reach 5,600 completes.

- With the short survey, we would only need 188 shifts which leads to the largest cost saving in shift cost.

Interviewer impact:

The following table shows how interviewers that only worked in the field trials performed compared to interviewers that worked in the field trials as well as the RREs. We compared number of recruits, number of completes and based on these two metrics, the response rate. The reason for this comparison was to get an understanding, if by chance, the interviewers used in the RRE were generally better at getting recruits and completes because that would have impacted on overall numbers.

Table 9: Comparison of interviewer performance in field trials and response rates experiments

	Number of recruits in field trials
Field trial interviewers (only)	8495
Interviewers used in both field trials and RREs	7141
	Number of completes in field trials
Field trial interviewer (only)	1463
Interviewers used in both field trials and RREs	1119
	Response rate in field trials
Field trial interviewer (only)	17%
Interviewers used in both field trials and RREs	16%

Key take outs:

- There is no indication that the interviewers we used for the RREs were better performing than the others we had used in the field trials. When looking at their performance in the field trials, they achieved even a slightly lower response rate than the other interviewers. From this we can infer that results are not influenced by interviewers' performance.

5. Our recommendations

Based on the results of the four interventions, there is a strong case to introduce a combination of enhancements to help improve response rate for the any future continuous survey. The UX enhancements showed to have a positive impact on recruitment and ultimately the number of completes per shift. These enhancements would be simple to implement and thus based on this we would recommend employing them in the future survey. The future pilot survey would also benefit from the introduction of an incentive. The results show they improve completes per shift even further (an uplift of 54%). The introduction of incentive may seem costly at first, but the cost per shift shows how the monthly cost of an incentive would be easily repaid by an increase in responses. Our recommendation would be to incorporate both the UX enhancements and the incentive to help boost the total number of completes per shift. These two interventions also had the least impact on completion before the end of the journey.

Focusing on the intervention where we removed the instruction to complete the survey at the end of the journey, our recommendation would be not to implement this intervention in the future survey. Our reasoning is based on a higher number of respondents completing the survey before they reach their arrival station and subsequently impacting the accuracy of the answers given. On a similar vein, the shorter 5 minute survey has the lowest number of respondents completing on time, likely due to the length of the survey, and would be our main reason for not going with this option. Nevertheless, this was an intervention where completes per shift increased greatly and this provides clear evidence of the positive impact on the completes per shift that is achieved with a shorter survey. From this we can infer that shortening the survey is likely to have an impact in the future and we recommend looking at ways to cut down the questionnaire for the any future ongoing survey if possible.

6. Appendix

Table 10: Age breakdown of Net Promoter Score for the short survey and user experience enhancements experiments

	Short survey five min							UX enhancements						
	16-24	25-34	35-44	45-54	55-64	65+	Prefer not to say	16-24	25-34	35-44	45-54	55-64	65+	Prefer not to say
NPS	26	23	23	24	24	base too low	base too low	9	22	21	9	26	base too low	base too low
Promoters	46%	44%	47%	48%	46%	base too low	base too low	36%	46%	47%	36%	42%	base too low	base too low
Passives	35%	34%	30%	29%	32%	base too low	base too low	37%	30%	26%	38%	42%	base too low	base too low
Detractors	20%	22%	23%	24%	22%	base too low	base too low	27%	24%	26%	27%	16%	base too low	base too low

Table 11: Gender breakdown of Net Promoter Score for the short survey and user experience enhancements experiments

	Short survey		UX enhancements	
	Female	Male	Female	Male
NPS	30	22	34	5
Promoters	51%	41%	50%	37%
Passives	28%	40%	34%	31%
Detractors	21%	19%	16%	32%