Transport Focus

Project Spotlight

HEADLINE RESULTS

DECEMBER 2019

Populus

Contents

1. Pilot performance

2. Overview of pilot data

3. Recommended design

4. Optimising the design

Pilot performance overview

Populus

Overview of Our Approach

Recruitment Pilot

6 pilot approaches

Face 2 face

PILOT 1:
Face to face interviewers located in station

PILOT 2:
Face to face interviewers on train

Online open link

PILOT 3:
Online survey to TF NRPS recontact panel

PILOT 4:
Online survey activated via on train Wi-Fi

Online panel

PILOT 5:
Daily online omnibus sample

PILOT 6:
Ad hoc online panellists

→

Recruit survey

10 question survey

Consistent micro survey for all pilots

- Recruitment
- Journey diary on most recent rail journey

All asked to participate in an ongoing diary with email address taken

→

1 month diary

Micro survey journey diary

4 week open diary

Reminder emails sent sent twice a week for 1 month (8 contacts), to capture their most recent rail journey

→

Dashboard

Visualising the data

All recruitment data and diary data uploaded into a dashboard

Summary of pilot approaches

	Recruitment Pilot	Recruit survey	1 month diary
Face 2 face (prize draw incentive)	PILOT 1: 9 X Face to face interviewers shifts in station 2 Waterloo, 3 Euston, 2 Liverpool St, 2 BNS*	Target a total minimum 200 face to face interviews	For those providing email address and agreeing to participate in diary
	PILOT 2: 10 X Face to face interviewers on train 5 shifts on Northern Route and 5 shifts GWR	Target minimum 200 face to face interviews	For those providing email address and agreeing to participate in diary
Online open link (prize draw incentive)	PILOT 3: Online survey to TF NRPS recontact panel	1749 email addresses invited to participate	All agreeing to participate in diary
	PILOT 4: Online survey activated via on train Wi-Fi	Note not progressed due to TOC based technical/logistic issues	
Online panel (panel incentive)	PILOT 5: Online omnibus survey sample, 2000 Nat Rep	Target recruit 200 people travelled in the past week	All recruited based on agreeing to participate in diary
	PILOT 6: Ad hoc online panellists – direct recruitment	Direct panel recruit of 100 people travelled in past week	All recruited based on agreeing to participate in diary

*we were unable to complete the final shift in Birmingham

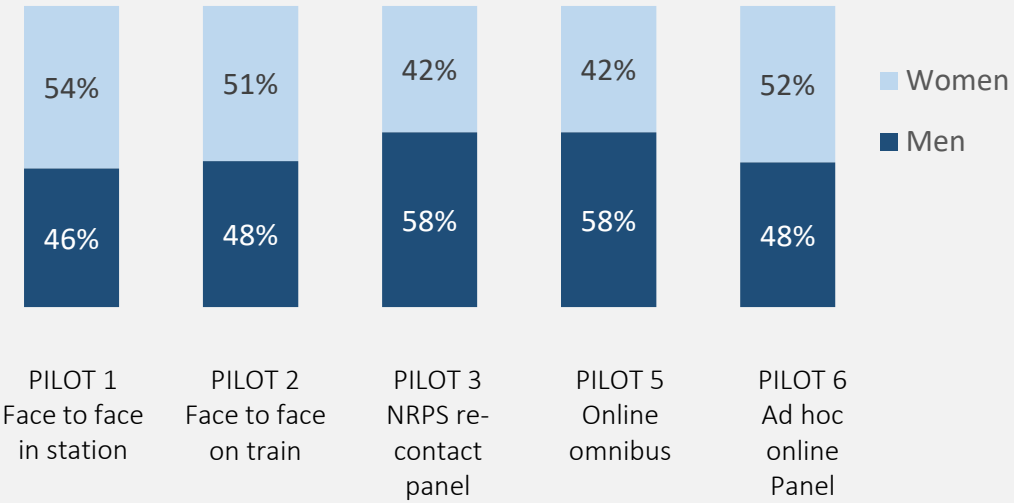
Summary of pilot performance

		PILOT 1 Face to face in station	PILOT 2 Face to face on train	PILOT 3 NRPS re-contact	PILOT 5 Omnibus sample	PILOT 6 Ad hoc online panellists
Recruitment	Number of people recruited	204	252	(1749 emails sent) 326 (19%)	(2000 Nat Rep) 320 past week (16%)	128
	Number of emails shared	92 (45%)	103 (41%)	253 (78%)	235 (73%)	100 (78%)
Online diary	Diary length	5 weeks	3 weeks	5 weeks	5 weeks	5 weeks
	Number of people completing online diary any week	12 (9%)	33 (32%)	185 (73%)	191 (81%)	80 (80%)
	Total number of journeys collected	27	66	735	2195	856
	Average numbers of journeys per active person per week	1.4	1.4	1.5	3.1	3.1
	Recommendation to use	NO	NO	YES	YES	YES

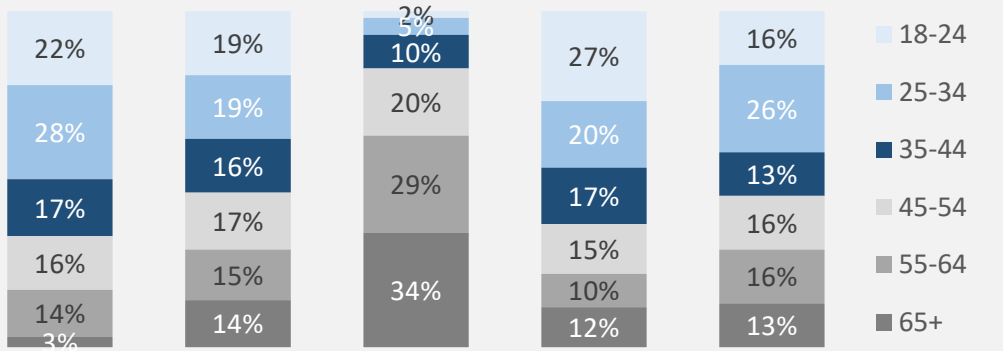
(xx%) = % conversion between levels

Profile of people recruited – all taken a journey in the last week

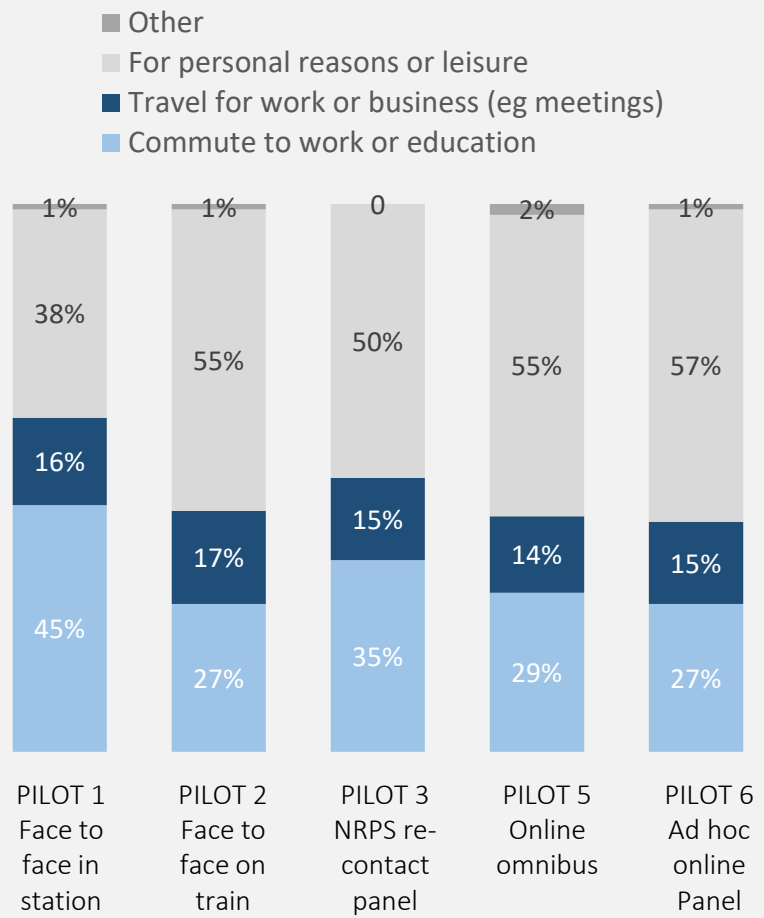
Gender



Age



Journey Purpose



Base: All who were recruited and took a journey in the last week.
Pilot 1 (180), Pilot 2 (247), Pilot 3 (242), Pilot 5 (308), Pilot 6 (128)

Number of active respondents by week

	PILOT 1		PILOT 2		PILOT 3		PILOT 5		PILOT 6	
	Face to face in station	% recruit active	Face to face on train	% recruit active	NRPS re-contact	% recruit active	Online omnibus	% recruit active	Ad hoc online	% recruit active
Recruitment	**92**		**103**		**253**		**235**		**100**	
Week 1	3	3%	16	16%	73	29%	165	70%	56	56%
Week 2	6	7%	19	18%	106	42%	148	63%	62	62%
Week 3	3	3%	9	9%	74	29%	126	54%	57	57%
Week 4	4	4%	6	6%	68	27%	107	46%	50	50%
Number completing any week	12	13%	33	32%	185	73%	191	81%	80	80%

Number of recorded journeys by week

	PILOT 1		PILOT 2		PILOT 3		PILOT 5		PILOT 6	
	Face to face in station	week on week drop	Face to face on train	week on week drop	NRPS re-contact	week on week drop	Online omnibus	week on week drop	Ad hoc online	week on week drop
Recruited people	**92**		**103**		**253**		**235**		**100**	
TOTAL JOURNEYS LOGGED										
Week 1	3		20		181		698		237	
Week 2	10	333%	21	105%	123	68%	448	64%	234	99%
Week 3	4	40%	16	76%	115	93%	329	73%	157	67%
Week 4	4	100%	9	56%	96	83%	272	83%	115	73%

Summary: pilot performance overview

- **Significant differences by pilot** in terms of recruitment and ongoing diary participation/entries
- The **Interviewer led i**n station and on train pilots are not viable given low diary participation and very high attrition between recruitment and the diary stage
    - In station sample skews to commuters reflecting time of interviewer shift
- The **NRPS recontact pilot** showed a relatively high email conversion rate, high diary participation with low week on week attrition. Considerably older sample.
- The **online panel omnibus and direct recruitment** approaches provided the highest volume of diary entries and lowest week on week attrition. Good age/gender spread
- **Recommendation:** use a combination of panel and NRPS recontact approaches

Overview of pilot data

NB. This is looking at journeys taken in the last week

Populus

Analysing the pilot results

1

What does journey satisfaction look like across the pilots?

2

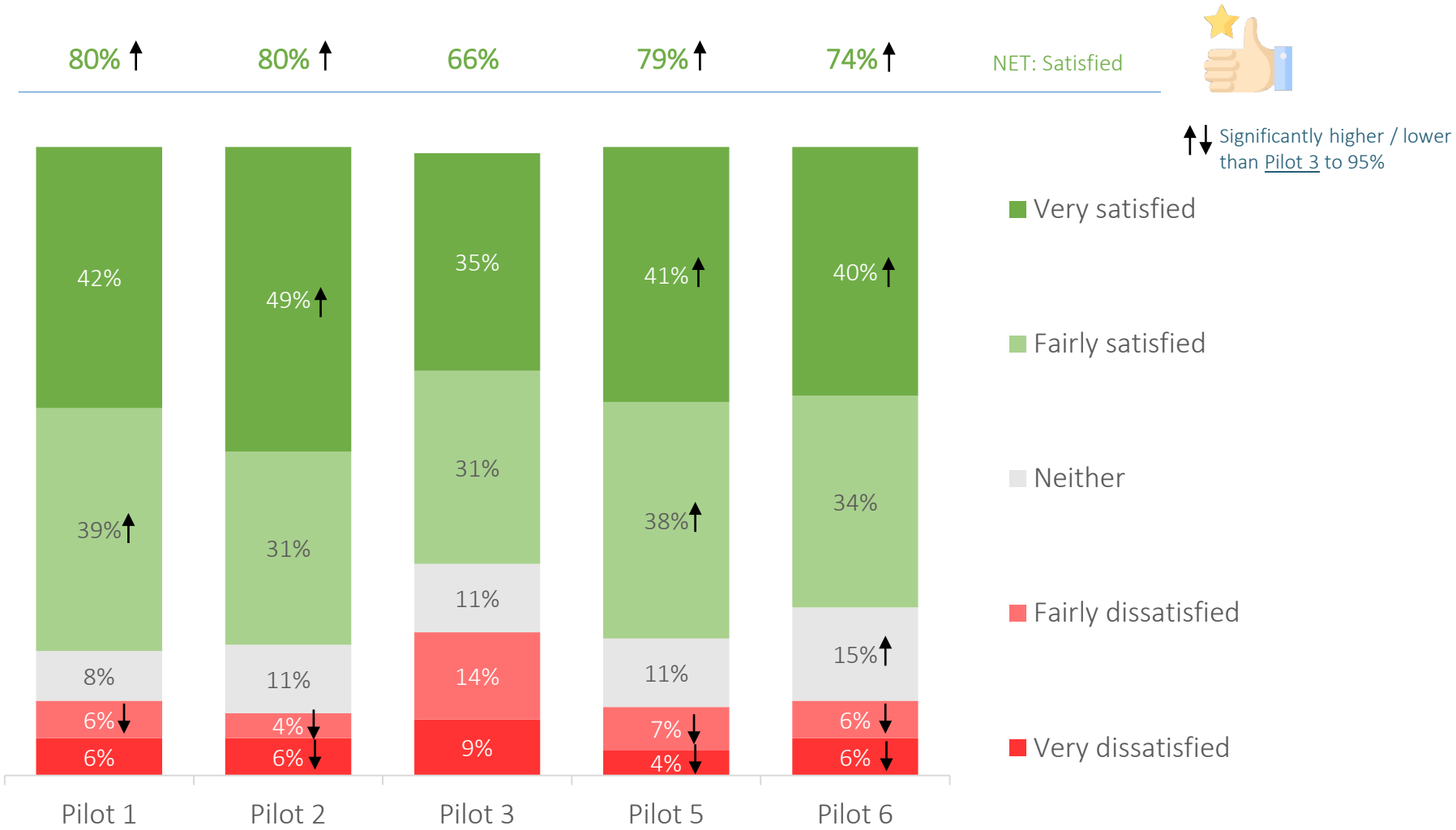
Do the pilots deliver a representative view of journeys taken in the last week?

3

How does pilot satisfaction compare to current NRPS results?

Journey satisfaction is high and consistent across pilots 1, 2 and 3. Journeys from pilot 3 aren't as positive

Overall satisfaction



Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base: All who took a journey in the last week. Total (4977), Pilot 1 (204), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984). NB. Spotlight data is Unweighted

There is a good spread of TOC usage across pilots 3, 5 and 6. Given the on train interviewing, there was very little variation in TOC usage in pilot 2

TOC's used for journeys: share of journeys

	PILOT 1 Face to face in station	PILOT 2 Face to face on train	PILOT 3 NRPS re-contact	PILOT 5 Omnibus sample	PILOT 6 Ad hoc online panellists
Northern	1%	**51%**	**11%**	**10%**	**11%**
Great Western Railway	2%	**40%**	9%	6%	**10%**
Southeastern	5%	1%	7%	**10%**	7%
South Western Railway	**16%**	1%	9%	5%	**12%**
ScotRail	-	-	6%	**10%**	4%
West Midlands Trains	**16%**	1%	5%	8%	8%
London Overground	**11%**	-	1%	7%	8%
Southern	6%	-	3%	6%	6%
Greater Anglia	**13%**	1%	3%	5%	8%
CrossCountry	5%	-	7%	5%	5%
Thameslink	2%	*	3%	4%	3%
c2c	1%	-	5%	2%	4%
Great Northern	1%	1%	2%	3%	2%
TfL Rail	2%	-	3%	3%	1%
Merseyrail	-	-	3%	2%	*
Other	**17%**	3%	**22%**	**14%**	9%

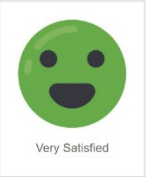
Populus

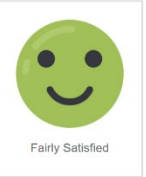
Q7. Which of these train companies did you use for this journey? (If more than one, which did you spend most time on?) Base: All who took a journey in the last week. Total (4977), Pilot 1 (204), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984) NB. Spotlight data is Unweighted

transportfocus

Top reasons for journey satisfaction is fairly consistent across pilots, although cleanliness is more top of mind for pilots 2 and 3

Reason for journey satisfaction (top 2 box) – top 4 topics mentioned





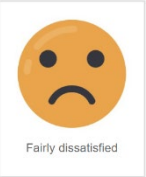
Top topic

PILOT 1 Face to face in station	PILOT 2 Face to face on train	PILOT 3 NRPS re-contact	PILOT 5 Omnibus sample	PILOT 6 Ad hoc online panellists
Punctuality (33%)	Punctuality (47%)	Punctuality (63%)	Punctuality (47%)	Punctuality (40%)
Speed / efficiency (20%)	Capacity / crowding (23%)	Capacity / crowding (31%)	Capacity / crowding (22%)	Pleasant / comfortable (23%)
Capacity / crowding (17%)	Pleasant / comfortable (18%)	Pleasant / comfortable (13%)	Pleasant / comfortable (21%)	Capacity / crowding (18%)
Pleasant / comfortable (14%)	Cleanliness (14%)	Cleanliness (11%)	Speed / efficiency (18%)	Speed / efficiency (16%)

Q9. Why do you feel this way? Base: All who took a journey in the last week and were satisfied / dissatisfied with their journey. Total (3754/649), Pilot 1 (164/24), Pilot 2 (250/30), Pilot 3 (643/226), Pilot 5 (1972/254), Pilot 6 (725/115) NB. Spotlight data is Unweighted

Reason for journey dissatisfaction is also quite similar across pilots

Reason for journey dissatisfaction (top 2 box) – top 4 topics mentioned



Top topic

PILOT 1 Face to face in station	PILOT 2 Face to face on train	PILOT 3 NRPS re-contact	PILOT 5 Omnibus sample	PILOT 6 Ad hoc online panellists
Punctuality (58%)	Punctuality (37%)	Punctuality (54%)	Punctuality (48%)	Capacity / crowding (57%)
Cancellations (21%)	Capacity / crowding (27%)	Capacity / crowding (27%)	Capacity / crowding (33%)	Punctuality (43%)
Quality of trains (17%)	Cancellations (13%)	Cancellations (22%)	Cancellations (13%)	Cost (7%)
Capacity / crowding (13%)	Quality of trains (10%)	Convenience / ease (15%)	Convenience / ease (7%)	Cancellations (7%)

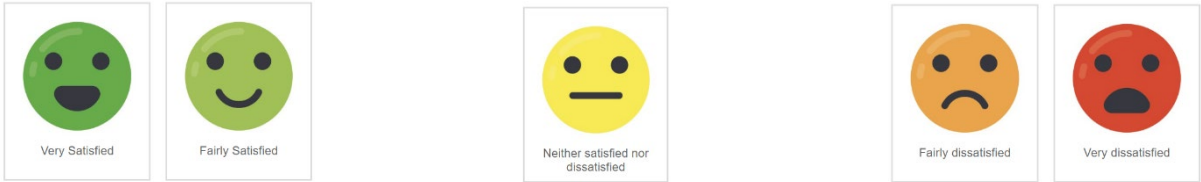
NB – low base size, indicative only

Q9. Why do you feel this way? Base: All who took a journey in the last week and were satisfied / dissatisfied with their journey. Total (3754/649), Pilot 1 (164/24), Pilot 2 (250/30), Pilot 3 (643/226), Pilot 5 (1972/254), Pilot 6 (725/115) NB. Spotlight data is Unweighted

The verbatim sentiment profile intuitively aligns with satisfaction, that said satisfactory journeys can still result in some level of negative response

Journey satisfaction by sentiment of satisfaction verbatim

Similar pattern across all the pilots



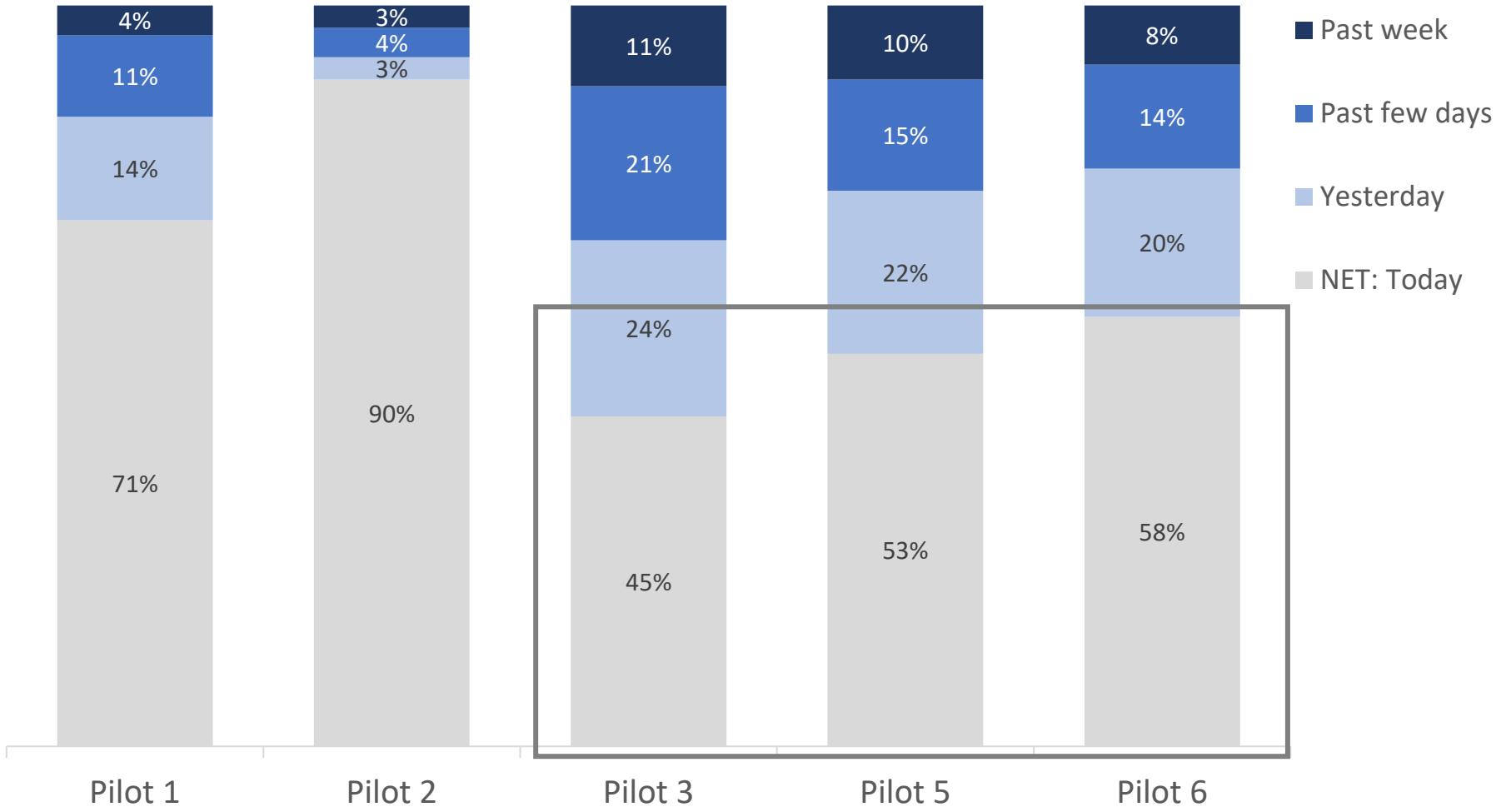
Why do you feel this level of satisfaction?

	Very / fairly satisfied	Neither	Very / fairly dissatisfied
Positive	84%	55%	35%
Neutral	4%	7%	2%
Negative	35%	68%	89%
Average number of sentiment themes	1.2	1.3	1.3

Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Q9. Why do you feel this way? Base: All who took a journey in the last week and satisfied / neutral / dissatisfied with their journey (3754 / 574 / 649) NB. Spotlight data is Unweighted

Given the methodology pilots 1 and 2 pick up the most recent journeys, that said the online pilots are also able to deliver feedback on more recent journey experiences

Journeys in the last week

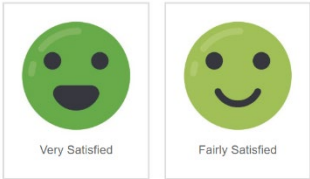


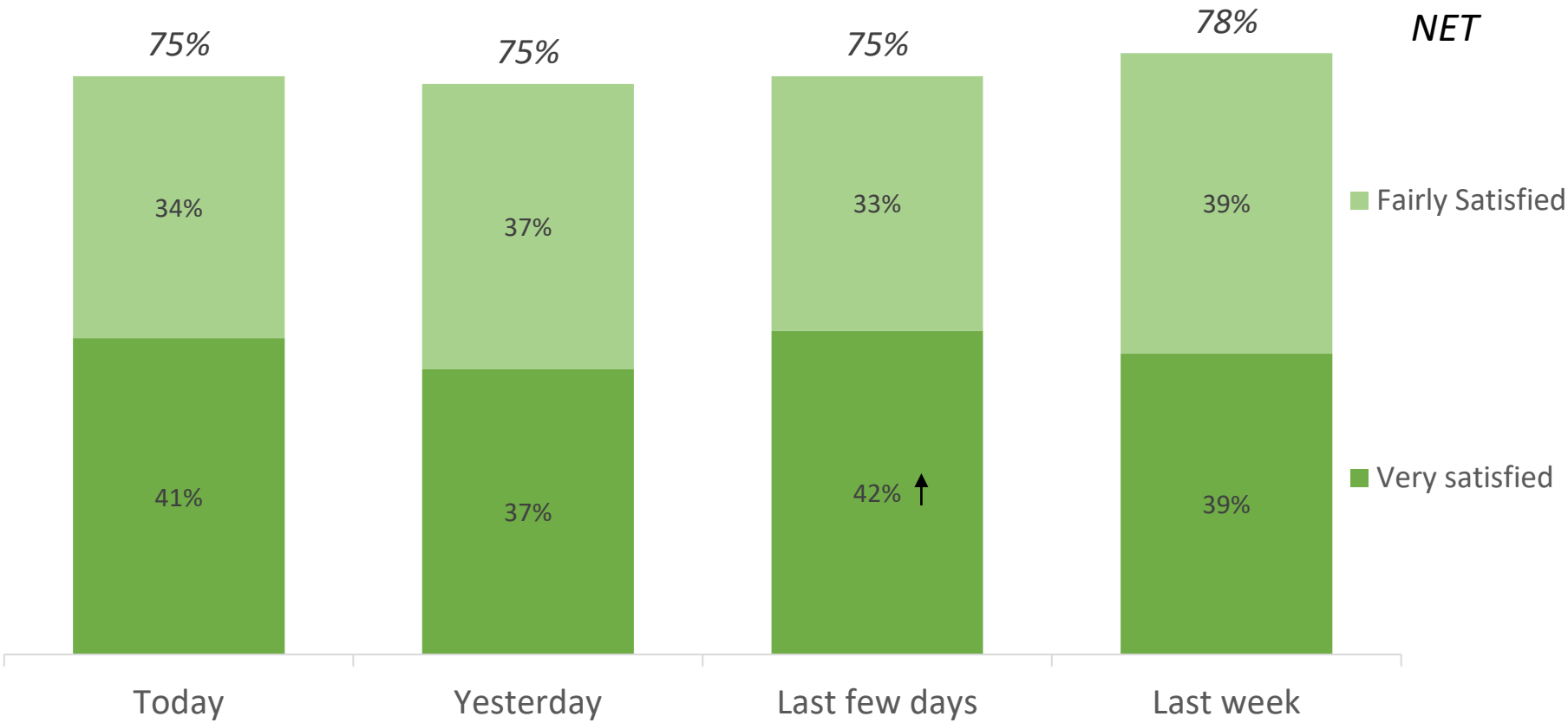
Q3. When did you last travel by rail (excluding London Underground, DLR, or light railway services? Please select one option.
Base: All who took a journey in the last week. Pilot 1 (204), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984)
NB. Spotlight data is Unweighted

Journey satisfaction is fairly consistent whether a journey is experienced today or in the last week

Journey Satisfaction (top 2 box) by recency of journey: all pilots combined

Similar pattern across all the pilots



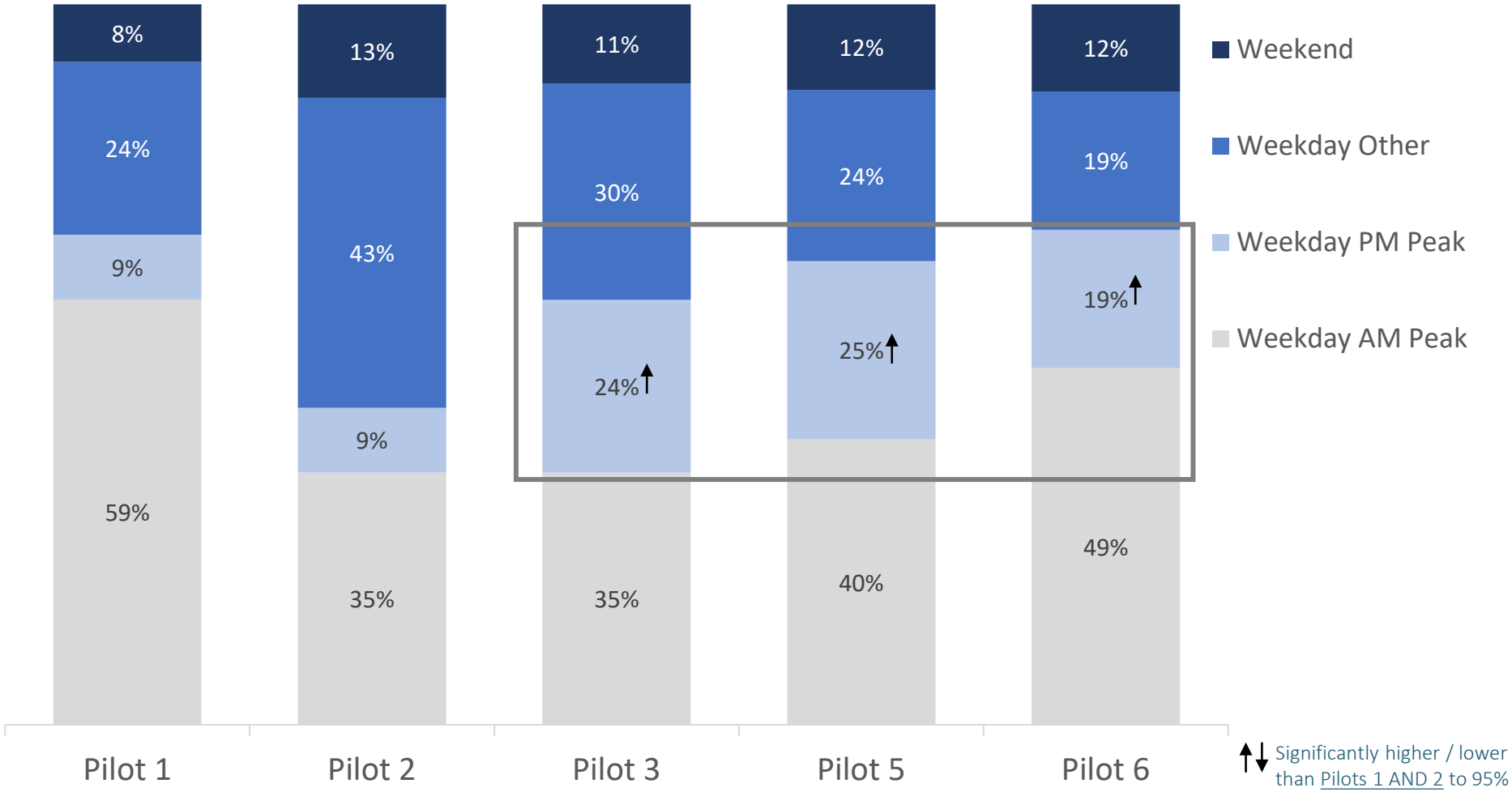


↑ Significantly higher than Yesterday to 95%

Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base: All who took a journey in the last week. Today (2745), Yesterday (1019), Last few days (759), Last week (454) NB. Spotlight data is Unweighted

Online pilots provide a robust read on peak journeys and captures more of the negative weekday PM peak trips

Time of journey

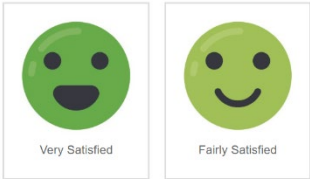


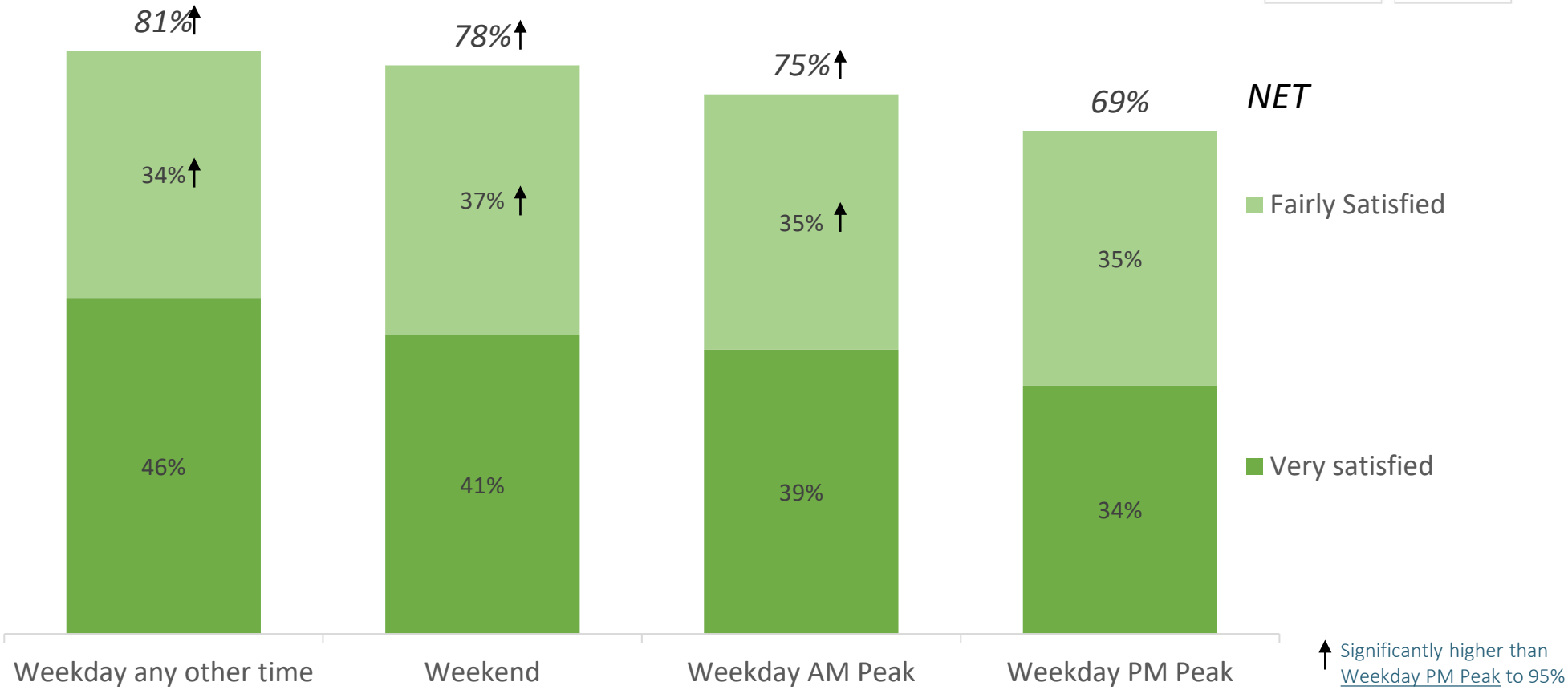
Q5a. When did your rail journey start? Please select one option. Base: All who took a journey in the last week. Pilot 1 (176*), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984). *Question added in when live so not everyone in pilot 1 were asked this question.
NB. Spotlight data is Unweighted

We know that journeys happening outside of peak weekday PM hours tend to be more positive

Journey Satisfaction (top 2 box) by time travelled: all pilots combined

Similar pattern across all the pilots

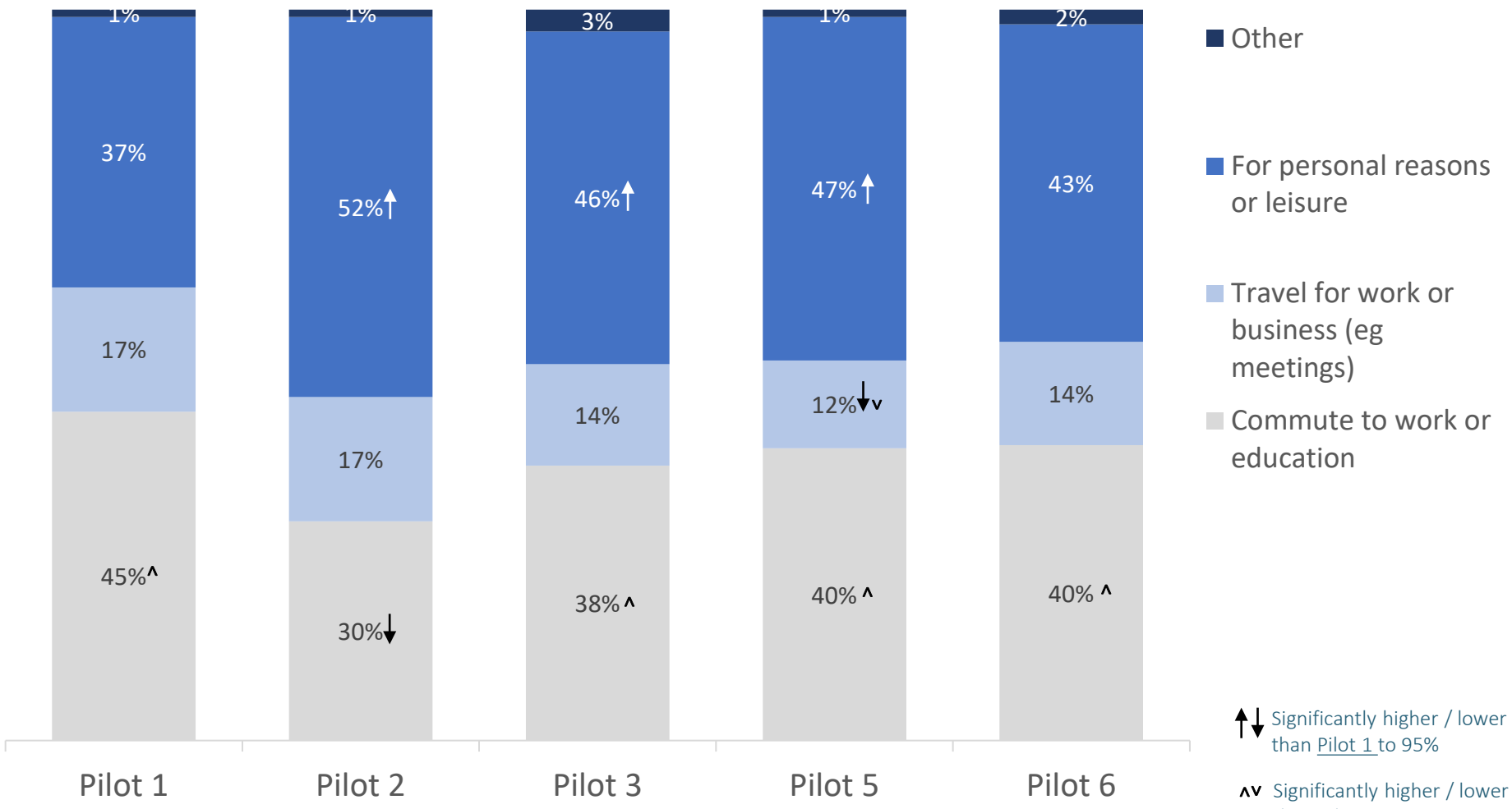




Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base: All who took a journey in the last week. Weekday PM peak (2033), Weekday AM peak (2033), Weekend (576), Weekday any other time (1249)
NB. Spotlight data is Unweighted

Across all pilots we achieved a mix of travel scenarios, though pilot 1 was better at reaching commuters which makes sense given the recruitment time periods*

Purpose of travel for journey taken



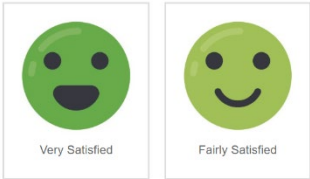
*majority weekday, shifts 8am-2pm and 2pm to 8pm

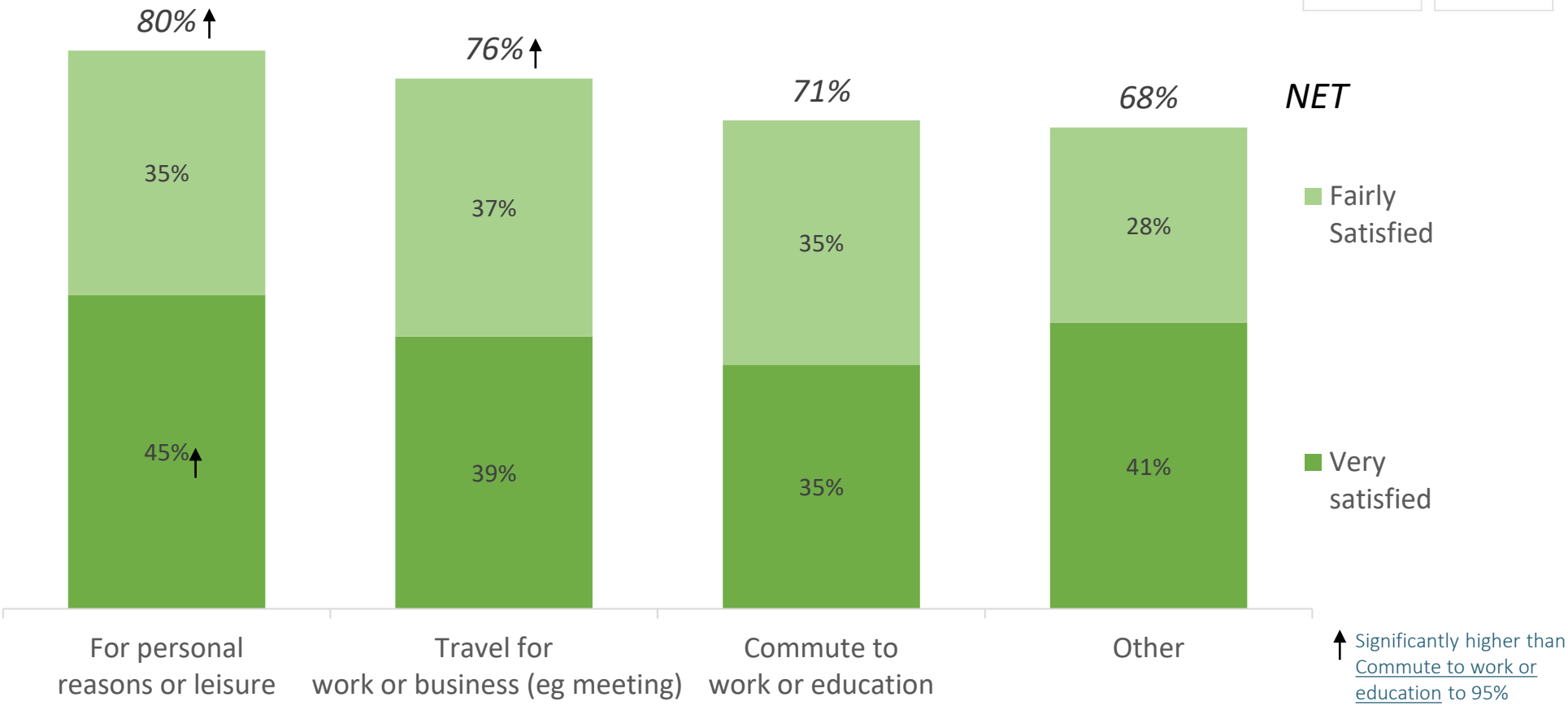
Q4. What was the main purpose of this journey? Base: All who took a journey in the last week. Pilot 1 (204), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984) NB. Spotlight data is Unweighted

Personal travel is more satisfying experience versus commuting for work or education

Journey Satisfaction (top 2 box) by type of traveller: all pilots combined

Similar pattern across all the pilots

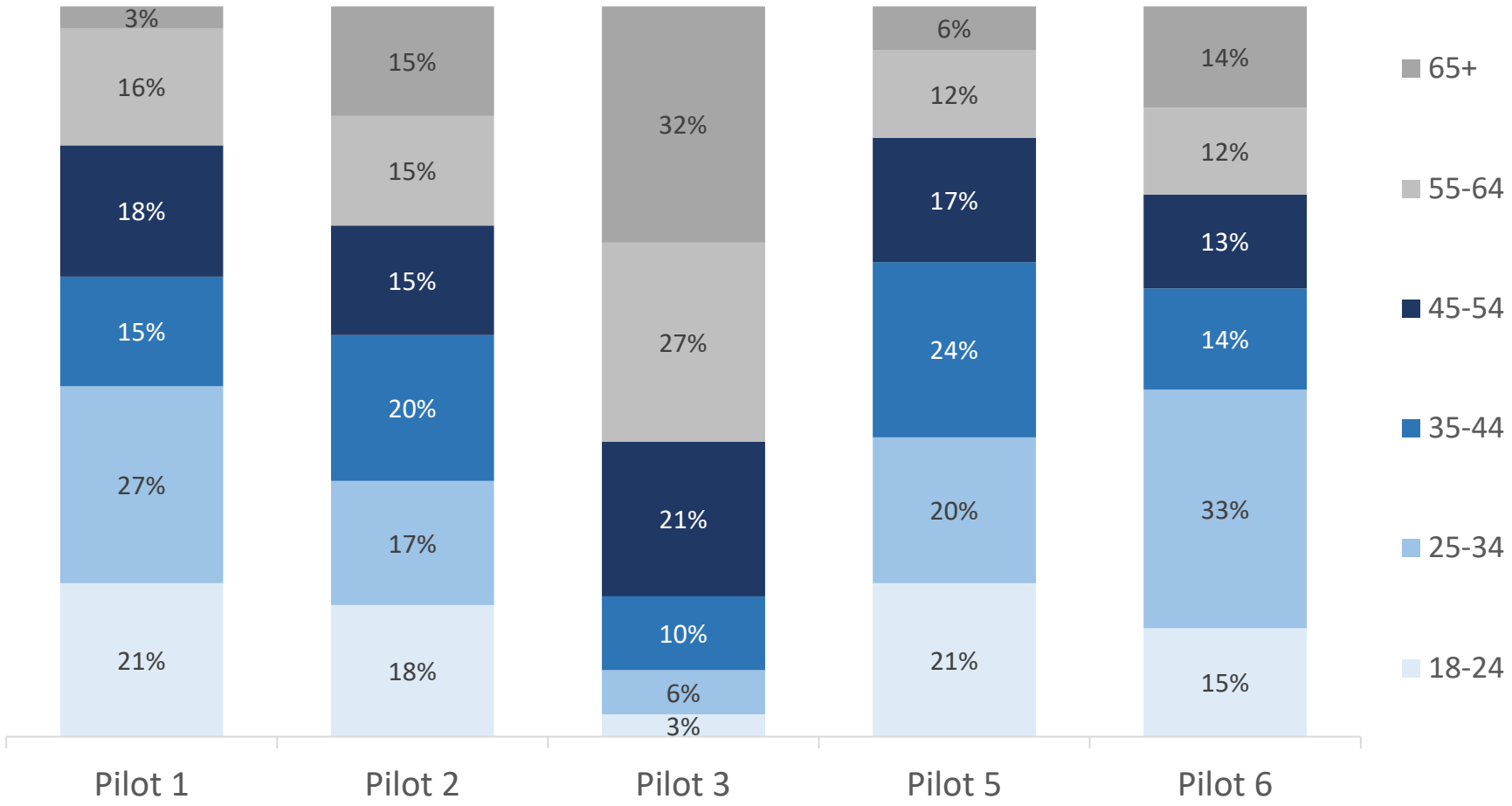




Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base: All who took a journey in the last week. For personal reasons or leisure (2287), Travel for work or business (661), Commute to work or education (1953), Other (76)
NB. Spotlight data is Unweighted

Online surveys give us access to younger travellers, while qualifying* NRPS recontacts are older

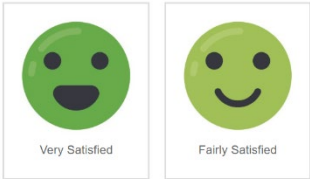
Age of traveller *who travelled in the last week

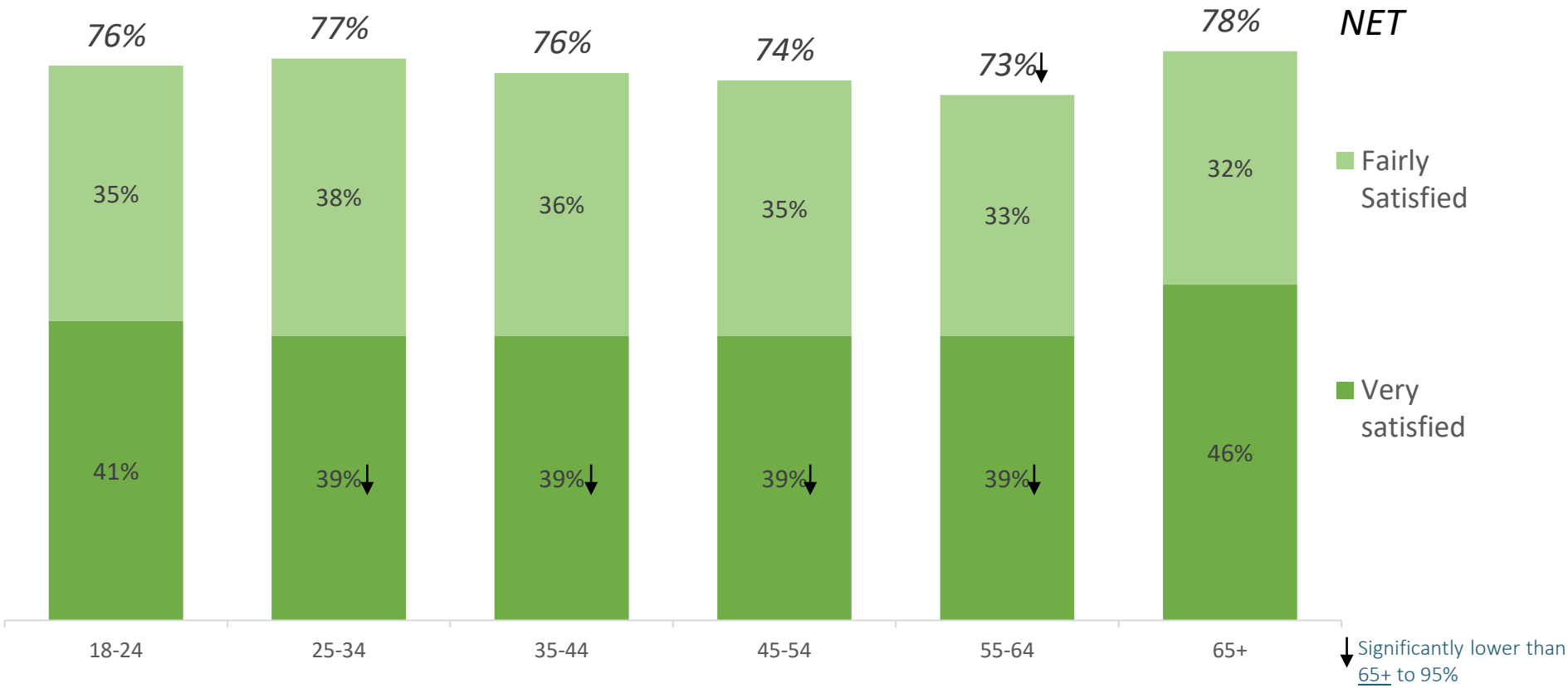


Q4. What was the main purpose of this journey? Base: All who took a journey in the last week. Pilot 1 (204), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984) NB. Spotlight data is Unweighted

Travelers aged 55-64 tend to have lower journey satisfaction

Journey Satisfaction (top 2 box) by age: all pilots combined

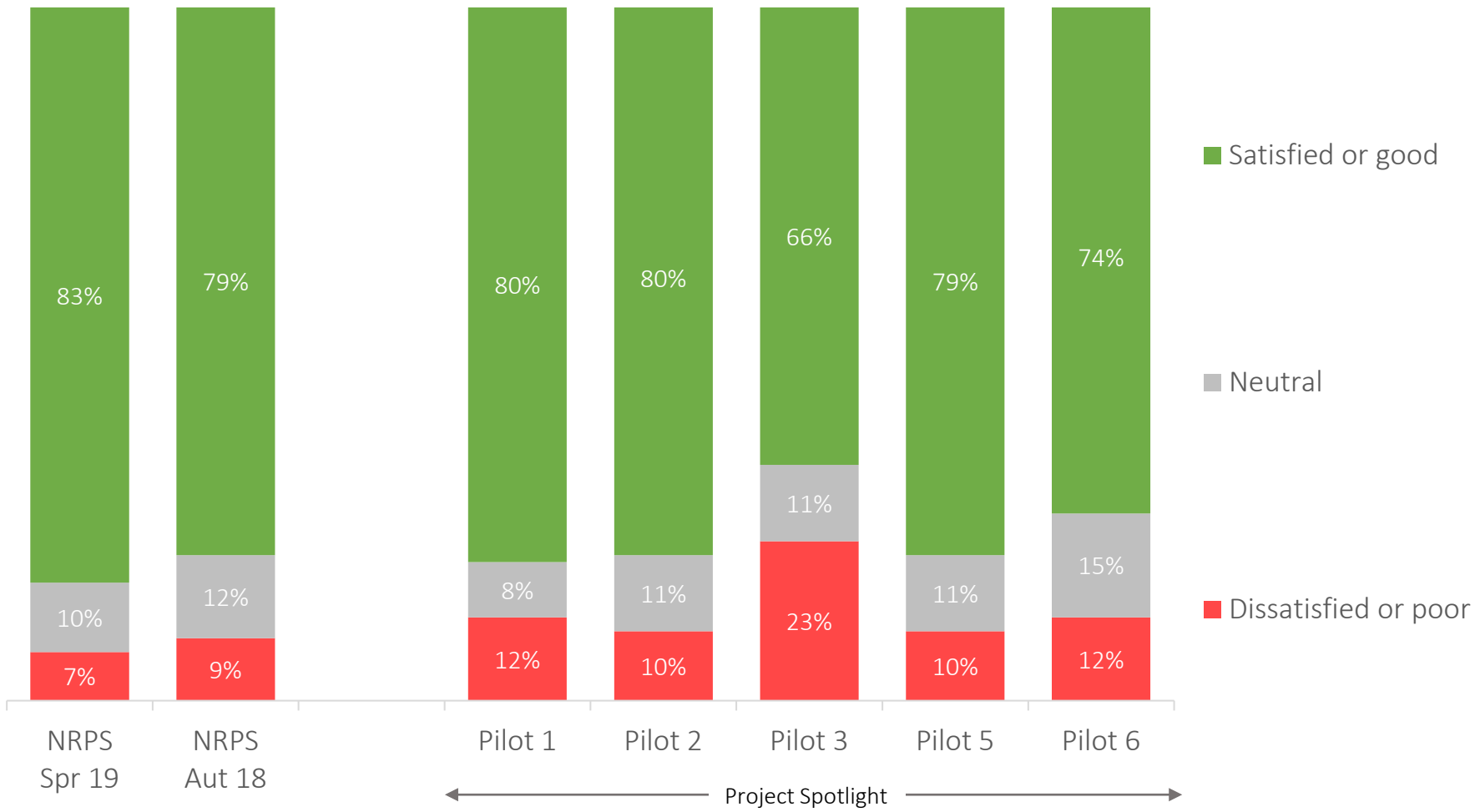




Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base: All who took a journey in the last week. Total (4993), 18-34 (787), 25-34 (981), 35-44 (933), 45-54 (836), 55-64 (775), 65+ (665) NB. Spotlight data is Unweighted

Overall satisfaction for NRPS for Autumn 2018 aren't too dissimilar to pilots 5, 1 and 2

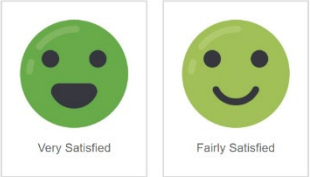
Comparisons to NRPS data

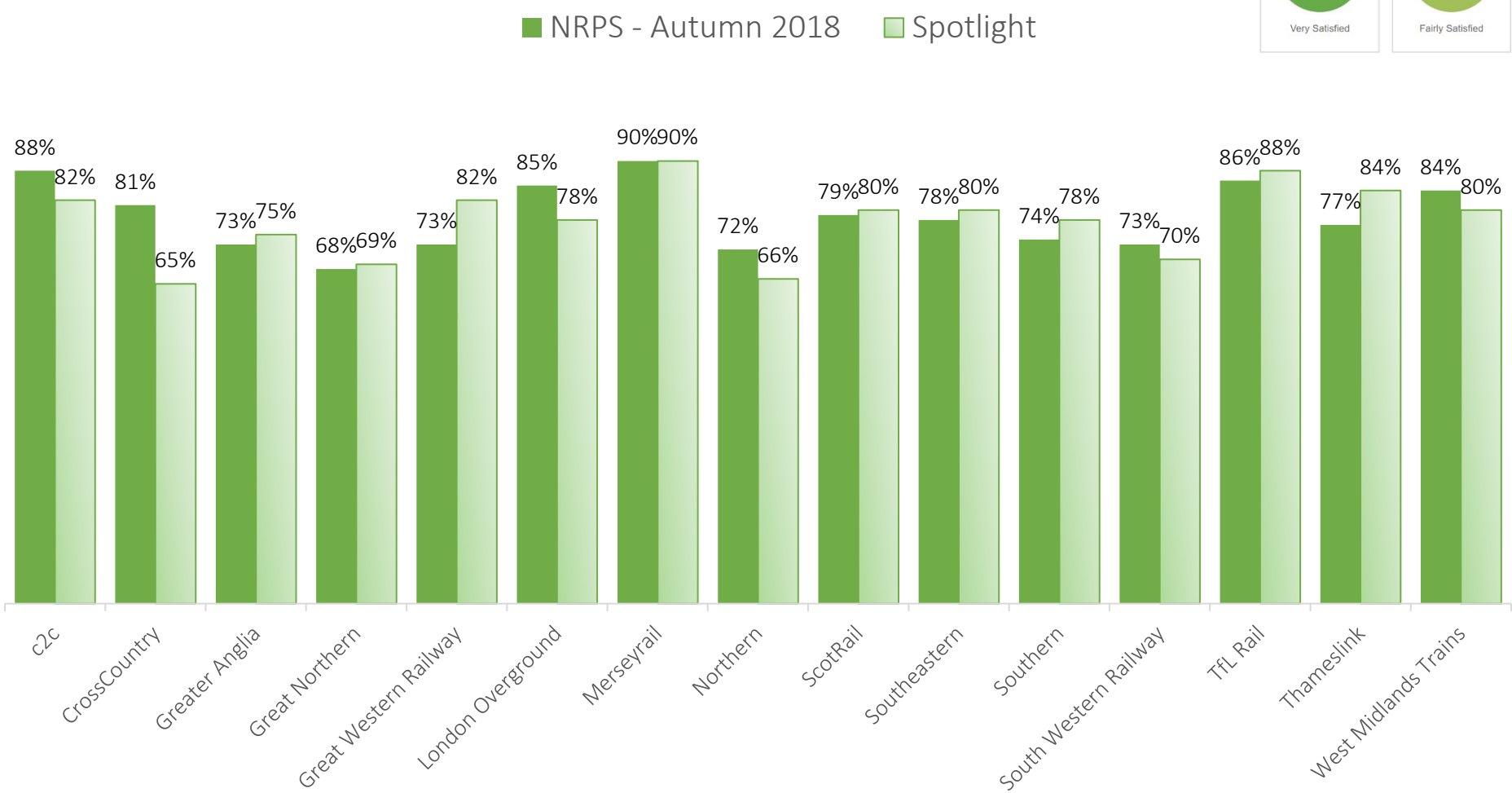


Overall satisfaction with journey. Base. NRPS Spring 2019 (26,494), NRPS Autumn 2018 (26,494). Project Spotlight Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base: All who took a journey in the last week. Pilot 1 (204), Pilot 2 (313), Pilot 3 (977), Pilot 5 (2499), Pilot 6 (984) NB. Spotlight data is Unweighted

At a TOC level overall satisfaction also follows a similar pattern with some differences in actual score

Journey Satisfaction (top 2 box) by TOC: all pilots combined

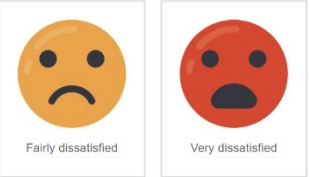


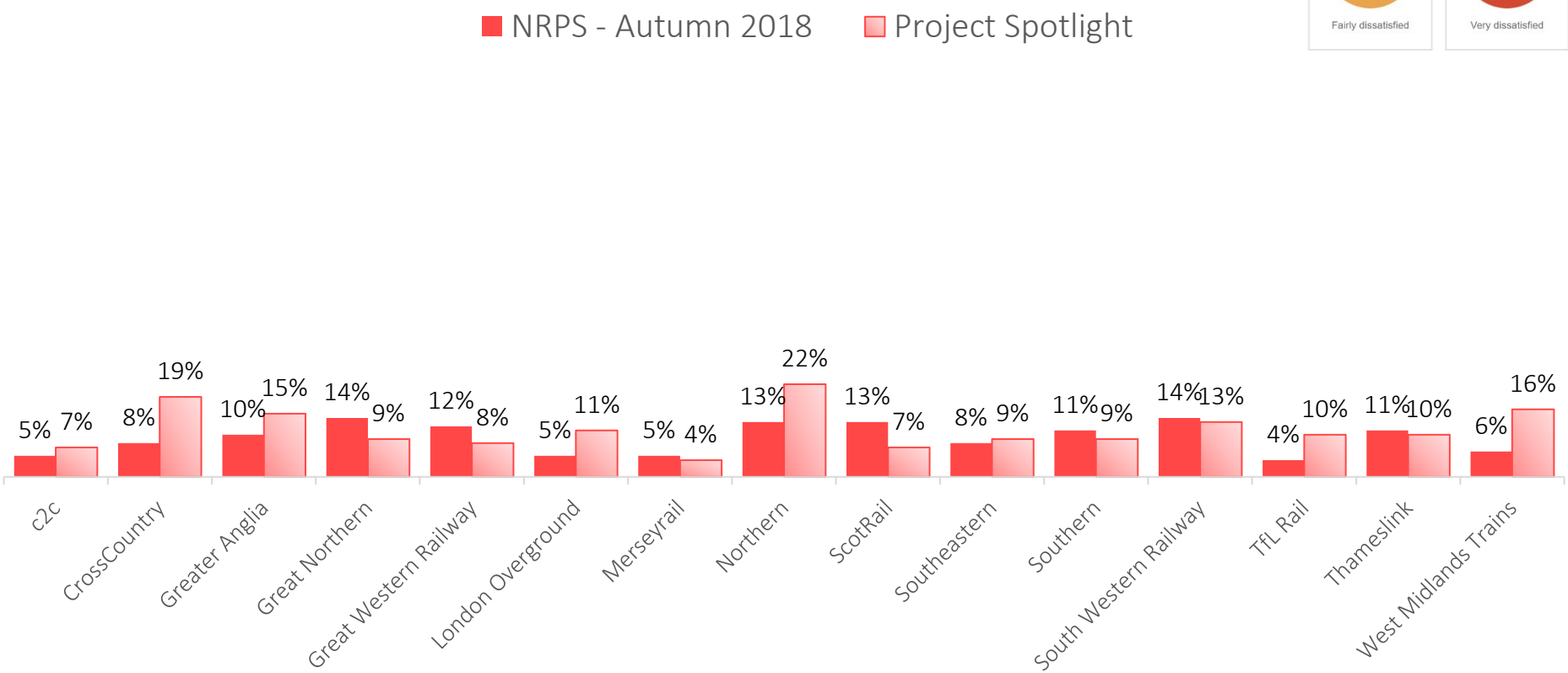


Overall satisfaction with journey. Base. NRPS Autumn 2018 min 700 per TOC. Project Spotlight Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base. All who took a journey in the last week all pilots combined. Min 77 per TOC

Dissatisfaction by TOC is also quite comparable across NRPS and Spotlight for most TOCs, though some see more differentiation

Journey Satisfaction (bottom 2 box) by TOC: all pilots combined





Overall satisfaction with journey. Base. NRPS Autumn 2018 min 700 per TOC. Project Spotlight Q8. Overall how satisfied were you with this journey [TEXT SUB Q4]? Base. All who took a journey in the last week all pilots combined. Min 77 per TOC

Summary: pilot performance overview

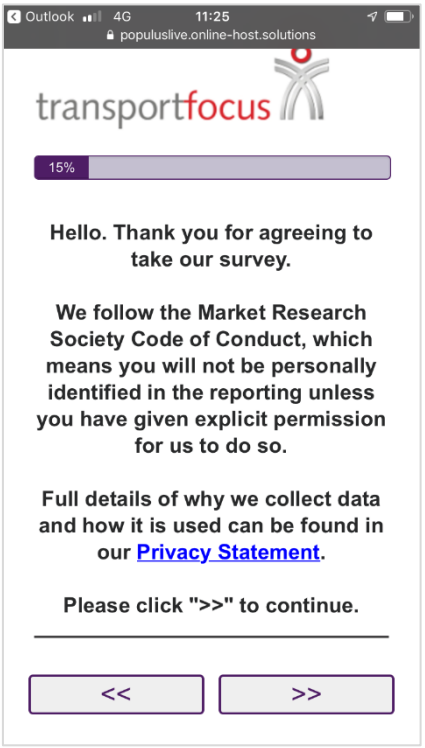
	PILOT 1 Face to face in station	PILOT 2 Face to face on train	PILOT 3 NRPS re-contact	PILOT 5 Omnibus sample	PILOT 6 Ad hoc online panellists
	NO	NO	YES	YES	YES
Strengths	✓ Collected most commuter journeys ✓ Most journeys are daily ✓ Good spread of age ✓ Overall satisfaction is similar to NRPS	✓ Most journeys are daily ✓ Good spread of age ✓ Overall satisfaction is similar to NRPS	✓ High email conversion rate, low week on week attrition ✓ Good mix of daily journeys ✓ Good spread of age and recent travel ✓ Collect more Weekday PM peak journeys ✓ Cheaper sample	✓ High email conversion rate, low week on week attrition ✓ Most journeys are daily ✓ Good spread of age and recent travel ✓ Collect more Weekday PM peak journeys ✓ Overall satisfaction is similar to NRPS	✓ High email conversion rate, low week on week attrition ✓ Most journeys are daily ✓ Good spread of age and recent travel ✓ Collect more Weekday PM peak journeys ✓ Overall satisfaction is similar to NRPS
Weaknesses	× Low diary participation × Very high attrition between recruitment and the diary stage × Costly sample collection	× More leisure trips × Low diary participation × Very high attrition between recruitment and the diary stage × Costly sample collection × Careful route planning so not to skew TOCs	× Considerably older sample × More negative about their journeys rail journeys	× Not an infinite supply of sample	× Not an infinite supply of sample × Smaller sample so need to target specific TOCs

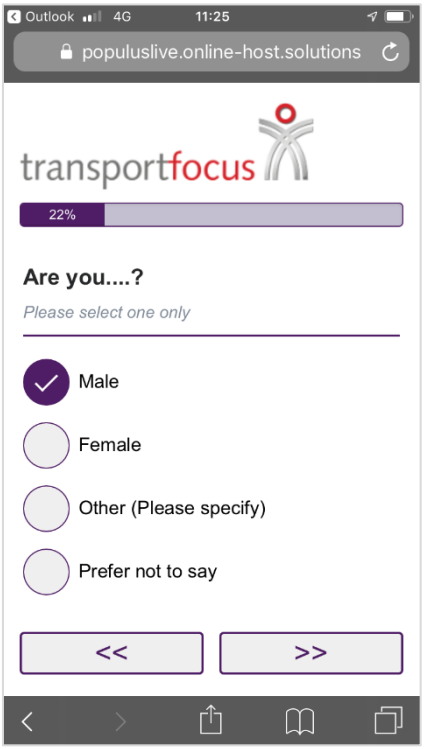
Optimising the current design

Populus

Optimising the questionnaire – recruitment specific questions

12 questions - 2 recruitment / 10 diary questions
Average interview length 2minutes
Mobile/tablet rendered 60%

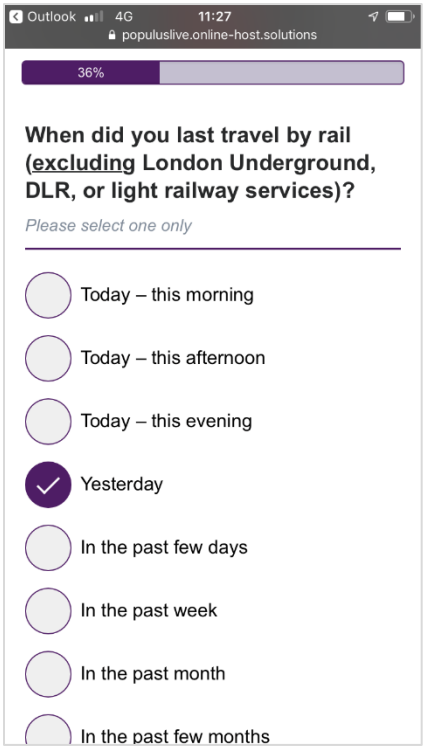




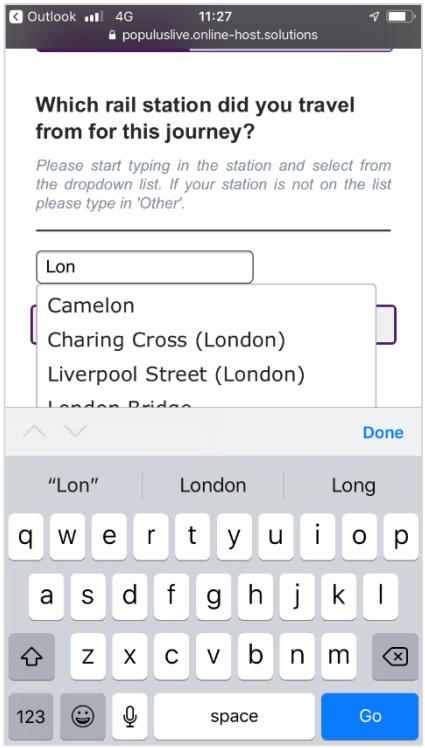
Are there other demographics at the recruitment stage to capture?

- Postcode
- Working status
- Kids in household
- Travel behaviour

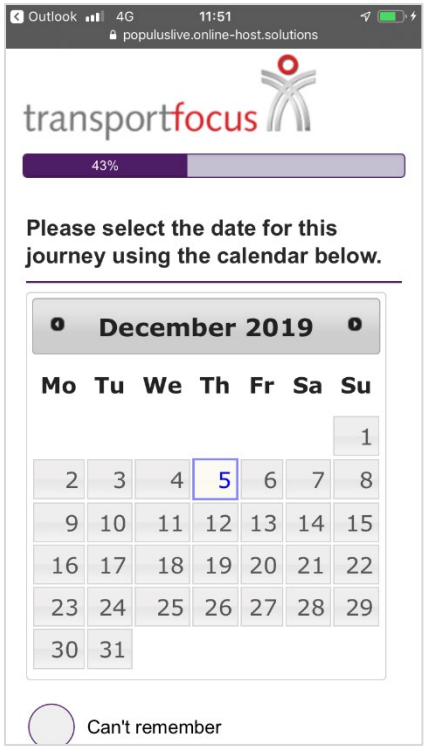
Optimising the questionnaire – journey details



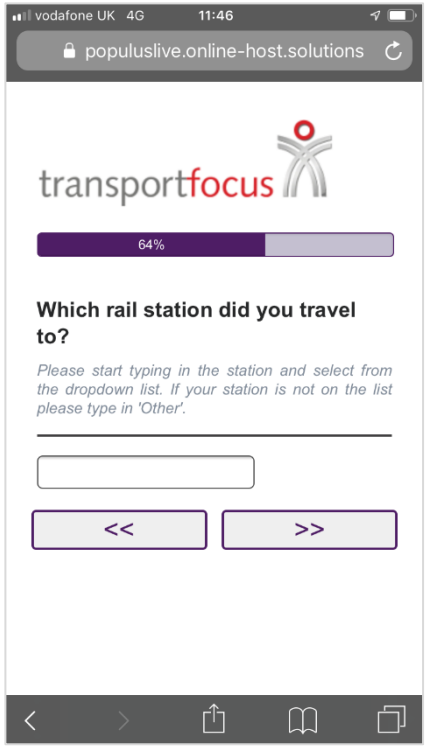
Screen out if travel more than 1 week ago



Departing station - text look up

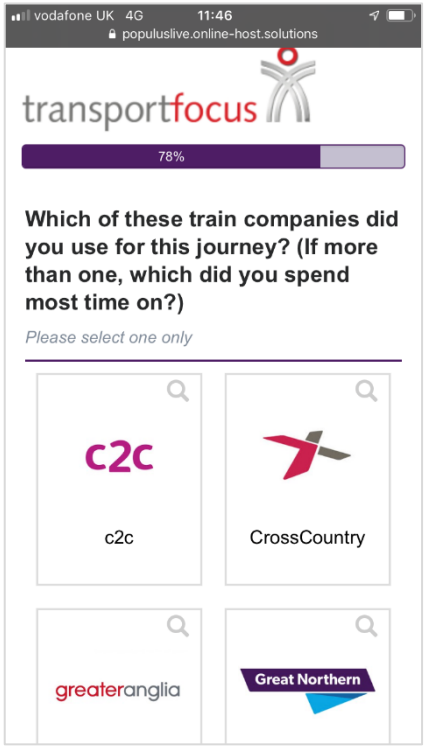


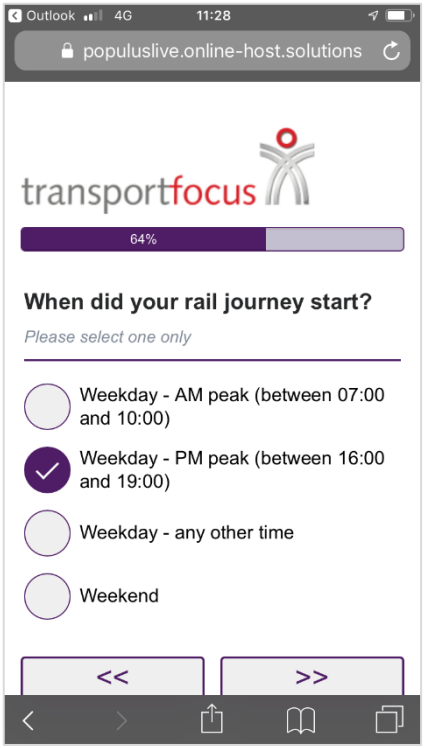
Only allow past 7 days to show on calendar

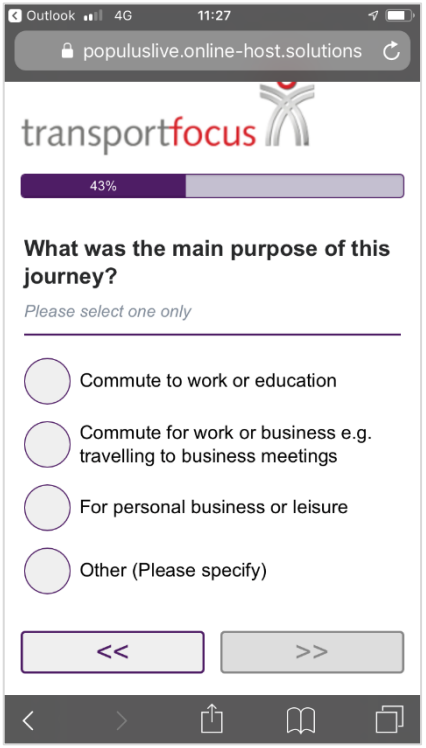


Destination station - text look up

Optimising the questionnaire – journey details





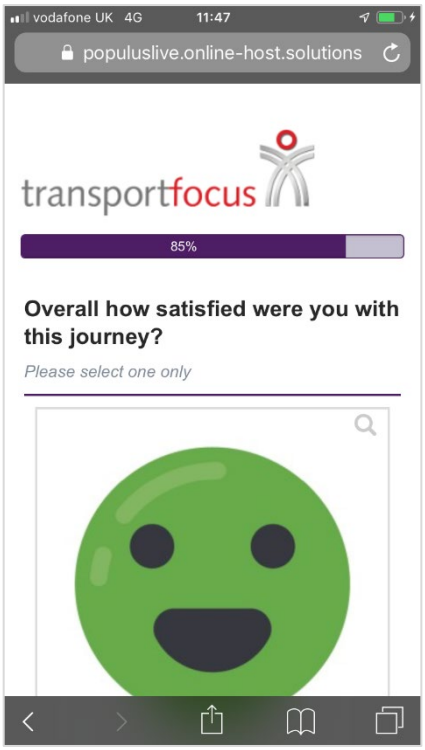


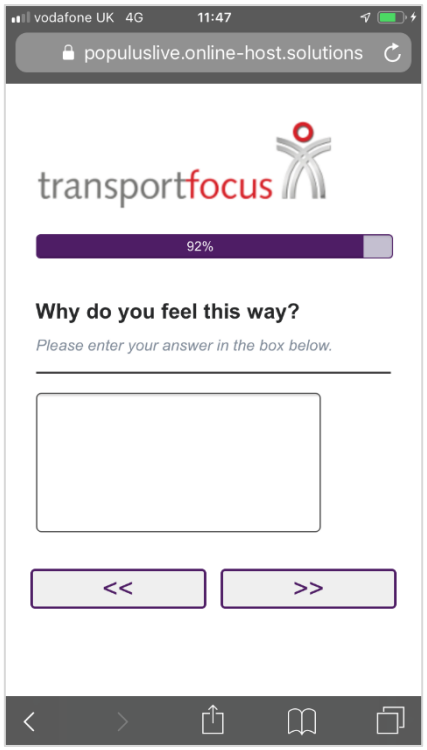
Other journey details to capture:

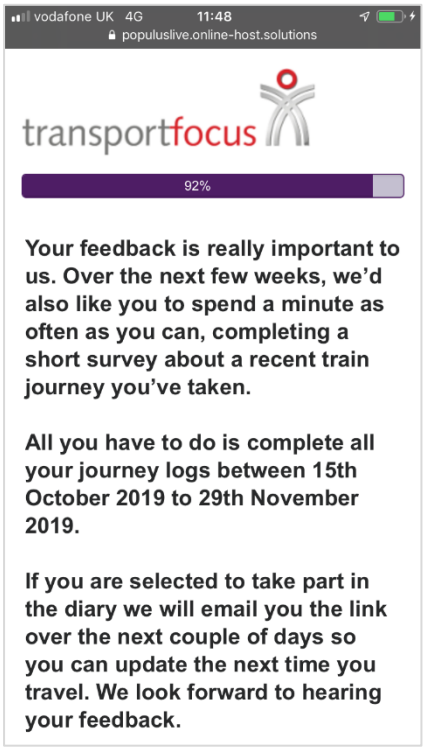
- Length of journey
- Fare type
- Delays
- Tactical questions

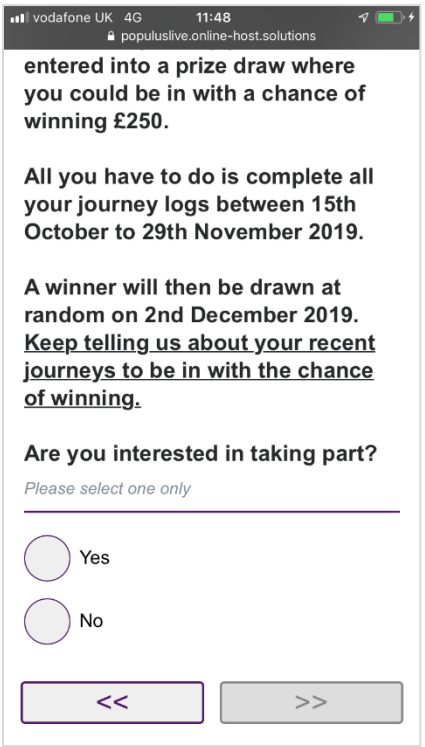
15 TOCs – show full 20 TOCc?

Optimising the questionnaire – journey experience









Scrollable emoji faces – use different images / emotions?

Does the context of the question provide the required response

Final screen at recruitment for non panel sample – prize draw of £250 for participation

Text analytics AI – overview of approach

Two separate AI models developed from the open text question – Context and Sentiment

1. Human coders code the first 500 verbatim comments and define the code frame
   - Context – 15 codes
   - Sentiment – Positive / negative / neutral

   Coded verbatims used to teach the AI how to code up topics and sentiment
2. AI predicts on next 1000 verbatim – this is checked by humans to validate
3. Iterative process to Improve the AI on the total 1500 coded verbatim
4. Full coding begins on remaining raw verbatims
5. Ongoing refinement

Text analytics AI – summary of performance

Two separate AI models developed from the open text question – Context and Sentiment

Context – 15 codes

	Precision	Recall	Accuracy	Samples
Punctuality	98.6%	93.5%	96.5%	682
Staff	81.7%	77.3%	98.0%	75
Cost	93.9%	82.8%	98.6%	93
Cleanliness	98.9%	93.0%	99.0%	186
Capacity/crowding	94.9%	81.7%	93.4%	437
Convenience/ease	86.2%	82.0%	98.7%	61
Speed/efficiency	64.3%	72.9%	84.2%	351
Pleasant/comfortable journey	82.4%	65.9%	88.2%	369
Cancellations	88.6%	100.0%	99.7%	39
Facilities	100.0%	90.5%	99.9%	21
Refreshments	100.0%	78.9%	99.7%	19
Quality of trains/ age of rolling stock	64.2%	73.9%	97.9%	46
Noise levels	100.0%	100.0%	100.0%	27
Other	45.2%	49.0%	82.3%	241
Weighted average	84.7%	79.5%	92.4%	2647

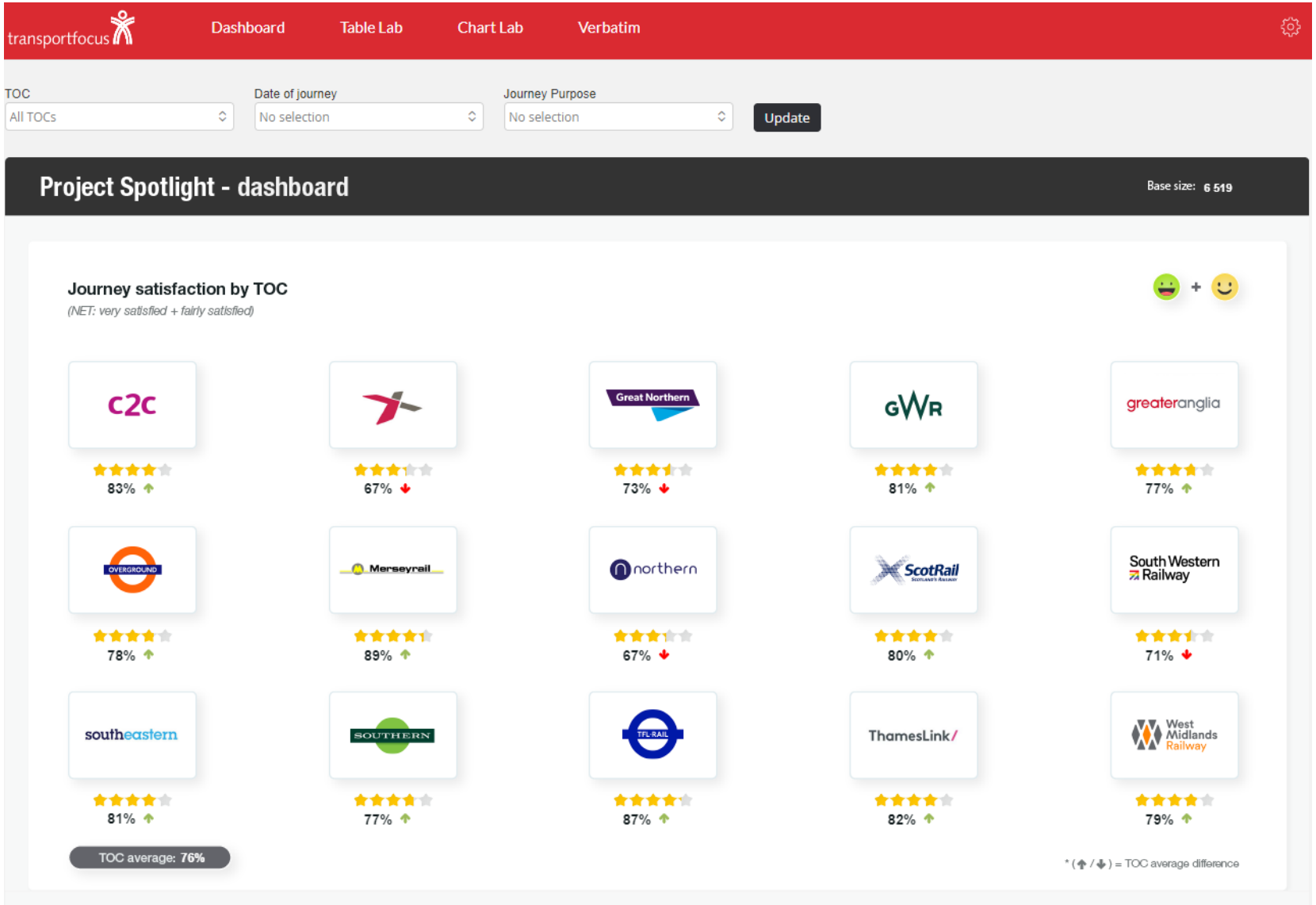
Sentiment

	Precision	Recall	Accuracy	Samples
Positive	82%	94%	83%	917
Neutral	76%	68%	96%	100
Negative	83%	84%	84%	654
Weighted average	82%	88%	84%	1671

- **Precision**: defined as the number of correct predictions out of all predictions made. Essentially this means: **How much of what the model says is correct**.

- **Recall**: defined as the number of actual labels retrieved by the model out of all actual labels. Essentially this means: **How much actual labels are captured.**

Optimising the portal



Optimising the portal

TOC detail - All TOCs

Base size: 6 519

Why satisfied

50%	Punctuality
24%	Capacity/ Crowding
24%	Pleasant/comfortable journey
19%	Speed/ Efficiency
12%	Cleanliness

Pleasent journey	*Pleasant/comfortable journey*
Every train comes on time. Everything is clean	*Punctuality, Cleanliness*
It was mostly comfortable with chargers and sometimes wifi	*Pleasant/comfortable journey, Facilities*

Why dissatisfied

55%	Punctuality
40%	Capacity/ Crowding
17%	Cancellation
14%	Other
10%	Convenience/ Ease

Delayed by 3 hours	*Punctuality*
Late everyday no sezys	*Punctuality*
It was delayed	*Punctuality*
Tain do run lot train not birminfham to viemingham	*Convenience/ Ease*

Journey satisfaction

38%	38%	12%	7%	5%

Journey satisfaction by daypart

Weekday AM peak	76%
Weekday PM peak	70%
Weekday non peak	80%
Weekend	78%

Summary: optimising the current design

- The **micro survey** worked well and was kept short at 2 minutes – can this be further shortened? Or have tactical questions that are asked every nth person
- Further testing on **emoji's + context question** needed – what is the best way to ask/represent journey experience in survey?
- Using **Artificial Intelligence** to code high volumes of raw open responses is feasible to a high level of accuracy – the AI model will need ongoing training to improve accuracy
- The purpose and scope of the **portal** needs further thought – would tracking data be useful and how much data should be shared?
    - How frequently should the data be uploaded – daily?