Train Operating Company Audits

—

Research agency report

Main Findings

19th December 2023

Fieldwork: 16th February to 22nd March 2022

Introduction

In Autumn 2021 Transport Focus and the Department for Transport (DfT) were keen to understand the current state of the rail network in terms of the condition and presentation of stations and trains managed by the 14 train companies (under DfT's jurisdiction) across the rail network in England. To do this Transport Focus undertook 'Network Standards' audits by means of mystery shopping in conjunction with an agency to give a baseline for a set of core factors that covered three key areas: stations, trains, and availability of staff.

Mystery shoppers from the fieldwork agency *LIVE*, led by *7th Sense* visited stations and travelled on board trains to assess the presentation and condition of facilities, and availability of staff. Most evaluations were recorded with scales between 0-10 for each aspect. They worked inconspicuously, with some rotation of different individuals for each train company in order that they did not become unusually familiar to staff.

On average 100 audits per Train Operating Company (TOC) were conducted (split equally between station and train), but flexibility was needed to tailor this by TOC - some having very few stations (or none in one case), whilst one TOC manages several hundred stations. Sample sizes were under 30 for some TOCs, but as the table on page 8 shows the % of station audits carried out by TOC was above 60% for five of the six TOCs with such low sample sizes (and in three cases was 96% to 100%). Note: For this research we covered stations served by each train company.

Introduction (cont'd)

In terms of train audits it is the case that passenger numbers and number of trains run by each train company vary significantly – hence the variation in the number of audits on trains by each TOC.

The audits were conducted between 16 February and 22 March 2022. Note: Initially the audits were planned to be done in December 2021 to January 2022 but the audit was delayed for a few months, mainly due to the pandemic.

The following report presents the results from the audits – in many cases by each train company. Note: Where sample sizes are small this is noted on each page. It should be noted that results are based on just one visit to each station, and results may vary over time.

In this condensed report it is not possible to present results for all factors, but it presents results for many of the key factors (including through comments on each page).

The overall principle for setting the sampling was to audit a number of stations or trains operated by a given TOC, based on the proportion of passenger journeys delivered and stations that they operate.

So, in principle, TOCs operating a large number of stations received a large number of assessments. A TOC delivering fewer customer journeys than others would receive fewer train assessments. Since 700 assessments were called for on trains and 700 on stations, the proportion of journeys delivered/stations operated was the basis for the percentage of 700 assessments that would be carried out by TOC. However, many factors influenced the final allocation:

Stations

- The number of stations of each size that the TOC operated

- The geographic spread of those stations

- The number of National Rail-managed stations that they operate from

- The preference for a minimum reportable value of n=50

Trains

- The need to cover different service times (peak/off peak/weekend)

- The principle that the proportion of assessments should be based on number of passenger journeys rather than number of services operated

- The preference for a minimum reportable value of n=50

General and specific considerations

General considerations for sample planning

Since this project was about the general customer experience of the trains and stations, sampling was based on the number of passenger journeys per TOC; the greater number of passengers a TOC has, the greater the impact of the condition and service received by those passengers will be.

- Volume of assessments was not based on passenger miles, since the condition of the customer-facing areas of a train are somewhat less impacted by the miles travelled as by the volume of people travelling on it, and the general experience does not alter significantly as the distance increases in a given journey.

- Since certain TOCs carry far more passengers than others, if proportions of passenger journeys were applied strictly, the 'smaller' TOCs would likely be given such low assessment numbers as to deliver a meaningless sample.

- Thus, the overall principle applied was that for 'smaller' TOCs all stations and trains should be assessed up to a mid point of n=50 assessments – if they operate that many stations – and after that, assessments were distributed according to volume of journeys or stations operated.

Specific considerations for individual TOCs

- Some TOCs with relatively high passenger volumes actually manage very few stations – of course, CrossCountry manage none at all. For much DfT work, a sample size under 50 is not reportable. However, if that criterion was applied and all TOCs were to receive at least 50 station audits, LNER (11 stations) and Avanti West Coast (16) would have received many multiple assessments at the same stations, which would not have revealed any new information and have used up assessments better spent on TOCs with a larger estate.

- With 700 station audits available, if every TOC was given 50 and all the 50 'spare' from CrossCountry went to Northern, only 20% of Northern's 477 stations would be assessed, compared with nine assessments for LNER's 11 stations – a much higher 81%.

- Since, statistically speaking, assessing every station operated by a given TOC is covering the entirety of their given universe, a single audit to each station was deemed by the clients (DfT and Transport Focus) to be acceptably representative.

Sample plan calculations and considerations

The TOCs that have fewer than 50 stations ('the small TOCs') Avanti West Coast, c2c, Chiltern Railways, LNER and TPE between them manage 106 stations. There are 1,629 stations managed by TOCs who have more than 50 stations but also some TOCs have staff at National Rail stations and these also needed to be assessed. The 700 station assessments were divided between the small TOCs at 106, and remainder of the 1,629 'left over' divided by the proportion those large TOCs manage.

However, this would have left EMR having around the same number of audits as Chiltern Railways, despite managing three times the number of stations, and having a sample size lower than 50. A redistribution was made to rebalance the number of assessments so that EMR were given a minimum of 50 assessments, and the other TOCs reduced by a relative amount.

For on-train audits, similar calculations were made to divide up the number of assessments per TOC:

- The sample plan was based on the fact that fieldwork for trains and station audits could be conjoined – that is, an assessor could complete an audit then leave the train and audit the station at which they alight. This was to ensure that, as far as possible, good value could be achieved for the taxpayer by maximising the use of the assessor's time. Therefore, as a starting point, the small TOCs were given the same number of on-train assessments as at-station assessments.

- A complicating factor was that CrossCountry services were to be assessed but, of course, no station audits were possible. Conversely the 20 Network Rail managed stations were to be assessed but Network Rail do not operate any trains.

- We needed to retain the principle that, as far as possible, n=50 would be an ideal minimum, but that some TOCs carry too few passengers to warrant that level of audits. Thus, some TOCs have quite a wide difference in the number of station audits to on-train audits.

- The proportion of station-to-train audits varies. Some TOCs also operate a large number of stations but deliver relatively few passenger journeys – particularly Northern who, after the small TOC stations are removed, operate 34% of remaining stations, compared with only 11% of remaining passenger journeys (once small TOCs are removed). Conversely Thameslink operate 17% of 'remaining' stations but deliver 30% of 'remaining' journeys.

- Balance of peak/off peak/weekend travel distribution for that TOC and station sizes (A-F) operated by that TOC were used to distribute the number of assessments for each journey time and station size. For station size a 'reasonable minimum' of n=10 was aimed for in order to deliver a good selection of each station size, except where (e.g. on A stations) there simply weren't enough stations of that type to assess.

Sample Definition

Station and Train Audit Sample Plan

TOC	Completed Station Audits	Completed Train Audits
All TOCs	680	700
Avanti West Coast (Avanti WC)	16	17
c2c	25	26
Chiltern Railways	35	36
CrossCountry	0	17
East Midlands Railway (EMR)	50	53
Great Northern	9	28
Greater Anglia	52	49
Great Western Railway (GWR)	73	48
London North Eastern Railway (LNER)	7	16
Northern	149	49
Southeastern	60	86
Southern	55	53
South Western Railway (SWR)	61	102
Thameslink	20	55
TransPennine Express (TPE)	19	22
West Midlands Trains (WMT)	49	43

Note: Fieldwork took place from 16th February to 22nd March 2022. 20 station audits were also conducted at Network Rail stations.

Sample Definition

% of Stations Audited

TOC	Completed Station Audits	Total Stations Operated	% of Stations Audited
All TOCs	**680**	**1,743**	**39%**
Avanti West Coast (Avanti WC)	16	16	100%
c2c	25	26	96%
Chiltern Railways	35	35	100%
CrossCountry	0	0	N/A
East Midlands Railway (EMR)	50	90	56%
Great Northern	9	54	17%
Greater Anglia	52	133	39%
Great Western Railway (GWR)	73	194	38%
London North Eastern Railway (LNER)	7	11	64%
Northern	149	477	31%
Southeastern	60	164	37%
Southern	55	156	35%
South Western Railway (SWR)	61	187	33%
Thameslink	20	32	63%
TransPennine Express (TPE)	19	19	100%
West Midlands Trains (WMT)	49	149	33%

Note: 20 station audits were also conducted at Network Rail stations.

—

Station Results

Overall Station Rating Scores

All Ratings Combined: By TOC

Mean scores for stations are relatively high. Overall, the TOCs are delivering a service that is rated at least 'good' (7-8) and (if above 9.0) 'excellent'. Mean scores below average are often associated with consistent poor scores across the assessment. Chiltern Railways, Great Northern, Southeastern, Southern, South Western Railway (SWR), Thameslink and West Midlands Trains (WMT) consistently post lower scores across many factors. TOCs performing well here are also very often the highest performers in most ratings.

All TOCs	Avanti WC	c2c	Chiltern Rail	EMR	Great Northern	Greater Anglia	GWR	LNER	Northern	Southeastern	Southern	SWR	Thameslink	TPE	WMT
8.32	8.45	8.33	7.97	8.91	7.50	8.59	8.73	9.05	8.59	7.93	8.05	7.92	7.82	8.62	7.51

Mean score (Rating 1-10)
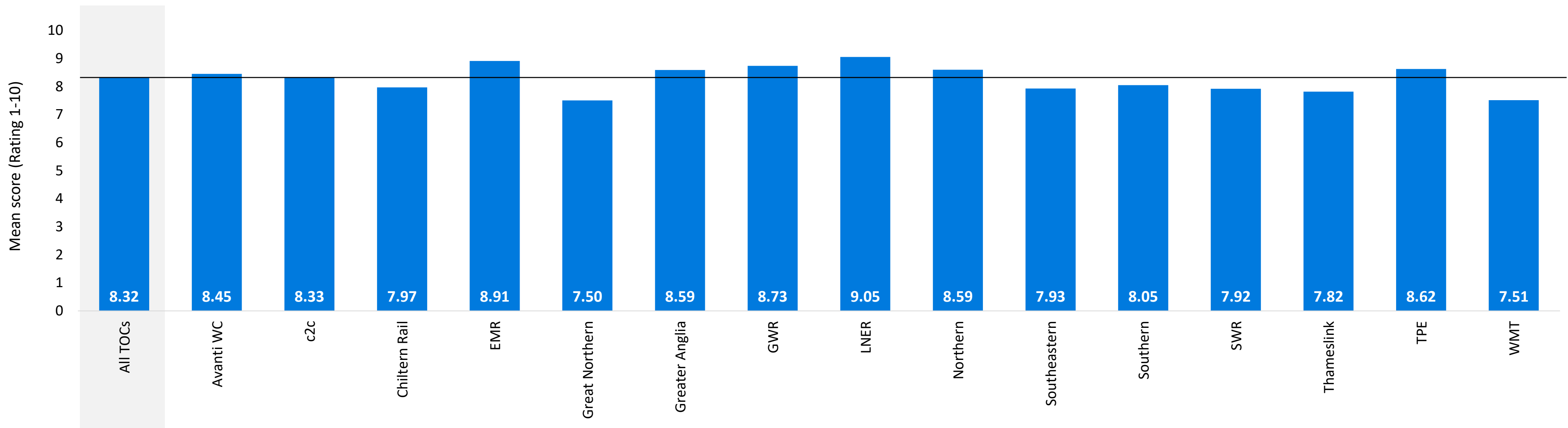
Q2a/3a/4a/5a/7a/9b/10a/11a/12a/14a/17b/24c/25a/27c/28b/29a/30a/34a/37b/38a/40a/41a/42a/45a/46c/47a/49c/51b/52a/54a/56a/57a

Total N = 680 (All TOCs) . Scores calculated are means of mean scores for each of the above rating questions.

Overall Station Rating Scores

Station Entrance: Cleanliness

Nearly 90% of station entrances are rated as good to excellent for cleanliness. In particular, c2c, East Midlands Railway (EMR), Greater Anglia, Great Western Railway and London North Eastern Railway (LNER) had over half their entrances rated 9-10. However, Great Northern, Southeastern, Southern, and SWR all posted fewer than 10% 'excellent' scores. Overall, ticket hall cleanliness had similar ratings to station entrance cleanliness.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Note
All TOCs	12%	51%	37%	
Avanti WC	6%	56%	38%	**16
c2c	4%	40%	56%	**25
Chiltern Rail	17%	60%	23%	
EMR	10%	16%	74%	
Great Northern	44%	56%		**9
Greater Anglia		38%	62%	
GWR	4%	41%	55%	
LNER		43%	57%	**7
Northern	6%	45%	49%	
Southeastern	8%	87%	5%	
Southern	25%	71%	4%	
SWR	25%	70%	5%	
Thameslink	20%	60%	20%	**20
TPE	26%	42%	32%	**19
WMT	22%	41%	37%	
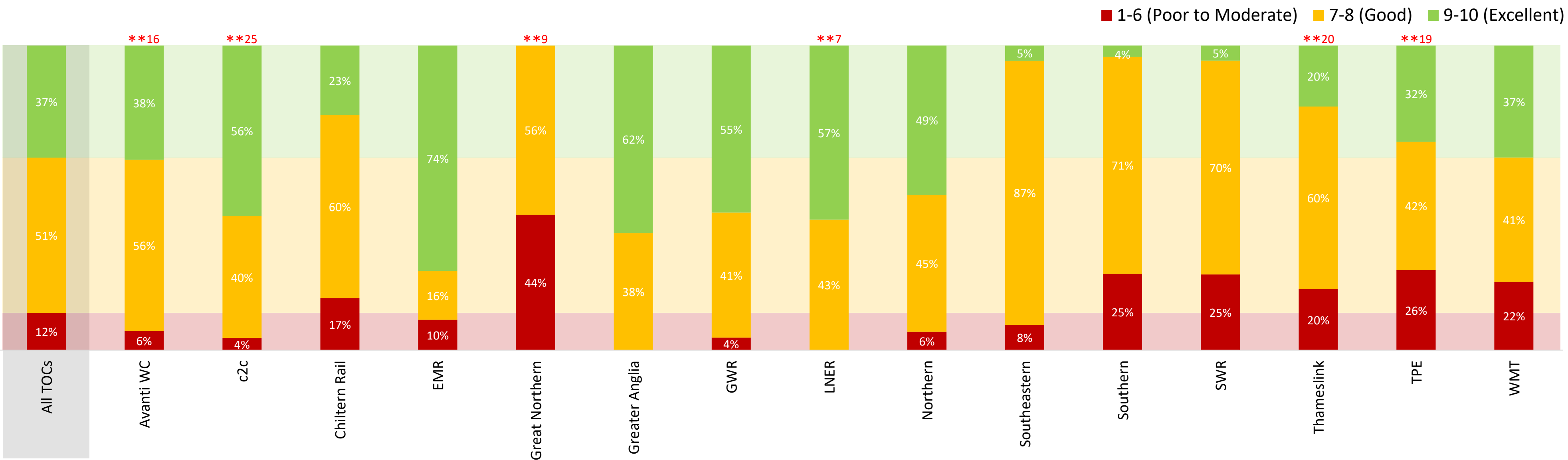
Q9b: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the cleanliness of the station entrance?

Total N = 680 (All TOCs) **indicates small sample size (under 30 stations)

Overall Station Rating Scores

Platform: Cleanliness

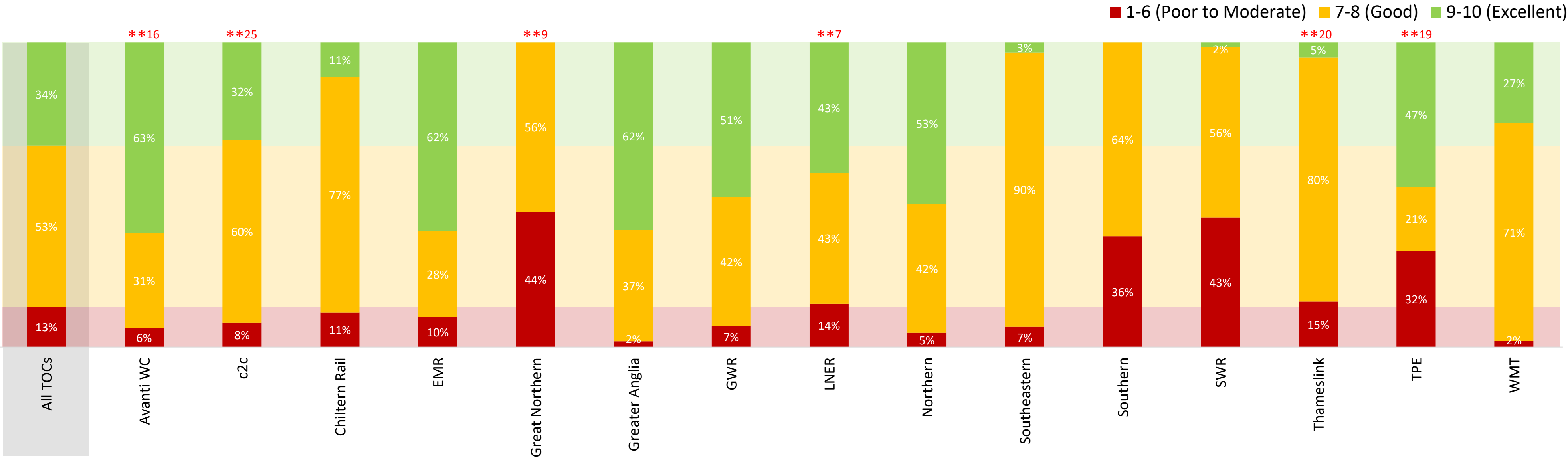Platform cleanliness was generally well or quite well maintained. Avanti West Coast, EMR and Greater Anglia all had more than 60% excellent scores. Great Northern, Southern, TransPennine Express (TPE) and SWR are rated poor to moderate around three times more than average.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Note
All TOCs	13%	53%	34%	
Avanti WC	6%	31%	63%	**16
c2c	8%	60%	32%	**25
Chiltern Rail	11%	77%	11%	
EMR	10%	28%	62%	
Great Northern	44%	56%		**9
Greater Anglia	2%	37%	62%	
GWR	7%	42%	51%	
LNER	14%	43%	43%	**7
Northern	5%	42%	53%	
Southeastern	7%	90%	3%	
Southern	36%	64%		
SWR	43%	56%	2%	
Thameslink	15%	80%	5%	**20
TPE	32%	21%	47%	**19
WMT	2%	71%	27%	

Q40a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the cleanliness of this platform?

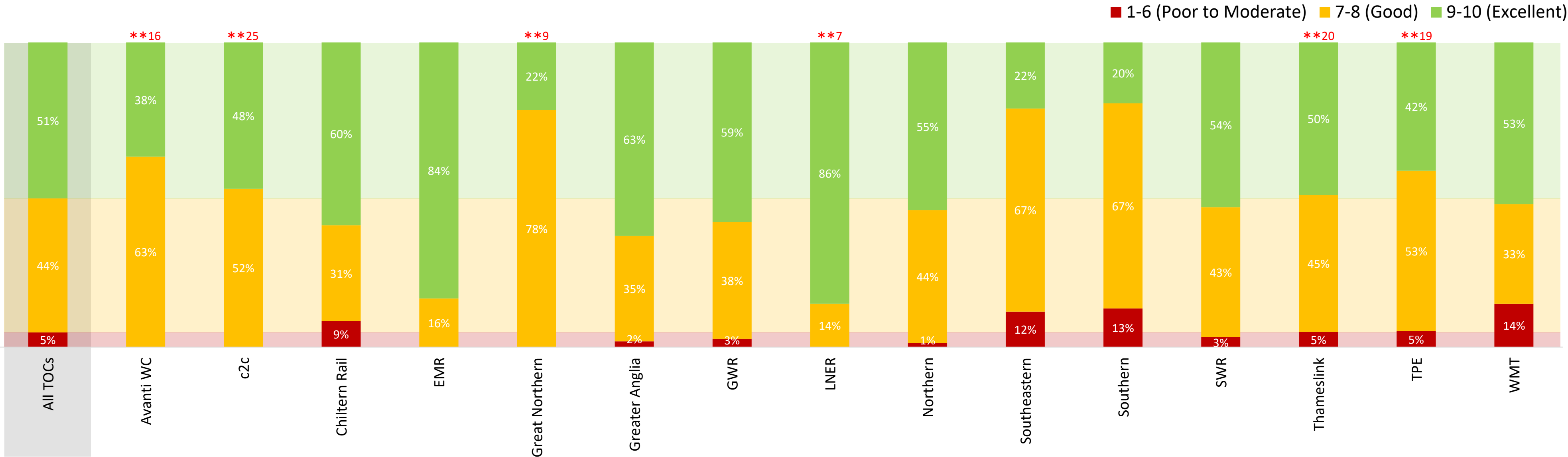Total N = 680 (All TOCs) **indicates small sample size (under 30 stations)

Overall Station Rating Scores
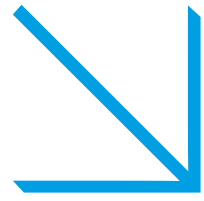
Station Entrance: Litter

Generally, stations across all TOCs performed fairly well in terms of entrance litter management, with over half of stations rated as excellent. EMR, LNER and Greater Anglia perform particularly well. Great Northern, Southeastern and Southern had a low percentage of excellent ratings. WMT had the highest percentage of poor to moderate ratings, though also had a high percent rated as excellent.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	9-10 (Excellent)	7-8 (Good)	1-6 (Poor to Moderate)
All TOCs	51%	44%	5%
Avanti WC **16	38%	63%	
c2c **25	48%	52%	
Chiltern Rail	60%	31%	9%
EMR	84%	16%	
Great Northern **9	22%	78%	2%
Greater Anglia	63%	35%	
GWR	59%	38%	3%
LNER **7	86%	14%	
Northern	55%	44%	1%
Southeastern	22%	67%	12%
Southern	20%	67%	13%
SWR	54%	43%	3%
Thameslink **20	50%	45%	5%
TPE **19	42%	53%	5%
WMT	53%	33%	14%

Q10a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the level of litter in the station entrance?

Total N = 680 (All TOCs) **indicates small sample size (under 30 stations)

Verbatims and Images

Station Entrance

Cleanliness and litter

'Excellent' images (10/10 scores)

Tilbury Town
c2c

Fiskerton
East Midlands Railway

'Poor' to 'Moderate' images and verbatims (1-6 scores)

"General dirt and damp patches across the station entrance building which give a cold and isolating feel."
Bishopstone (Sussex), Southern

"It looks a bit shabby. There are cigarette butts scattered, only weeds in the flower boxes and looks like it will need a lick of paint soon."
Northfleet, Southeastern

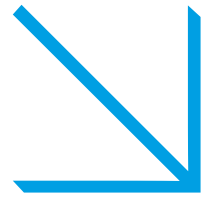*"Tiny wrappers all over the place as well as chewing gum all over the floor."*
Newhaven Town, Southern

Q9c/9d/10b/10c: If your score is 5 or lower, please explain why you have given this rating. Please upload one image to illustrate your rating (scores of 1-5 or 10).

Ticket Buying Facilities and Ticket Office

Ticket buying facilities available

Ticket office	Ticket vending machine(s)	Permit to ride machine(s)	Purchase on board/at destination sign	None
59%	82%	4%	31%	5%
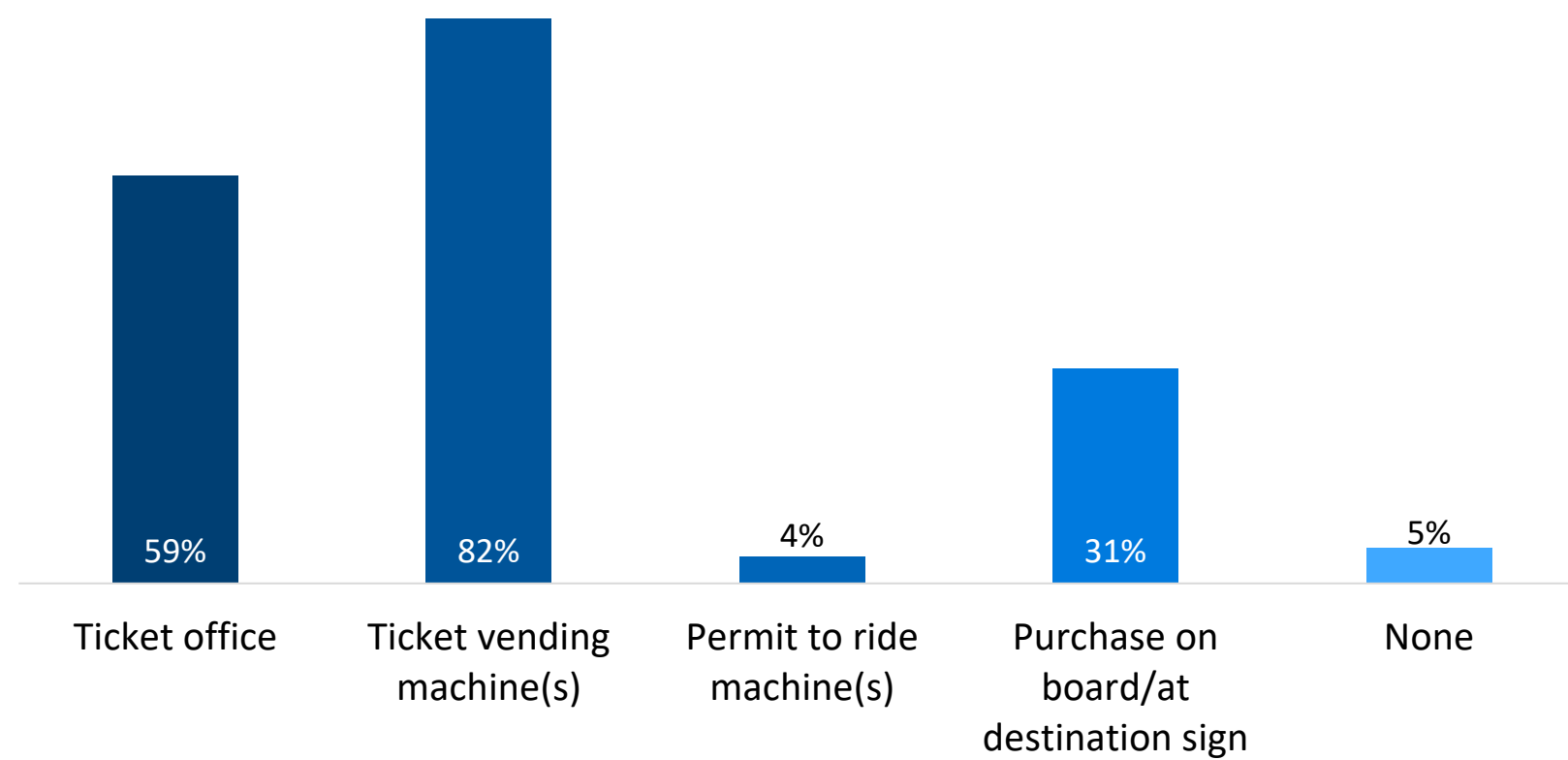
Ticket office open/closed in line with advertised hours?

■ Yes, as advertised ■ No, not as advertised

6%

94%
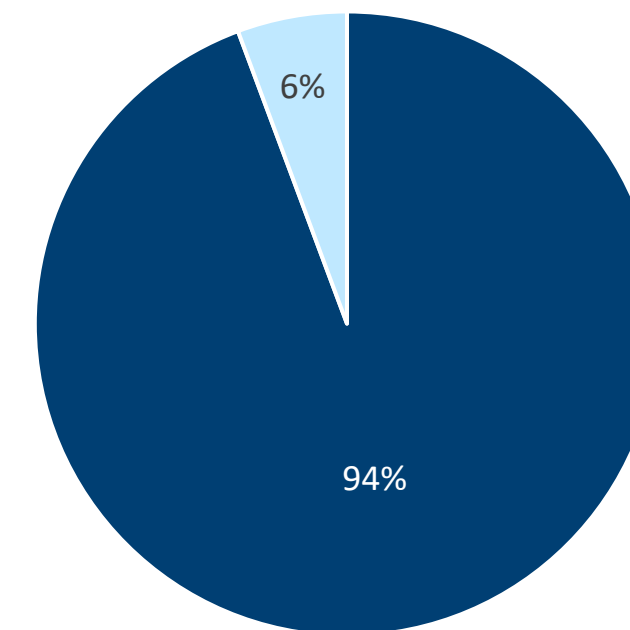
Q15a: Which of the following ticket buying facilities are available at this station? 15b: Is the ticket office open in line with its advertised opening hours?

N = 680 (15a), N = 404 (15b)

Ticket Vending Machines: Functionality

Other consists of:
- Advises speaking to staff member or calling listed telephone number **(50%)**
- Button unresponsive/faulty/just shows serial numbers **(36%)**
- Says 'no service operators available to assist at this time' **(11%)**
- Other **(3%)**

Ticket vending machines working?

■ Yes ■ No

1%

99%

Help button/option available?

■ Yes ■ No

53%

47%

Ways of obtaining assistance

Button to agent, answered	Button to agent, not answered	Button to onscreen help	Other
26%	7%	32%	38%
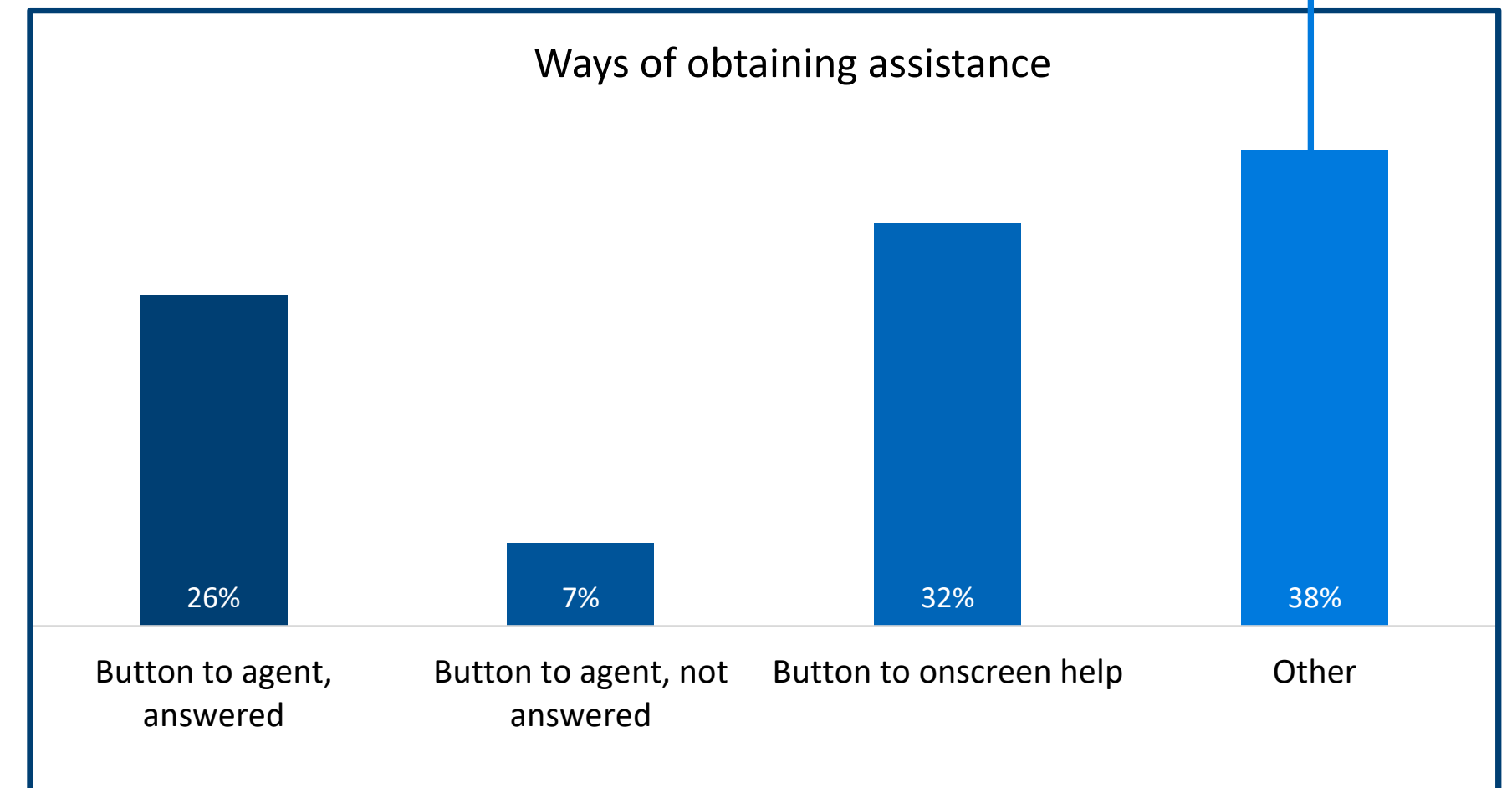
Q15c: Is at least one ticket vending machine currently working? 15d: Is there a help button/option available on the TVM? 15e: In what ways can you obtain assistance via the help button/option?

N = 560 (15c), N = 552 (15d), N = 261 (15e)

Help Point Availability

Verbatims regarding help point answered/not answered options

Is there a help point available on platform/in ticket hall and is it answered?

- Yes, available and answered: 52%
- Yes, available but not answered: 7%
- Cannot see a help point: 41%
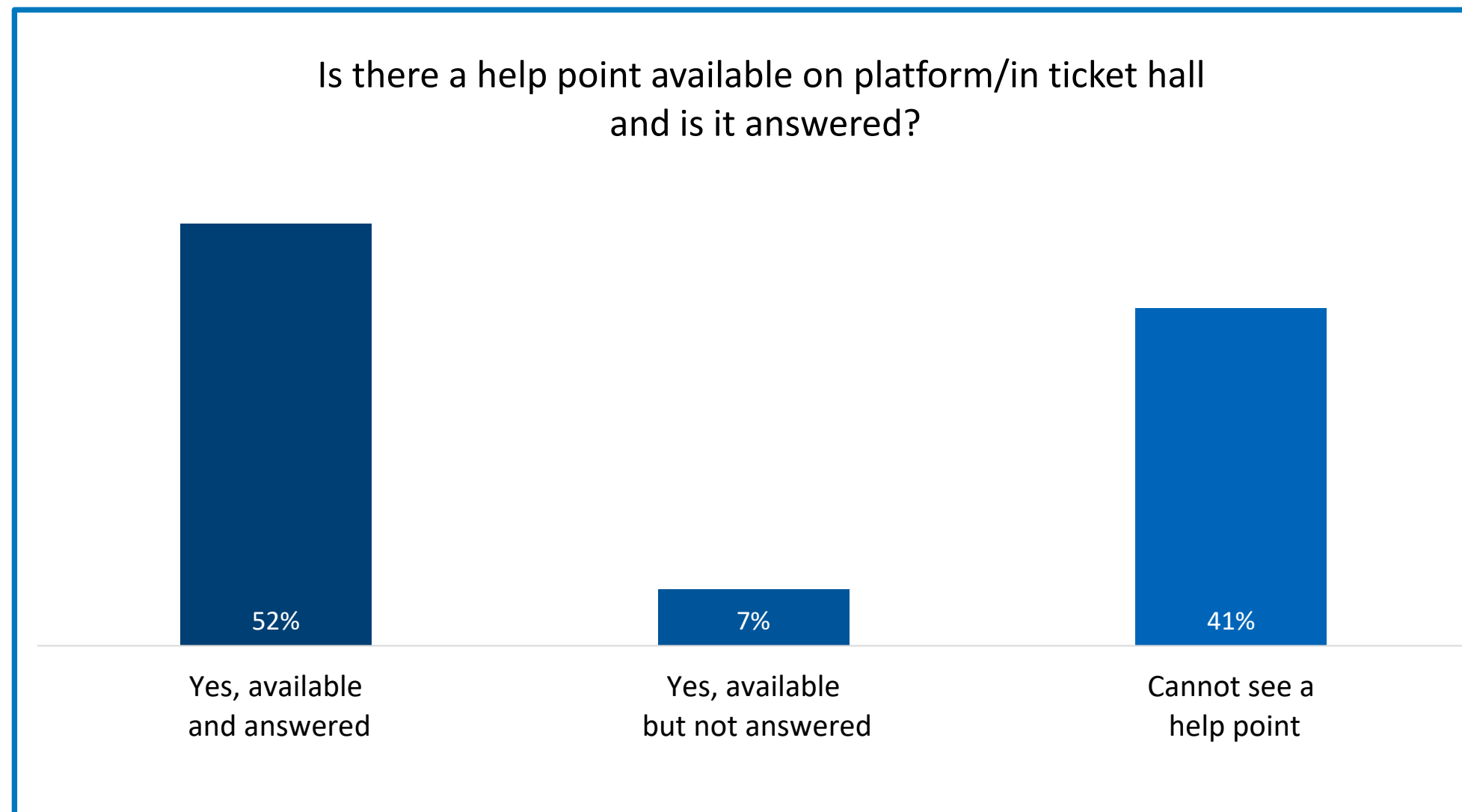
"After a short wait, I was able to get the information for the next train. The quality wasn't great though, and had to listen very intently to discern the answer. Very crackly and noisy response."
Wool, South Western Railway

"The help point didn't work when I pushed the button to ask for the next train."
Luxulyan, Great Western Railway

"I pushed the button and someone answered me straight away, within 10 seconds."
Northfleet, Southeastern

"It's not that it isn't answered, it's actually broken. It has a constant humming/buzzing sound and nothing happens when the button is pressed."
Clock House, Southeastern

Q60a: Is there a help point available in the station building/approach/on this platform, and is it answered? 60b Please explain why you have given this answer
N = 29 (60a) , N = 17 (60b)

Platform Access: Ticket Gates

Ticket gates working as planned?

Fully operational (closed with staff)	Fully operational (open w/o staff)	Partially operational (some gates can't be used)	Non-operational	No ticket gates at this station
18%	6%	1%	0.15%	75%
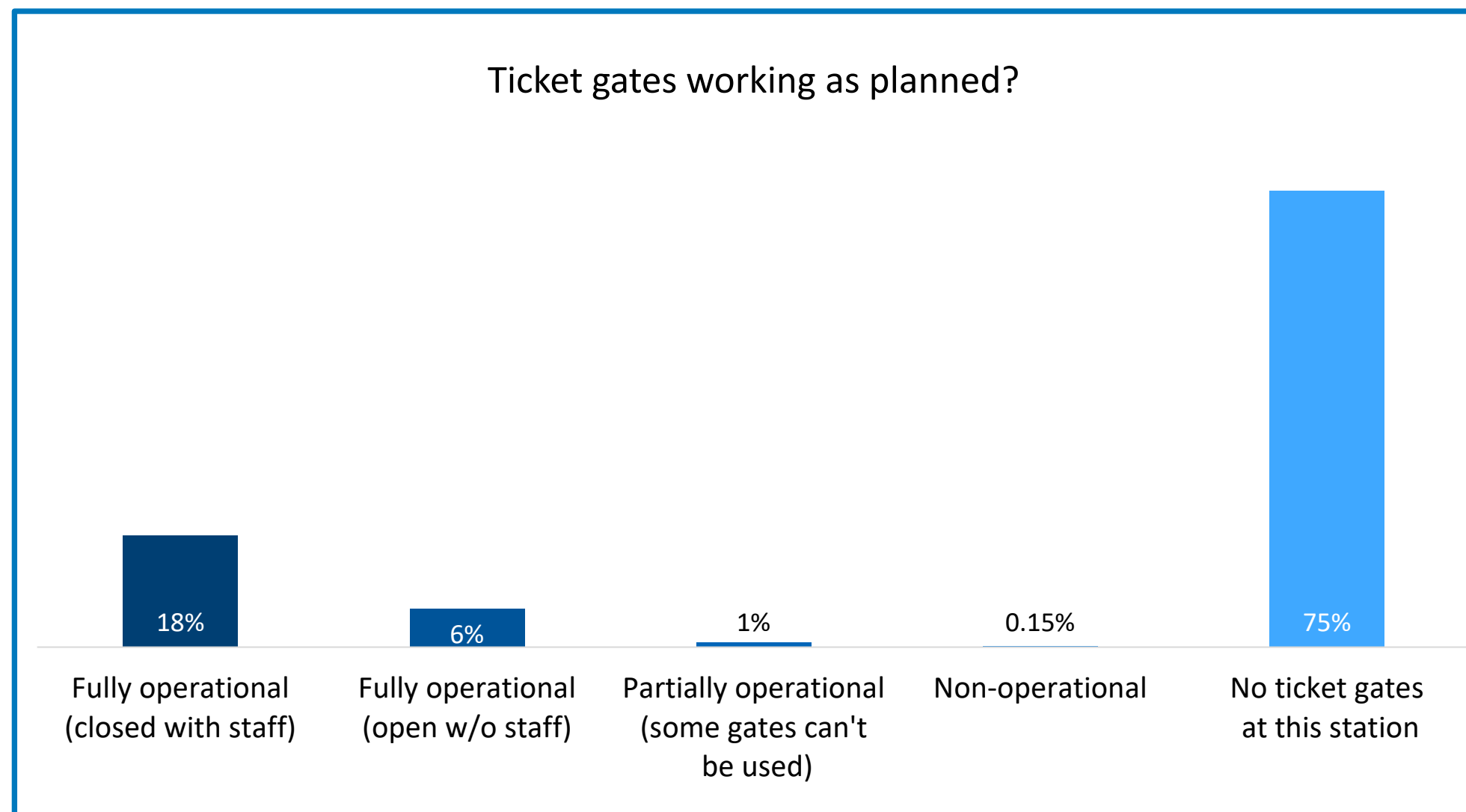
Verbatims regarding partial/non-operational ticket gates

"There were staff present however, the large accessible gate was open with an X sign showing on either side so it looked as though it was disabled by staff. It wouldn't be possible to touch out of this particular gate."
Slough, Great Western Railway

"One out of the four gates are being repaired on platform 1's entrance."
Grays, c2c

"Although all the gates are working, one of the doors of the accessible gate is slightly broken. The door keeps swinging open and shut slightly as though the door mechanism is glitching."
Tooting, Thameslink

Q35a: To what extent are the ticket gates working as planned? 35b: Please explain why you have given this answer.
N = 680 (35a), N = 6 (35b)

Overall Station Rating Scores
Platform: Customer Information Screens

Platform information screens were on average, excellent, with many TOCs at, or close to, 100% rated excellent. However, Chiltern Railways, Great Western Railway (GWR) and West Midlands Trains (WMT) had a few poor to moderate screens on their platforms and Southeastern and West Midlands Trains also had a high proportion (more than 75%) that are rated 'only' as good. Customer Information Screens in ticket halls/concourses had similar scores to on the platform.

■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Note
All TOCs	3%	20%	77%	
Avanti WC		31%	69%	**16
c2c	8%		92%	**25
Chiltern Rail	9%	17%	74%	
EMR	4%	4%	93%	**28
Great Northern		11%	89%	**9
Greater Anglia		8%	92%	
GWR	8%	12%	80%	
LNER			100%	**7
Northern	2%	2%	96%	
Southeastern	2%	78%	21%	
Southern	2%	8%	90%	
SWR	2%	7%	92%	
Thameslink		10%	90%	**20
TPE			100%	**19
WMT	8%	76%	16%	
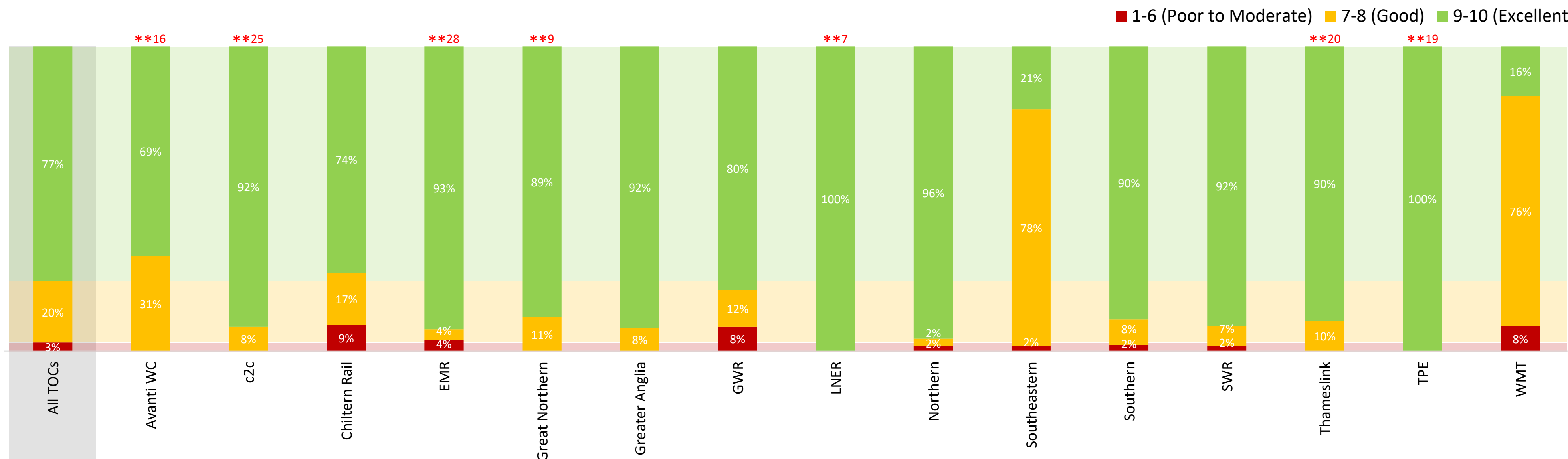
Q54a: To what extent are the customer information screens working on the platform?

Total N = 598 (All TOCs) **indicates small sample size (under 30 stations)

Overall Station Rating Scores
Public Announcement Clarity: Sound

Most stations across all TOCs perform well regarding the sound (and speech) of their public announcement system. However, WMT and Southeastern had far fewer excellent ratings than other TOCs, and WMT was rated poor to moderate more than four times the average for all TOCs.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Note
All TOCs	4%	21%	75%	
Avanti WC		31%	69%	**16
c2c	5%	14%	82%	**22
Chiltern Rail	9%	30%	61%	**23
EMR		21%	79%	
Great Northern	50%		50%	**2
Greater Anglia		5%	95%	
GWR			100%	
LNER		17%	83%	**6
Northern	3%	4%	93%	
Southeastern	11%	68%	20%	
Southern		5%	95%	
SWR		12%	88%	
Thameslink	22%	33%	44%	**9
TPE			100%	**19
WMT	19%	69%	11%	
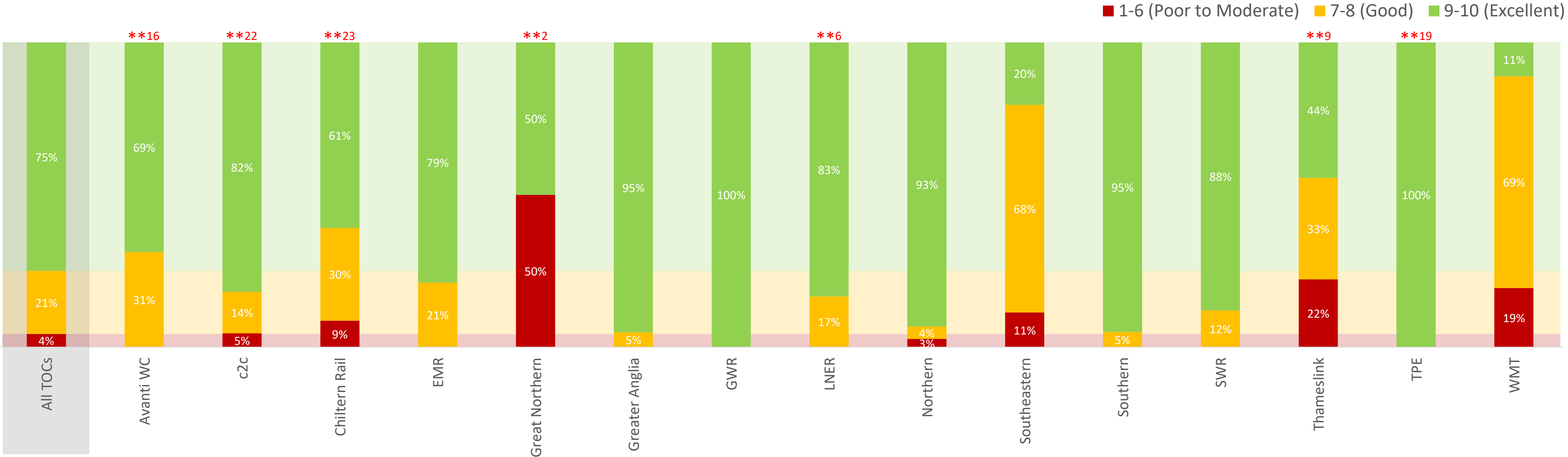
Q56a: To what extent was the announcement made clearly and audibly?

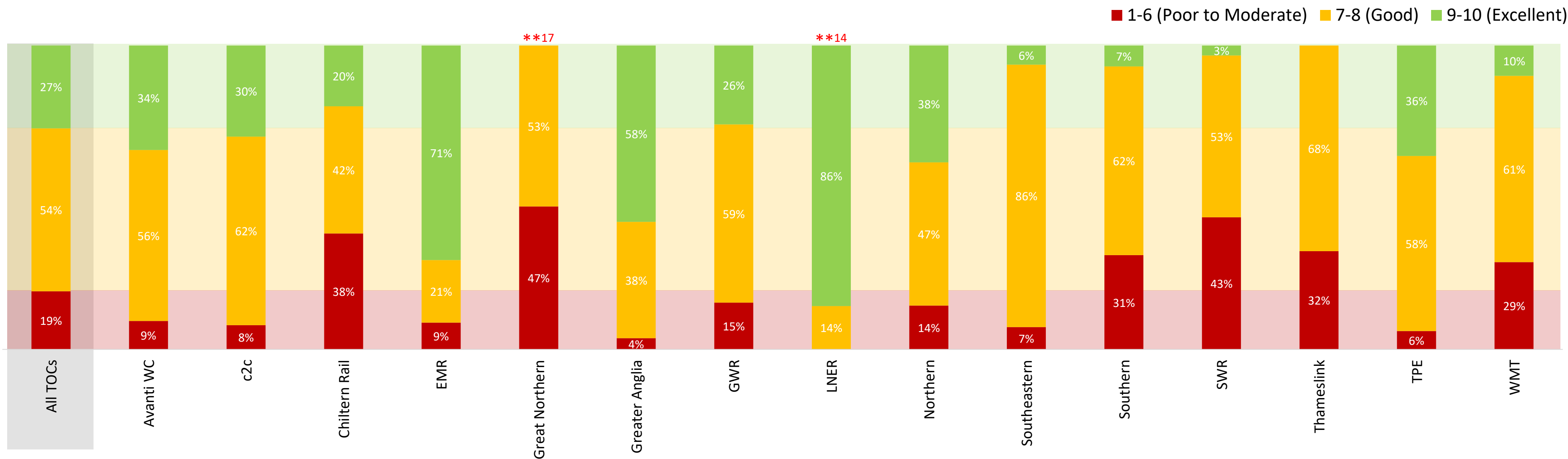Total N = 459 (All TOCs) **indicates small sample size (under 30 stations)

Overall Station Rating Scores
Station Building/Platform: Upkeep and Repair

Upkeep and repair of stations was rated as good (54%) or excellent (27%) across all TOCs combined. EMR, Greater Anglia, and LNER received more than double the average excellent ratings. However, Chiltern Railways, Great Northern, Southern, SWR, Thameslink and WMT all had much higher than average poor to moderate ratings and much lower than average excellent ratings.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

	All TOCs	Avanti WC	c2c	Chiltern Rail	EMR	Great Northern **17	Greater Anglia	GWR	LNER **14	Northern	Southeastern	Southern	SWR	Thameslink	TPE	WMT
9-10 (Excellent)	27%	34%	30%	20%	71%		58%	26%	86%	38%	6%	7%	3%		36%	10%
7-8 (Good)	54%	56%	62%	42%	21%	53%	38%	59%	14%	47%	86%	62%	53%	68%	58%	61%
1-6 (Poor to Moderate)	19%	9%	8%	38%	9%	47%	4%	15%	14%	14%	7%	31%	43%	32%	6%	29%

Q30a/52a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the upkeep and repair of this station? On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the upkeep and repair of this platform? Total N = 1050 ratings **indicates small sample size (under 30 stations)
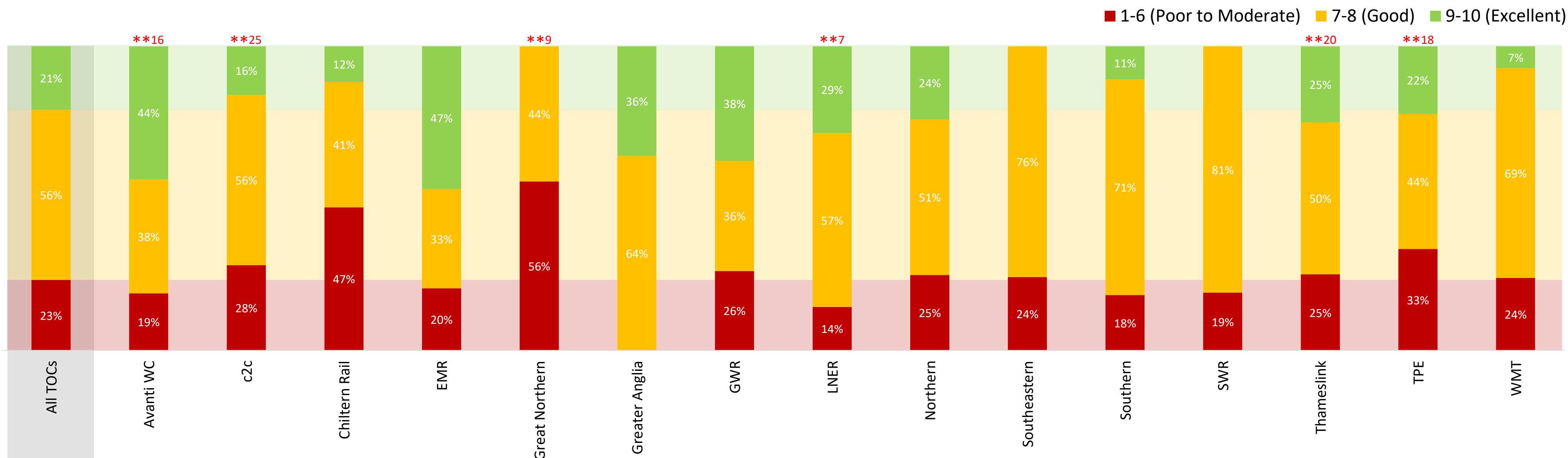
Overall Station Rating Scores
Platform Seating Condition

Avanti West Coast and EMR had lower than average poor to moderate ratings but more than double the average excellent scores. Chiltern Railways in particular had a very high % of poor to moderate scores. Great Northern had a high number of poor to moderate ratings (from a small sample). Southeastern, SWR and WMT had few or no excellent ratings. Greater Anglia had no poor to moderate ratings.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

	All TOCs	Avanti WC **16	c2c **25	Chiltern Rail	EMR	Great Northern **9	Greater Anglia	GWR	LNER **7	Northern	Southeastern	Southern	SWR	Thameslink **20	TPE **18	WMT
9-10 (Excellent)	21%	44%	16%	12%	47%	44%	36%	38%	29%	24%		11%		25%	22%	7%
7-8 (Good)	56%	38%	56%	41%	33%		64%	36%	57%	51%	76%	71%	81%	50%	44%	69%
1-6 (Poor to Moderate)	23%	19%	28%	47%	20%	56%		26%	14%	25%	24%	18%	19%	25%	33%	24%

Q45a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the condition of the seats on this platform?

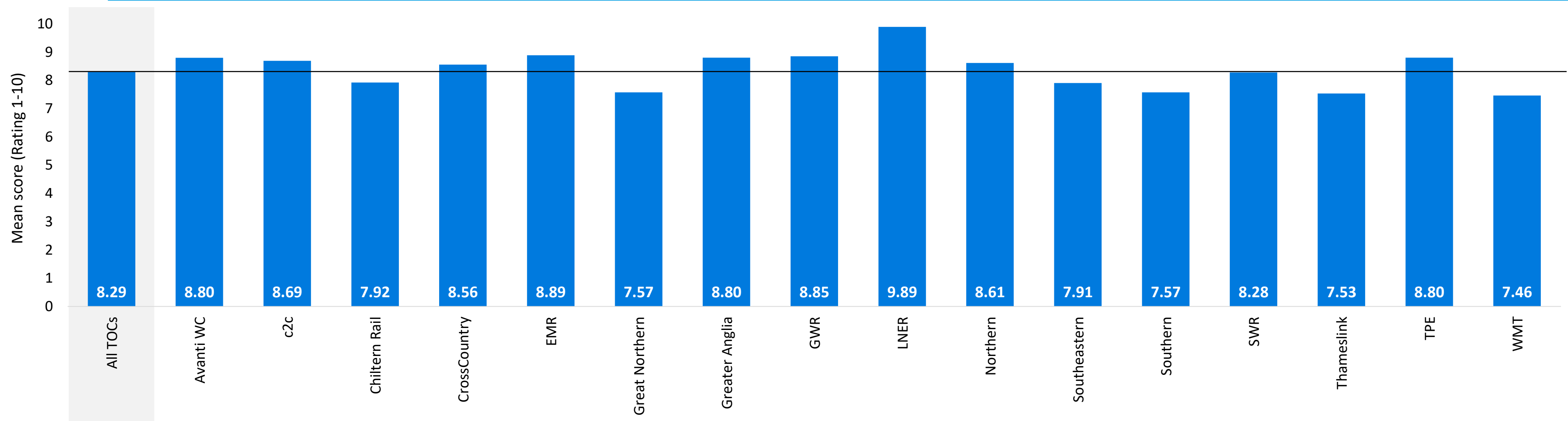Total N = 643 (All TOCs) **indicates small sample size (under 30 stations)

—

Train Results

Overall Train Rating Scores

All Ratings Combined: By TOC

Mean scores for trains show more variation than for stations. Some TOCs – particularly LNER – post very high scores but from a low base in each individual question. However, once multiple questions have been combined, these means then represent a high enough number of ratings to be meaningful measures of TOC performance. As a 'high sample' TOC, East Midlands Railway do very well, as do GWR and Greater Anglia. Thameslink post many very low scores but compensate with a number of excellent scores; whereas West Midlands Trains often score 7-8 out of 10, but rarely 9-10 or 1-6 and, thus, have a mean score similar to Thameslink.
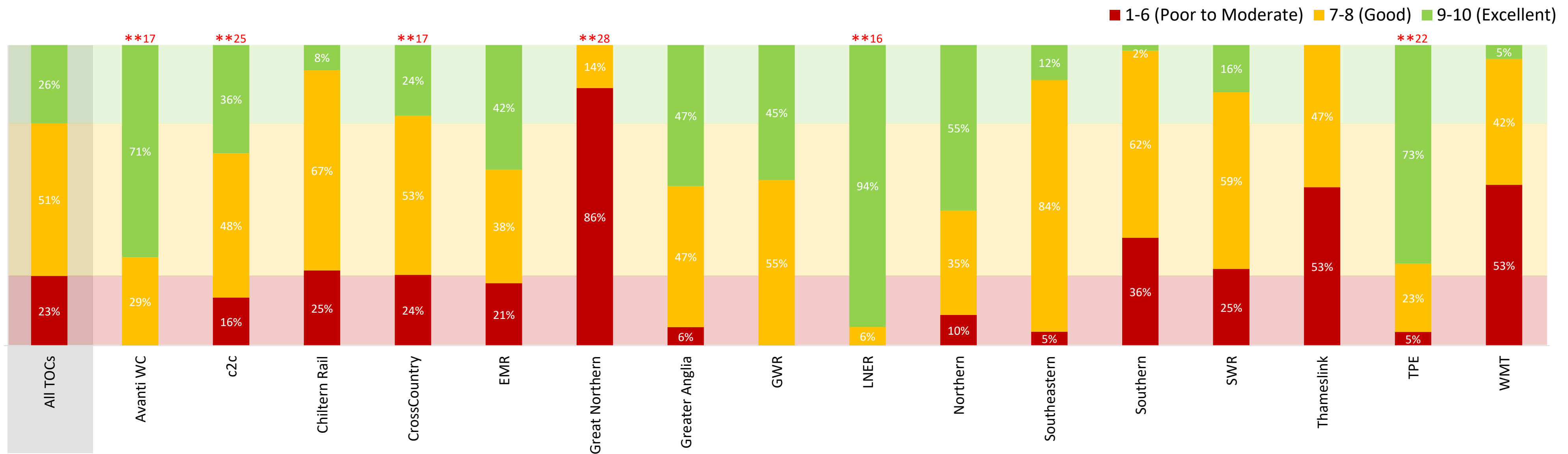

All TOCs	Avanti WC	c2c	Chiltern Rail	CrossCountry	EMR	Great Northern	Greater Anglia	GWR	LNER	Northern	Southeastern	Southern	SWR	Thameslink	TPE	WMT
8.29	8.80	8.69	7.92	8.56	8.89	7.57	8.80	8.85	9.89	8.61	7.91	7.57	8.28	7.53	8.80	7.46

Y-axis: Mean score (Rating 1-10)

Q2a/3a/4a/5a/6a/10b/11a/12a/14a/15a/17a/23a/24a

Total N = 700 (All TOCs). Scores calculated are means of mean scores for each of the above rating questions.

Overall Train Rating Scores
Carriage Interior Condition: Floor Cleanliness

Great Northern post a noticeably low score from 28 ratings. There are also problems with floor cleanliness at Southern, Thameslink and WMT. Avanti West Coast, LNER and TPE are all rated highly but from a low sample size. GWR is a notable performance for a high sample TOC, with no poor/moderate ratings. Northern gets a high 55% excellent and under half average poor/moderate ratings.

■ 1-6 (Poor to Moderate)   ■ 7-8 (Good)   ■ 9-10 (Excellent)

|	All TOCs	Avanti WC **17	c2c **25	Chiltern Rail	CrossCountry **17	EMR	Great Northern **28	Greater Anglia	GWR	LNER **16	Northern	Southeastern	Southern	SWR	Thameslink	TPE **22	WMT
9-10 (Excellent)	26%		36%	8%	24%	42%	14%	47%	45%		55%	12%	2%	16%	47%	73%	5%
7-8 (Good)	51%	71%	48%	67%	53%	38%		47%	55%	94%	35%	84%	62%	59%	53%	23%	42%
1-6 (Poor to Moderate)	23%	29%	16%	25%	24%	21%	86%	6%		6%	10%	5%	36%	25%		5%	53%

Q2a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the cleanliness of this train carriage floor?

Total N = 700 (All TOCs) **indicates small sample size (under 30 train journeys)
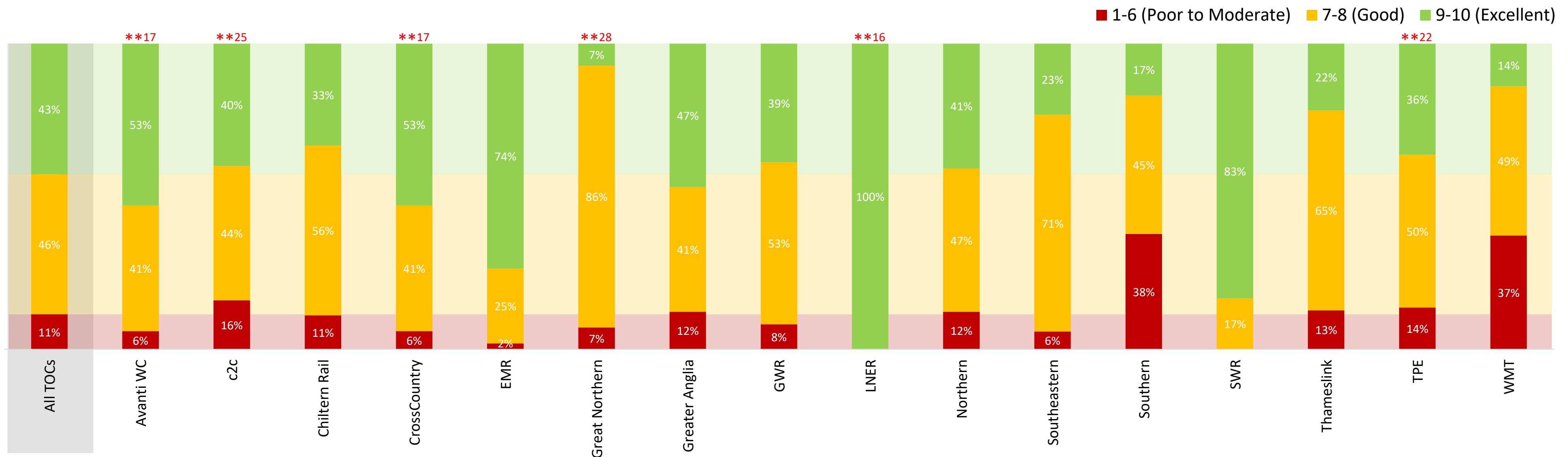
Overall Train Rating Scores

Carriage Interior Condition: Litter
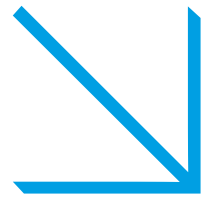
South Western Railway performs particularly well in terms of handling litter on board. However, Southern and West Midlands Trains do not: in nearly two in five cases, the litter situation is rated as poor/moderate. East Midlands Railway perform very strongly (74% excellent) as do LNER (100% excellent) but from a smaller base (n=16).

■ 1-6 (Poor to Moderate)   ■ 7-8 (Good)   ■ 9-10 (Excellent)

	All TOCs	Avanti WC **17	c2c **25	Chiltern Rail	CrossCountry **17	EMR	Great Northern **28	Greater Anglia	GWR	LNER **16	Northern	Southeastern	Southern	SWR	Thameslink	TPE **22	WMT
9-10 (Excellent)	43%	53%	40%	33%	53%	74%	7%	47%	39%	100%	41%	23%	17%	83%	22%	36%	14%
7-8 (Good)	46%	41%	44%	56%	41%	25%	86%	41%	53%		47%	71%	45%	17%	65%	50%	49%
1-6 (Poor to Moderate)	11%	6%	16%	11%	6%	2%	7%	12%	8%		12%	6%	38%		13%	14%	37%

Q6a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the level of litter inside this carriage?

Total N = 700 (All TOCs) **indicates small sample size (under 30 train journeys)
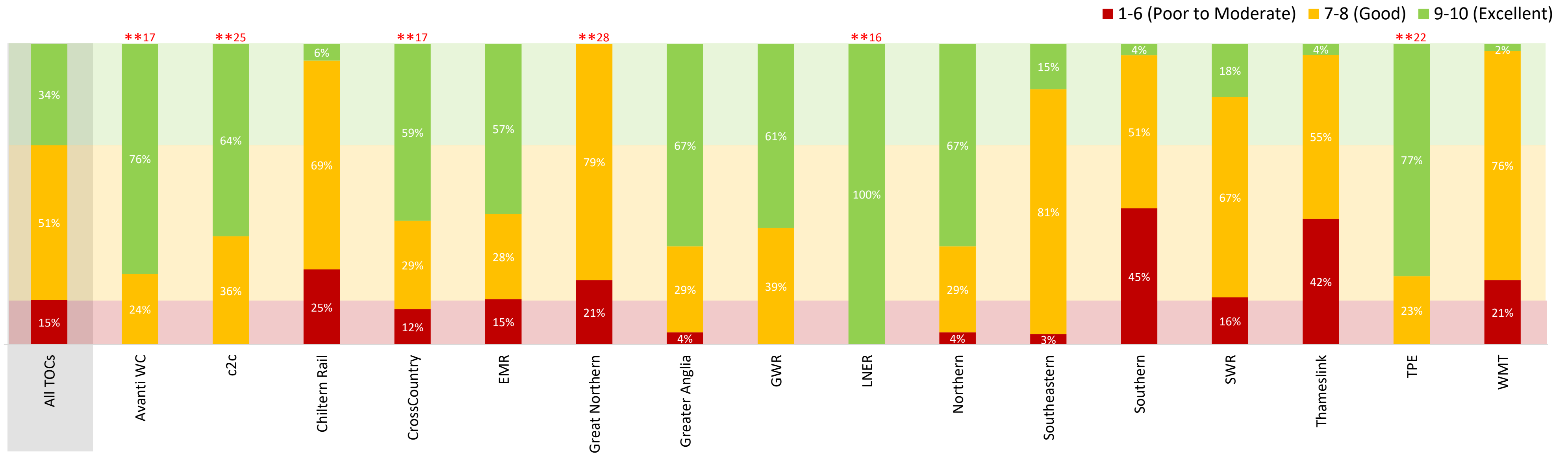
Overall Train Rating Scores

Seating Cleanliness

Avanti West Coast, c2c, GWR, LNER and TransPennine Express achieve high 1-10, and no 1-6, scores. Greater Anglia and Northern also post high 1-10 ratings – with minimal 1-6. Thameslink and Southern have not managed to deliver this, with both posting poor to moderate ratings around three times higher than average, with minimal excellent ratings.

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Note
All TOCs	15%	51%	34%	
Avanti WC	24%	76%		**17
c2c	36%	64%		**25
Chiltern Rail	25%	69%	6%	
CrossCountry	12%	29%	59%	**17
EMR	15%	28%	57%	
Great Northern	21%	79%		**28
Greater Anglia	4%	29%	67%	
GWR		39%	61%	
LNER			100%	**16
Northern	4%	29%	67%	
Southeastern	3%	81%	15%	
Southern	45%	51%	4%	
SWR	16%	67%	18%	
Thameslink	42%	55%	4%	
TPE		23%	77%	**22
WMT	21%	76%	2%	

Q3a: On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the cleanliness of the seats in this train carriage?

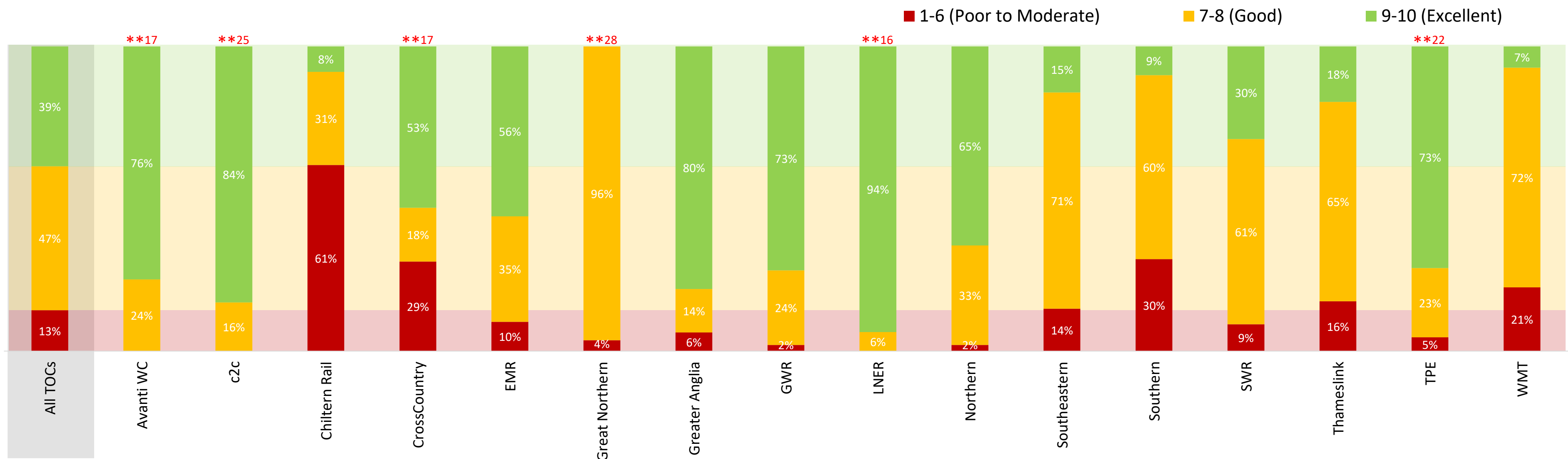Total N = 699 (All TOCs) **indicates small sample size (under 30 train journeys)

Overall Train Rating Scores
Seating Condition

Whilst CrossCountry deliver clean seats, the condition of some of their seats is not as good – they scored poor/moderate more than twice the average. Southern, West Midlands Trains and, especially, Chiltern Railways stand out as having seats in poor condition. Avanti West Coast, c2c and LNER received no 1-6 scores in any of their assessments; Greater Anglia and GWR also score very highly.
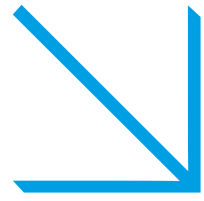

Legend: ■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

TOC	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Note
All TOCs	13%	47%	39%	
Avanti WC	24%	76%		**17
c2c	16%	84%		**25
Chiltern Rail	61%	31%	8%	
CrossCountry	29%	18%	53%	**17
EMR	10%	35%	56%	
Great Northern	4%	96%		**28
Greater Anglia	6%	14%	80%	
GWR	2%	24%	73%	
LNER	6%	94%		**16
Northern	2%	33%	65%	
Southeastern	14%	71%	15%	
Southern	30%	60%	9%	
SWR	9%	61%	30%	
Thameslink	16%	65%	18%	
TPE	5%	23%	73%	**22
WMT	21%	72%	7%	

Q4a: Now focusing specifically on the physical condition of the seats, on a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the seats in this train carriage?

Total N = 699 (All TOCs) **indicates small sample size (under 30 train journeys)

'Excellent' images (10/10 scores)

Avanti West Coast

LNER

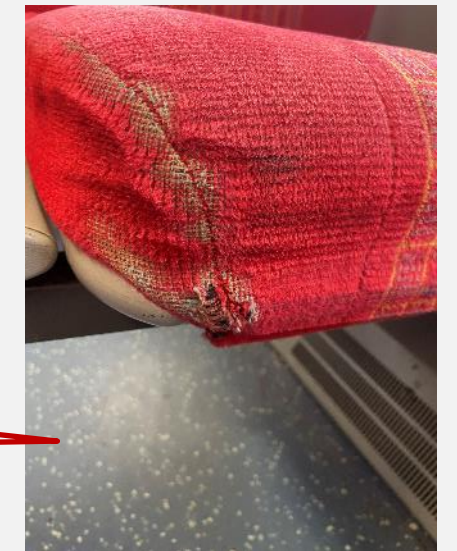### 'Poor' to 'Moderate' images and verbatims (1-6 scores)

"There's lots of dirt built up in the corners of the chairs as well as some light staining on them too."
Southern

"Most of the chairs in this carriage have wear on the corners, some worse than others. The stuffing is coming out and you can see a fair amount of frayed edges."
South Western Railway

"Old, faded, frayed and torn in places - most of the seats were in a poor state."
Chiltern Railways

Q3b/3c/4b/4c: If your score is 5 or lower, please explain why you have given this rating. Please upload one image to illustrate your rating (scores of 1-5 or 10).

Overall Train Rating Scores
Passenger Information Screens

Generally, the passenger information screens work well inside trains. In this case, not scoring excellent can be seen as low performance. Clearly, Southeastern and West Midlands Trains are not as advanced as other TOCs. CrossCountry and East Midlands Railway are generally excellent but have a higher than average series of poor to moderate ratings.

■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

	1-6 (Poor to Moderate)	7-8 (Good)	9-10 (Excellent)	Small sample
All TOCs	5%	19%	76%	
Avanti WC		18%	82%	**17
c2c	4%	20%	76%	**25
Chiltern Rail	8%		92%	
CrossCountry	18%	12%	71%	**17
EMR	25%		75%	
Great Northern			100%	**28
Greater Anglia	6%	4%	90%	
GWR	6%	21%	72%	
LNER			100%	**16
Northern	2%	2%	96%	
Southeastern	2%	83%	15%	
Southern	8%	8%	85%	
SWR			100%	
Thameslink		4%	96%	
TPE	10%		90%	**21
WMT	9%	72%	19%	
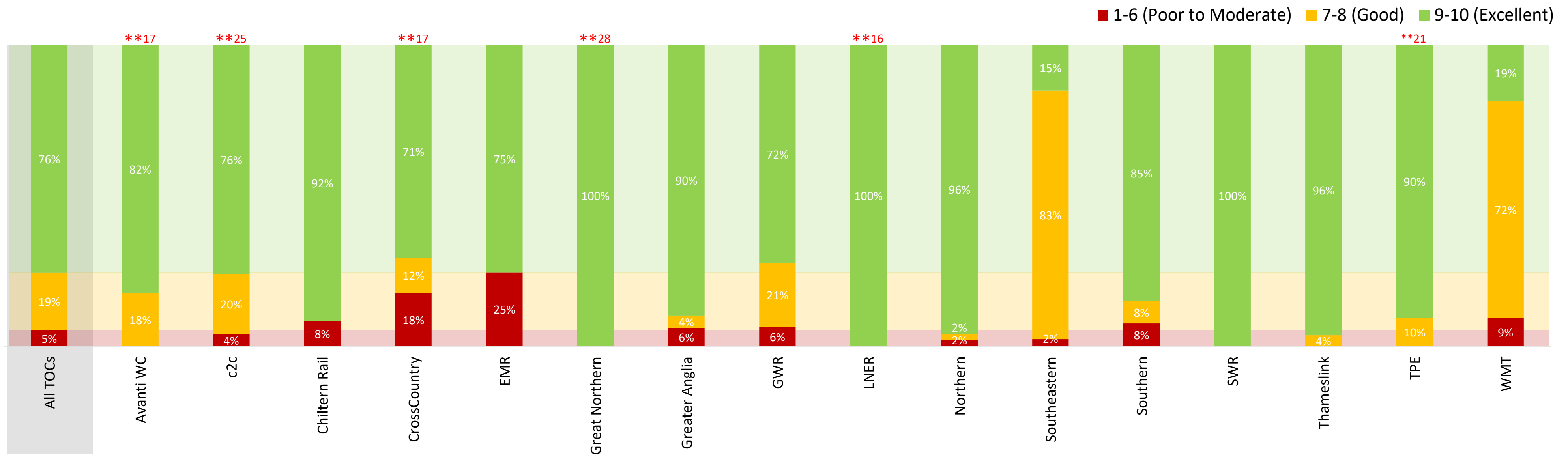
Q12a: To what extent are the passenger information screen(s) working on this train?

Total N = 694 (All TOCs) **indicates small sample size (under 30 train journeys)

Overall Train Rating Scores
Public Announcement Clarity: Speech

In line with the findings around sound of public announcements, West Midlands Trains and Southeastern are the two TOCs that might improve public announcement on-train, as both speech and sound receive a higher proportion of 7-8 scores, whereas other TOCs' assessments are near uniformly excellent.

■ 1-6 (Poor to Moderate) ■ 7-8 (Good) ■ 9-10 (Excellent)

	All TOCs	Avanti WC	c2c	Chiltern Rail	CrossCountry	EMR	Great Northern	Greater Anglia	GWR	LNER	Northern	Southeastern	Southern	SWR	Thameslink	TPE	WMT
		**16	**25		**17		**28			**16						**22	
9-10	77%	94%	100%	94%	82%	85%	100%	96%	83%	94%	87%	18%	92%	86%	98%	100%	16%
												82%					77%
7-8	21%	6%		3%	18%	13%		4%	15%	6%	6%		4%	10%	2%		
1-6	2%			3%		2%			2%		6%		4%	4%			7%
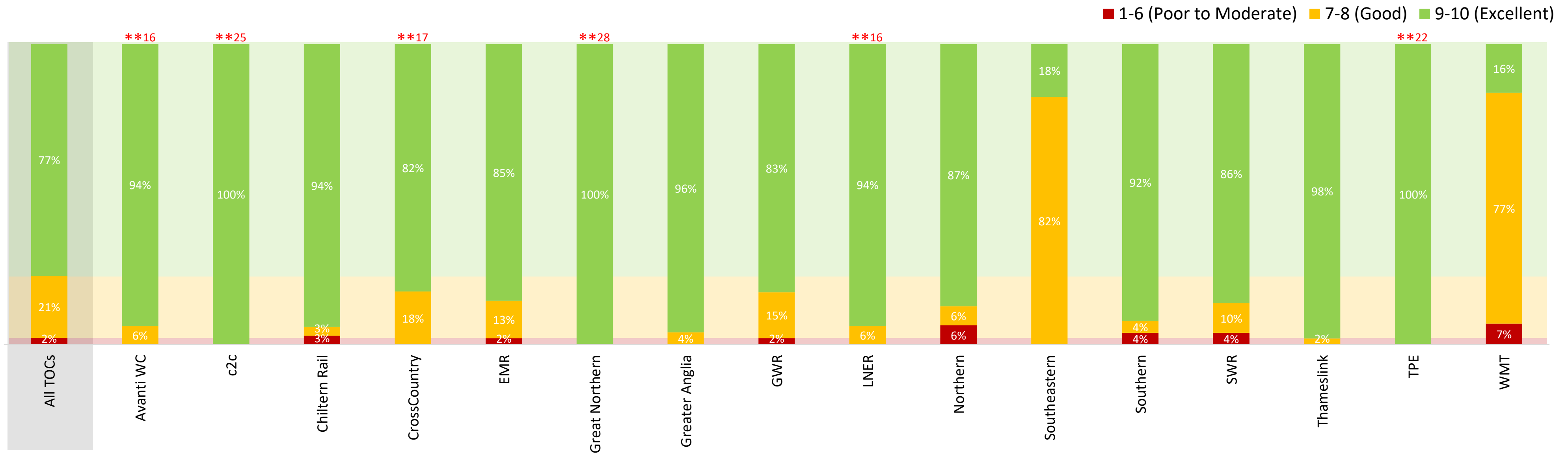
Q15a: To what extent did the announcer speak clearly?

Total N = 678 (All TOCs) **indicates small sample size (under 30 train journeys)

Summary of key findings

The overall findings from this research were positive in that the overall average (or mean) score for stations, and also on trains for all appropriate measures combined both approximately scored 8.30 (scores of 7 or 8 were classed as 'good', whilst 9 or 10's were 'excellent').

However, there was quite a bit of variation between individual TOCs: On stations the average scores varied between 9.05 (LNER) and 7.50 (Great Northern). On trains there was even more variation, ranging from 9.89 (LNER) and 7.46 (West Midlands Trains). Operators which seemed to struggle more than others were Great Northern and West Midlands Trains (at stations and on trains). Thameslink and Southern were two TOCs that also scored less well – especially on train factors.

Generally high average scores meant the TOC scored highly on most ratings, whereas lower scores meant a TOC performed less well on several factors.

In terms of factors where scores were particularly high; both at station and on train public announcements scored highly (both sound and speech), and passenger/customer information screens (more than 70% recorded as 'excellent' for all these factors).

Factors where lower scores were recorded (where there appeared to be room for improvement) were station buildings – upkeep and repair, platform seating condition, train floor cleanliness, toilet cleanliness, and exterior train cleanliness. All these factors had less than 30% scoring 'excellent'.

Process for ensuring quality and consistency

To ensure quality and consistency between all assessors, the agency used the following process:

- They had a small team of 12 assessors, who conducted the 1,400 audits. A larger team would have allowed for the possibility of wider variance in scoring and, thus, a longer period of training/re-training to adjust to a consistent delivery.

- Each assessor covered at least one TOC; several conducted audits across two or more TOCs. This allowed assessors to build awareness quickly of the differences between stations and train services and, thus, identifying where scoring might need to be adjusted to maintain a level of consistency.

- Assessors (in groups of 4-6) each took part in a 3-hour training session to allow plenty of opportunity for individual questions and clarifications.

- After the training, each assessor conducted several 'dummy runs' where they tested the questionnaire and submitted results for the fieldwork team to check. Answers were reviewed for understanding of each question and consistency of scoring.

- For each of the 1-10 scoring questions in the audit questionnaire, assessors were provided with a framework of typical reasons for giving that score (see examples on next page). This helped to ensure that they understood the full context of that question and their scoring was in alignment with all other assessors.

- Assessors who gave a score of 1-5 (poor to moderate) for any rating question were required to submit a comment and photo as evidence of why they gave that score. From this we could check that scoring remained consistent with other audits. Assessors also had the option to upload a photo for a score of 10 (excellent) to indicate best practice.

- Before the audit data was sent to 7th Sense for analysis, their fieldwork agency, LIVE, ran daily quality checks to ensure it was both complete and valid. LIVE's account manager checked every submitted report for accuracy and completeness. If they found any issues (e.g. anomalous scoring), they contacted the assessor for clarification or to ask them to edit their submission. LIVE's account manager reviewed all audits and thus was able to ensure that scoring was consistent across all TOCs, based on the evidence provided by the assessors (verbal feedback, verbatim comments and photo evidence provided in each assessment).

- As a second layer of quality assurance, 7th Sense's project manager reviewed every piece of assessor feedback (answers, comments and photos) for completeness and consistency. Any queries were flagged with LIVE and the assessor was contacted for further clarification/the score was adjusted accordingly.

Auditor Quality and Consistency Checks

Examples of 1-10 rating score frameworks

Please see below some example scoring frameworks/guidance for assessors for the 1-10 rating questions.

Example question from station audit questionnaire.

Q11a TICKET HALL/CONCOURSE: CLEANLINESS
On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the cleanliness of the ticket hall/concourse?

Please include the following in your assessment: all dust, dirt, staining, marks, chewing gum/gum residue, body fluids, spillages, and bird/animal droppings etc. Do not include graffiti or litter

Excellent: no evidence of dirt, marks or impact on cleanliness from any of the sources listed	10
	9
Good: minor evidence of dirt, marks or impact on cleanliness from the sources listed	8
	7
Moderate: some evidence of dirt, marks or impact on cleanliness from the sources listed	6
	5
Poor: extensive evidence of dirt, marks or impact on cleanliness from the sources listed	4
	3
Very poor: major evidence of dirt, marks or impact on cleanliness from the sources listed	2
Unacceptable: extreme evidence of dirt, marks or impact on cleanliness from the sources listed, with risk of soiling upon contact	1

Example question from train audit questionnaire.

Q11a INSIDE TRAIN: CONDITION/FUNCTION OF CUSTOMER TOILETS
On a scale of 1-10 (where 1 = unacceptable, 10 = excellent) how do you rate the condition/function of the toilet on this train?

Excellent: working fully (including door, locks, the toilet itself, water, hand-dryer if present). Soap, mirror, toilet paper, and towels (if no dryer) all present and plentiful; bins easily accessible and not full	10
	9
Good: working fully (including door, locks, the toilet itself, water, hand-dryer if present). Soap, toilet paper, and towels (if no dryer) all present but may be starting to run low (will need replenishing soon); bins easily accessible but will need emptying soon	8
	7
Moderate: working well (including door, locks, the toilet itself, water). Soap and toilet paper available, but no provision to dry hands (dryer not working and/or no towels); bin easily accessible but will need emptying soon	6
	5
Poor: any problems with closing/opening door or lock, or with flushing toilet OR either soap, toilet paper, or water are missing/not working OR bins too full, making the area unpleasant	4
	3
Very poor: any problems with closing/opening door or lock, or with flushing the toilet OR two or more of soap, toilet paper, or water are missing/not working AND bin too full (or not present), making the area unpleasant	2
Unacceptable: any problems with closing/opening the door or lock, or with flushing the toilet AND two or more of soap, toilet paper, or water are missing/not working AND bin too full (or not present), making the area unpleasant	1

Analysis of audit data, verbatim and photographs

- For every station and train audit, the assessor completed an online quantitative questionnaire. It included a number of open-ended responses where assessors could provide extra insight on the reasoning behind their answers (particularly if giving a 1-5 score for the ratings questions).

- After the data was checked for quality and consistency, data tables were created with the individual TOCs as a crossbreak. One set of tables was created for the station audits; one set for the train audits.

- In the reporting, for the overall mean score slides, we combined all rating questions (32 in the station audits; 13 in the train audits) to calculate a mean of the means scores – this then allowed us to compare overall means by TOC/other measures.

- The content of the rest of the report was then grouped by assessor journey touchpoints (e.g. station car park, station entrance, ticket hall etc..). Some ratings questions were shown as single questions on a slide; others were combined. The combined ratings were often because assessors only needed to rate a touchpoint once, e.g. the customer toilets at a station. If toilets were in both the ticket hall AND on the platform, assessors only needed to provide a rating for one location. The findings on the combined slides are therefore illustrating the scores for the toilets as a whole at that station.

- If an assessor gave a rating of 10 they had the option to upload a photo. The photos in the report illustrating this score were selected based on the commentary provided at the top of the rating slides, to give a flavour of excellent performing TOCs. These are simply a selection to show a range of points – the complete database of photographs submitted by assessors has been provided to the relevant TOCs separately to our reporting.

- If an assessor gave a rating of 1-5, they were required to provide a comment and a photo as evidence. Again, a selection of photos and comments have been provided in the reporting to represent the commentary on the rating slides – a full database of photos has been provided to each individual TOC.

- Other, non-rating score, slides are reported at a total sample level. Full data tables and raw data have been provided by 7th Sense, should the DfT/Transport Focus wish to conduct further analysis by TOC. Verbatim comments/photos provided on these slides have been chosen as they illustrate the range of themes mentioned across the assessor comments – the full list of comments can be found in the raw data files.

- All photos and verbatim used in the report were merely chosen for illustrative purposes, purely for picture quality and appropriateness/articulation of response.

- Where open-ended questions asked for factual responses (not assessor opinion) – for example "What type of Covid-19 signage did you see?" – a thematic codeframe was created. We counted the number of times assessors saw each type of signage and those counts were converted into percentages of the total number of items mentioned.