

Railfuture response to Office of Rail and Road consultation on a new assessment framework for Rail Passenger Assistance: benchmarking train operators' performance

(1) Do you agree with our plan to publish a report that benchmarks operators' performance levels in providing assistance to rail passengers? Explain your answer.

Yes. The significant number of partial or full assistance failures (24%) make it essential to publish such a report as soon as practicable. Several high profile failures over the last few months reinforce the need to demonstrate that action is being taken.

(2) Are the two headline performance categories (delivery and capability to improve) clear and meaningful? Is there anything else that should be included?

The major issue in terms of delivery is communication – both from operators to passengers who have requested assistance. Communication of assistance requests within each operator's teams and between operators where connections are involved are also a major area of concern. This has been demonstrated by several communication failures where passengers have been abandoned on trains or on platforms. Some operators' staffing levels and staff locations will have a big impact on their capability to improve. This is particularly important for passengers requiring assistance to start or end their journey, or to change trains, at unstaffed stations. Should staffing levels be built into the assessment?

(3) Are the performance level descriptors clear and meaningful? How should we name these performance levels (if at all)? Is three the right number of levels?

We strongly support focussing on the percentage of respondents who report not having received any of the assistance that they had booked. However, how will the experience of those who only received assistance for part of their journey be taken into account? (this is particularly important where passengers need to change trains and/or operators). Three levels would appear adequate. Upper, middle and lower can be used.

(4) Is the proposed reliability metric clear and sufficiently reflective of reliability performance to allow comparison?

The basis of the weighting is unclear. Data for the first year (2021-22) will be affected by the impact of Covid-19 on rail use. Has this been taken into account?

(5) Are the proposed reliability performance level thresholds set appropriately?

The limited number of levels does seem quite restrictive. Perhaps four levels, with a third one ranging from 11% to 15% and a fourth from 16% to 20% might be more meaningful?

(6) Is the minimum sample size threshold for inclusion in this area of the assessment set appropriately?

The sample size of 50 seems quite a limited one for a meaningful assessment – how representative is it? It is concerning that, even with this limited number, Cross Country (a major

www.railfuture.org.uk www.railfuturescotland.org.uk www.railfuturewales.org.uk
www.railwatch.org.uk

national operator) is excluded. Use of assistance by passengers with this operator is likely to be fairly high, given the length of potential direct journeys and their attraction to elderly or disabled travellers. More data is clearly needed!

(7) Is the proposed Satisfaction metric clear, and sufficiently reflective of performance to allow comparison?

We have similar concerns about the metric as for the one for reliability (question 4).

(8) Are the Satisfaction performance level thresholds set appropriately?

It is surprising that the thresholds are so close together and thus not clear that they will provide a true picture of the operators' performance.

(9) Is the minimum sample size threshold for inclusion in this area of the assessment set appropriately?

We have similar concerns about the minimum sample size as we do for reliability - see response to question 6.

(10) Is the proposed Training metric clear (taking into account the relevant detail provided in the ATP guidance on mandatory training outcomes)?

The Training metric is unclear. It needs to have much clearer, measurable requirements.

(11) Are the Training performance level thresholds set appropriately?

The three proposed levels seem appropriate, but how these are measured remains a concern – see response to question 10.

(12) Taken together, do these criteria capture the key factors that determine an operator's capability to effectively manage and improve its provision of assistance? Is there anything missing, and is there anything listed that is not relevant?

We have concerns that all three levels rely on the operator conducting its own data quality assurance and using comparator data sources "where appropriate". There is a need for ORR to be involved in these processes and also to approve the comparator sources to be used. Data collection needs to distinguish journeys involving changes of train and each operator involved in the journey needs to record the full journey details. All three levels need to cover both booked and unbooked assistance. Collection of data on unbooked assistance is likely to be patchy at best – specific arrangements for this need to be set out by operators and agreed with ORR.

(13) Can these criteria be critically evaluated, based on the assessment of a proportionate evidence gathering, submission and assessment process?

It is not clear whether these criteria can be critically evaluated, as they appear to rely heavily on qualitative and subjective assessment methods.

(14) Please indicate how the assistance benchmarking framework might be used to even further advance equality of opportunity and reduce discrimination for groups protected under the Equality Act.

Some form of categorisation of assistance requests eg nature of help needed would be required.