Transport for London

Developing a reliability metric for LU customers

11046

September 2011



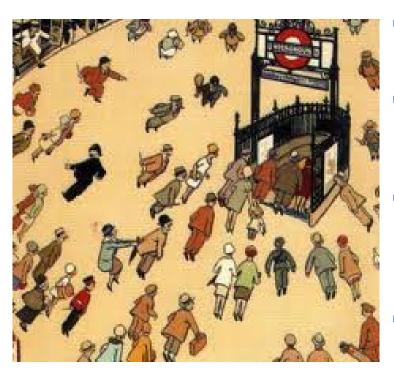
Confidentiality

- Please note that the copyright in the attached report is owned by TfL and the provision of information under Freedom of Information Act does not give the recipient a right to re-use the information in a way that would infringe copyright (for example, by publishing and issuing copies to the public).
- Brief extracts of the material may be reproduced under the fair dealing provisions of the Copyright, Designs and Patents Act 1988 for the purposes of research for noncommercial purposes, private study, criticism, review and news reporting.
- Details of the arrangements for reusing the material owned by TfL for any other purpose can be obtained by contacting us at enquire@tfl.gov.uk.
- Research conducted by 2CV

Agenda

11.00-11.15	Background, objectives, methodology and executive summary
11.15-11.35	Defining and managing reliability: the customers' perspective
11.35-11.45	Reliability in context: perceptions of LU and media backdrop
11.45-12.05	Towards a reliability metric: customer response to the metrics
12.05-12.20	Break / refreshments
12.20-12.35	Discuss findings and implications
12.35-1.15	Workstream groups to agree implications and actions

Background



- London Underground (LU) are working to improve reliability performance
- One workstream is specifically tasked with developing a key metric on reliability – that can be monitored over time and potentially communicated to customers
- We aim to achieve a secondary goal of the reliability metric workstream to achieve greater transparency and accountability with customers and stakeholders about performance
- To achieve credibility, this metric needs to resonate with LU customers' perceptions and experiences of reliability
- Research is required to explore LU customer expectations and perceptions of reliability in detail, and the potential role of a reliability metric



Research approach

- An iterative research process has been designed to develop and explore customer response to metrics
- This document presents findings of the first stage of a two stage research and development process

STAGE 2 (to be completed) STAGE 1 Research (completed) Six focus groups amongst a debrief and fresh set of Tube users Six focus groups amongst Tube development of Research metrics for debrief To test response to metric testing in communication ideas and To explore customer experience, second stage provide guidance and understanding and response to feedback for refinement a range of metrics for future development

- Stage 1 research conducted by 2CV in August 2011
- Stage 2 expected to be conducted in September 2011 objectives and stimulus plan to be agreed by end w/c 5th September. Final stimulus agreed by end of w/c 12th September

Research objectives

- Explore customer experience and understanding of reliability on the Tube – identifying customer language, experiences, emotions, beliefs, wants, expectations, etc
- Identify the type of performance reporting that customers want and what will be credible
- Understand customer response to a range of current and potential metrics identified by LU to help inform the development of a reliability metric

Methodology and sample

6 x 2 hour focus groups

	Life stage	Primary journey type	Location
1	Pre-family	Commuters	Inner London
2	Family	Leisure	Inner London
3	Post-family	Commuters	Inner London
4	Pre-family	Leisure	Outer London
5	Family	Commuters	Outer London
6	Post-family	Leisure	Outer London

- Additional criteria: all ABC1; equal mix of genders; good representation of main line usage; at least half to have been affected by disruptions in the last six months; mix of advocacy towards LU but no rejecters
- Research conducted by 2CV in August 2011

Executive summary: headlines from the research

Headlines from the research

- 1. Customers cannot isolate reliability from their whole journey experience and attitudes to the Tube
- 2. Customers define LU reliability as performance across both Operational (platform wait time, as expected, frequent) and Customer Care (making me feel secure, keeping me informed) domains
 - > Both aspects should be equally considered in improving customer reliability perceptions and experiences
- 3. Customers cannot have absolute certainty in their ability to predict or forecast their day to day experiences and have therefore developed personal and ever-changing strategies to try to stay in control. This is more complex and individual than a time based buffer zone
- 4. Reliability metrics must be developed to be complementary to the customer mind-set, be relatable to personal experiences and help customers hone their individual habits there is a need to be wary of any potential discord or rejection from customers
 - > Whilst opportunities exist for development, if there is any dissonance between a LU metric and customer mind-set, it would potentially do more harm than good
- 5. Customers do call for greater transparency and communication from LU, however, this information must relate to **their** current or future journeys

Defining reliability: customer attitudes and experiences

The core qualities of 'reliability' are predictability, consistency and dependability

Customers were asked to define 'reliability' as a quality (beyond travel, transport and LU):

A reliable Somecretionating is departable and trustmorthy that leads to a precious experience.

Reliable is --
Something that's depositionly trust wathy havest reassuring and something you can count on.

Sandthing or someone your own depend on without

Customers consider their Tube experiences and expectations holistically

When considering LU 'reliability' customers automatically articulate all aspects of the service experience and their travel psyche / habits:

LU Service Factors

Customer factors

- Frequency of trains
- Comfort levels aircon / capacity / getting a seat
- Value for money
- System capacity / overcrowding / ability to get a seat / busyness
- Other system users
- System disruptions works, day to day disruptions
- Staff visibility / communication
- Information services and communication
- Inner vs Outer London regularity of service and alternative modes available

- Personal habits / travel identity and heuristics
 - Early vs late person
 - Planner vs non-planner
- Personal familiarity and confidence with the network
- Typical journeys stations, lines
- Specific journeys and the overarching experience:
 - The journey route / destination
 - Journey purpose
- Travelling alone or with others (particularly with small children)

Reliability is deeply interwoven with many other facets of experience and personal habits

Customers can't isolate reliability from their whole journey experience and attitudes to the Tube



It is challenging to raise the issue of 'reliability' for LU without rousing other emotive topics

For customers LU 'reliability' performance is multifaceted and considered across two equally important domains



Operational Reliability

- 'On time' (defined by customers as wait time on platform)
- Regular service
- As <u>frequent</u> as possible to reduce overcrowding / improve experience
- Predictable / as expected
- Consistent performance the same experience everyday



Providing good service - on time, comfortable, value for money

Customer Care Reliability



- Clear and transparent communication
- Real time updates
- Helping customers make good travel decisions – eg live updates, weekend closures, journey planning, time estimates
- Providing reassurance when customers experience disruptions / closures
- Being 'trusted' to deliver emotional reassurance



Keeping customers **informed** and **helping** when things go wrong

These elements work in tandem and are interrelated. Whilst customers feel there are sometimes small improvements in both domains, there is still much 'unreliability' in the service

Customer Care and Operational reliability should compensate for each other's shortcomings, but often when performance is disrupted on one measure, the other does not respond



Operational Reliability

"If it's all running normally, you don't really need any information or care – your train turns up on time, you get a seat, you read the paper and you're done – it doesn't get better than that"

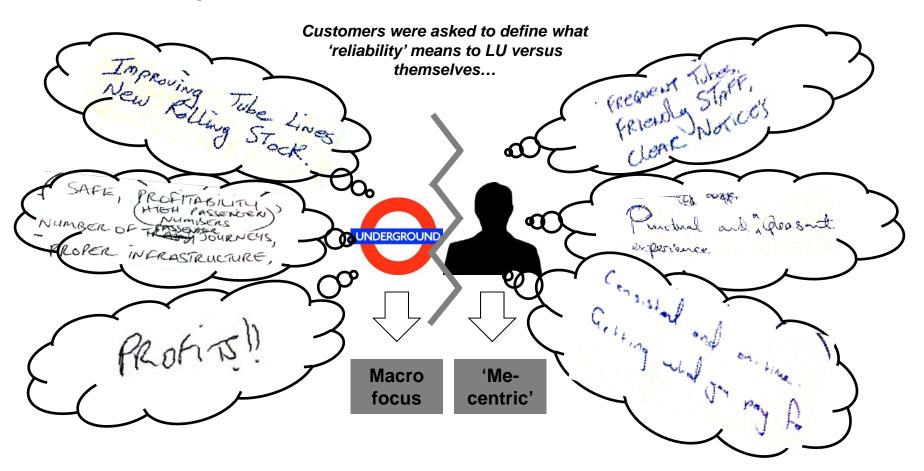
Customer Care Reliability

"I got stuck in a tunnel for two and a half hours, when we got out, they hadn't arranged any water – they had two hours to think about it – why didn't they do anything?"

"What does 'severe delays' mean anyway? They tell you there's problems when it's running smoothly and say it's all running normally when you're stuck in a tunnel for 5 minutes between stops – how are you meant to know what is happening?"

Currently, customers cannot place absolute confidence in LU's performance on either factor – leading to poor perceptions of 'reliability'

Customers believe LU's approach to reliability does not match their needs and expectations



Managing reliability: customer expectations and strategies

Customers base their perceptions of LU reliability on how easy it is to predict day to day experiences

0% reliability < → 100% reliability Perceived reliability: 60-90% "It's not like the NY metro that "0% would be never "I think it runs as expected runs 24 hours, or turning up on time, about 60% of the time -"I think it's quite the Beijing one trains constantly there's occasionally a few good overall, which is broken, irregular minutes late and the trains every 2-3 completely new service without a weekend closures can minutes in but it will never timetable " really mess your plans" peak, 5 minutes be those things" in off peak and they help you "I get the train most mornings, I get to work on when there's "It's the most time, it's a bit crowded, but it is peak hour, problems - I'd expensive in the occasionally it completely stops - but it is a big say about 90%" world - they keep network I quess" raising the fares

Based on a host of personal perceptions and experiences. Customers are resigned to a degree of unreliability, believing the ideal (95-100% reliability) is unachievable within the current system

and for what?"

Day to day, customers are monitoring 'reliability' automatically along their journeys











Pre-journey

- Have I left on time
- What is the purpose of my journey
- What other commitments do I have (eg childcare)
- What is my mood
- What is happening in London

Journey to the Tube:

- Are there any issues
- How busy are the streets
- Do I get distracted / waylayed

In the station entrance:

- How many people are there
- What do the notices say
- What is the 'vibe'

On the platform:

- How long to the next train (dot matrix)
- How many people are there
- How busy are the trains
- Are there any announcements (system and line)

On the train:

- Does it seem to be running normally?
- How long are any delays
- Are they telling me what happens?

Customers are constantly (and automatically) aware of how 'normal' the service is – prepared to alter their travel mood and plans if required

<u>Personal</u> experiences sit at the heart of customers' knowledge of how the system works

Customers have their own robust data set of reliability experiences and perceptions

This data set is subconsciously recalibrated during each moment of travel

This is influenced by many personal factors:

- Commuter vs Leisure
- How tolerant they are of the Tube environment
- How they pass the time
- Their pressures at home (childcare etc)
- Their mindset at the moment of travel
- Their destination
- Their travel habits
- The luck of the draw / good timing
- Their travel identity I am a Tube user vs I am a car driver who uses the Tube to get to work
- Customer reliability an early vs late person, planner vs non-planner

Therefore, customers' approaches to coping with LU's unreliability are more complex than time based buffer zones

- Customers build in personal strategies that are not just about how the system runs but account for and are driven by travel habits and personal coping mechanisms
- 'Buffer Zones' account for both quantity and quality of time.

"I leave an hour extra at weekend to entirely re-route if need be" "I don't use it at weekends because it is so often closed"

"I don't build in any time because I am not worried about being late to work..."

"I know I'll only get the third train in the morning" "If it seems really busy or backed up, I can walk a station further and have a better experience"

"I try to avoid the peak hours"

"I leave 10 minutes earlier to make my journey less crowded and more relaxing" "Sometimes when it is really bad I get on the train going in the opposite direction just so I can get on at all" "I know that on a good day when nothing goes wrong I can make it in 40 minutes. I am not that bothered about being late, I just don't want to spend more time than I absolutely have to"

Strategies are deep seated, personal and based on a wealth of first hand experiences. Personal confidence in these strategies help customers feel more in control

Reliability in context: perceptions of LU and media backdrop

Customers feel passionate about many issues affecting them as Tube users

- When discussing reliability and LU it is clear that customers hold strong and deep seated opinions:
 - All customers simultaneously hold both positive and negative opinions of the Tube
 - Both quality and quantity of time are important
 - They are passionate about the network and have very clear expectations of the service they want and expect.
 And they want specific issues addressed by LU (air con, wait time, cost of service)
 - They feel that they are 'Stakeholders' and should be treated as such, but often they end up feeling like victims at the mercy of an omniscient power

"It's a monopoly – they don't have to compete – they put up prices but it doesn't improve at the same rate, we don't get any more for our money"

"I think they need to look at the state of the system in its entirety – what customers experience, and what they need to get better service"

> "It's our Tube – we deserve better than we get"

On balance, customers are experiencing tangible improvements on the Tube



Perceptions of LU are improving over time

"I have seen information about Tube upgrades so they are working on it but it will take some time" "The new Victoria trains are much nicer - I always hope for a new one when I take the Tube"

OUR UPGRADE PLAN

"They're making more announcements than before, it helps you make decisions"

"There are more trains on the Jubilee line than there used to be – you don't have to wait as long in the morning"

Improved perceptions driven by both Operational and Customer Care factors

Perceived tangible improvements are being positively supported by the TUP communications



Conscious mentions

- Customers referencing the posters and updates – LU providing better information and keeping customers informed
- Customers mention posters explaining escalator improvements – the number of times around the world / to the moon
- Low conscious mentions references from the TUP campaign are entering customers lexicon:
 - 'Oldest in the world'
 - 'More passengers than ever before'
 - 'They have a plan'
 - 'There's more information about closures'
 - 'Improvements'
 - 'Bringing the Tube into 21st century'

But the media environment is still weighted towards the negative

Selection of media, Wednesday 10th August – Thursday 19th August

Crush hour on railways gets worse

City Half Carnespondent

Fare rises are heartless



Taxpayers foot 'scandalous' bill for full-time TfL union activists ATIL spokesman said: "The member of trade union representatives at TIL described the cost as "scambalous". It also emerged that the Met employs Peter Dominiczak

Towards a reliability metric: reviewing response to the metrics

A note on the research approach

To inform development of a reliability metric for LU, the research took a three-pronged approach:

- 1. First, customers worked to come up with their own metrics / demands for LU reliability:
 - What would they like to see improved
 - How would they track reliability over time
- Secondly, customers were introduced a range of current and potential metrics for review:
 - Evaluated in terms of credibility, relevance, interest levels and expectations
- 3. Finally, we explored customer preference in the logistics of delivery:
 - How, what, how often, what level of personalisation

List of current and potential metrics

The research evaluated a number of different metrics in order to identify the most appealing approach

Current Metrics

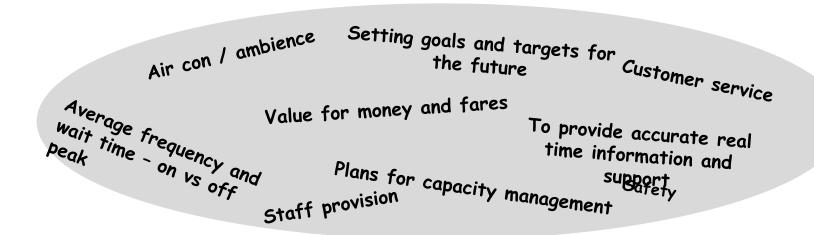
- 99% of trains in service
- 48 out of 260 stations closed more than 15 minutes
- 44 minutes total journey time on average
- 6 minutes excess journey time on average
- 79 out of 100 overall score on customer satisfaction survey (January to March 2011)
- 1,065 million total number of journeys carried each year
- 98% km operated
- 97% of lifts in service
- 96% of escalators in service

Potential / additional metrics

- Platform wait time averages across the network,
- % of journeys where passengers wait on a platform for 5 minutes more than they should
- % of journeys taking 5/10/15/30 minutes more than
- Number of delays on train that are more than 5/10/15/30 minutes
- Number of trains stuck between stations for more than 30 minutes
- Number of detrainments

Customers do not separate 'reliability' metrics from other performance objectives when developing their own list

- Customers were asked to list key reliability performance metrics for LU, but could not, and did not want to 'isolate' reliability from other performance factors
- LU reliability performance reporting should include...



Response to LU developed metrics often evoked an emotional response from customers

Whilst customers like the idea of communication and information from LU and see LU as a trusted source, many of the metrics felt very disconnected from their personal needs and expectations:

"Where do I start? This is just ridiculous – have they ever even taken the Tube during a rush hour?" "I want to see improvements!"

"It's ok, but how are they going to improve the experience for the customers?"

"I might not go to the website but it's good to know they have the information available"

"This is just typical... why can't they just change things instead of telling us about it?" "I don't want more statistics – I want a better commute in to work. I pay enough for that at least"

"I want to know about what they're doing in the future, what we will be getting"

"I don't have a choice. I have to use the Tube, so why are they trying to make me feel even worse?"

"They should think about the comfort of the journey, not just the numbers"

"Some of this is good, but a lot of it is a bit meaningless, it's hard to know what to do"

Customers call for transparency and greater communication with LU, but are quick to push back when information is not suited to their needs

Information is not being projected onto a blank slate but a complex web of personal experiences, expectations and cultural context

Cultural Mythology:

A web of deeply embedded cultural and media discourses (it's a British sport to criticise the Tube, a love / hate relationship, other metros are better, the media)

Customer Ideal:

The ideal performance and service expectations that customers may recognise are unrealistic, but are relevant nonetheless – the Tube should be almost entirely perfect, regular, comfortable and never let me down

Customerknowledge:

How the system is expected to run under reasonable conditions—the unwritten customer Tube rules (wait times peak vs offpeak, capacity, likely disruptions

Customer Experience:

A combination of most recent journeys, totality of experience, network performance, any big events or disruptions, personal traumas, how lucky / unlucky they are

Any information that doesn't complement this picture can evoke a tirade of criticism and counter-evidence

This web of information is important to customers as it helps them stay in control of their experiences

Customers want to stay 'in control' of the expected experience at all times

Customers have their own robust data set of reliability experiences, perceptions and personal travel strategies

This data set is subconsciously recalibrated during each moment of travel

External data presented to
customers can trigger cognitive
dissonance as customers
compare external 'system' data
with their complex personal
experiences of reliability

Information must be in line with customer expectations and desires to be accepted into the web

Customers identified five groups of metrics



Operational

- 97% of lifts in service
- 96% of escalators in service
- 99% of trains in service
- 98% km operated



Negative

- Number of delays on train that are more than 5/10/15/30 minutes
- Number of detrainments
- Number of trains stuck between stations for more than 30 minutes
- 48 out of 260 stations closed more than 15 minutes



Retrospective

- 6 minutes excess journey time on average
- 44 minutes total journey time on average
- 79 out of 100 overall score on customer satisfaction survey
- % of journeys taking 5/10/15/30 minutes more than they should
- % of journeys where passengers wait on a platform for 5 minutes more than they should



Big Facts

- Top causes of delays*
- 1,065 million total number of journeys carried each year



Future Focused

- Train frequency*
- Platform wait time averages across the network, line, station, at peak and off peak
- # of signal failures*

* New metrics suggested by customers in this research

Each group triggered a different response from customers



Operational

Rejected because information relies on system knowledge for interpretation and doesn't relate to customer experience

"How many trains are there anyway? And how does it impact my experience?"



Negative

Rejected because they make customers feel insecure by bringing to the fore the most negative potential experiences

"How many times does this happen anyway? I think I'd stop using the Tube"



Retrospective

Journey data
presented in a way
that discords with
customers' personal
'subjective'
experiences or
doesn't speak in
customer language

"What do they mean 'more than they should?' I don't want to wait more than two minutes"



Big Facts

These help customers understand the system / how it works and adds to their overall system knowledge

"If it carries that many people a year, I think – wow – that's impressive and it must work"



Future Focused

Welcomed as information that can be translated into personal strategies or used to improve future journies

"If I know how often trains come during off-peak, and when an increase in frequency occurs, I can work around that"

Operational metrics measure performance at a system level, the opposite to customers' 'me-centric' viewpoint



Operational

- 97% of lifts in service
- 96% of escalators in service
- 99% of trains in service
- 98% km operated

"Does that mean that they didn't cover 2% of the network?"

"This group are all absolutely meaningless to me"

"What if you only have one lift at your station and it is part of the 3% - then 100% of your lifts are out of service"

> "But there are hardly any lifts in the first place! They should be talking about making it accessible"

"Yeah, great if you don't have children and aren't travelling with buggies"

"There always appears to be lots of unused trains as you catch the train out of Paddington – that looks like a lot more than 1%?"

- Customers don't have the contextual facts to interpret the data
 - eg How many lifts are there? How many trains are there? What does KM operated mean?
- They also are not interested in learning about the technical 'operational' qualities of the system as they cannot see how it could ever translate to their journeys or forecast reliability
- Furthermore, customers believe this operational data easily conceals bad personal experiences and is not credible on a personal level

This system data does not relate to customers' definition or experience of reliability

Negative facts make customers feel unhappy

VERT, VERT, VERT DISAPPOINTED

Negative

- Number of detrainments
- Number of trains stuck between stations for more than 30 minutes
- Number of delays on train that are more than 5/10/15/30 minutes
- 48 out of 260 stations closed more than 15 minutes

"If I knew this...I would never get on a train again"

"What are they doing about it?"

"What is a detrainment – it doesn't sound like a good thing"

"Why are they stuck?!"

"20% is awful, why would they tell you that?"

"I understand how they would be useful internally – they should know and do something about these things – but it's not really for customers to know"

- These facts bring poor performance to forefront of customers' minds and are extrapolated to be severe failures and 'catastrophes' of performance
- These 'facts' are also reminiscent of negative media discourses
- Customers' projected tolerance for system 'failures' is lower than the reality – they don't want to have to imagine the experience, and would prefer to bring in coping strategies in the moment
- Furthermore, these are viewed as more occasional disruptions on the system which customers don't typically plan for
- They make customers feel out of 'control' (upsetting) and can evoke feelings of victimisation

Customers expect LU to be vigilantly managing these issues. But these should not be directly presented to customers as they have strong potential to damage LU's reputation

Retrospective performance, customers struggle to reconcile personal historical data with system averages



Retrospective

- 6 minutes excess journey time on average
- % of journeys taking 5/10/15/30 minutes more than they should
- % of journeys where passengers wait on a platform for 5 minutes more than they should
- 44 minutes total journey time on average
- 79 out of 100 overall score on customer satisfaction survey

"What do they mean excess journey time? That doesn't make any sense to me"

"Who defines how long a journey 'should' take? Surely it should take the same amount every time? I'd rather know how long my journey takes on average"

"What time of the day does this apply to? Five minutes in peak is awful, in off-peak that's ok"

"44 minutes on the Tube sounds like a long time, and I don't care about 'average' journeys anyway – each journey is different"

"It's an ok fact but who did they ask? They didn't ask me?!"

- Customers automatically consider statistics against their own databank and find it easy to invalidate the statistic based on their personal experience
- Customers don't want to be compared to the average, or told what to feel – they use the system regularly enough to have their own data
- System averages are hard to relate to individual experience
 - Average journey time is unique to each user
 - Personal thresholds for wait time, excess journey time etc vary
 - Expectations also vary by peak vs off-peak
 Exacerbated by a disconnect with customer language, they don't think in terms of 'excess journey time' or 'waiting longer than they should'

Terminology and language also fail to relate. But could these be re-articulated?

Big Facts can be interesting and engaging for customers, helping put reliability performance into context



"What's holding everything up anyway? What kind of problems are they dealing with?"

- Top causes of delays (new)
- 1,065 million total number of journeys carried each year

"It would be nice to know what is causing all the problems. You hear about delays all the time but what is the reason?"

"Maybe a comparison to last year would be good to know and perhaps a forecast for next year too"

"It wouldn't make a difference to my journey but a billion journeys is pretty impressive" "They could tell you about the different issues on different lines and stations – what causes the issues at Earls Court"?

- This information helps customers increase their knowledge of the system:
 - Helps put individual experience into context
 - Enabling them to feel more in control when things go wrong – making system failures more of a known quantity
- Makes the system feel large and well managed:
 - Have the potential to evoke pride and empathy, which in turn builds tolerance
 - Can make it feel more 'reliable' in terms of reach and system coverage if not day to day performance

These could also be used as a backdrop to create interest and contextualise metrics

Future Focused metrics can quickly be translated to help customers be more in control



Future Focused

- Train frequency
 (esp. by peak / off
 peak, lines/stations)
 (new)
- Platform wait time averages across the network, line, station, at peak and off peak
- # of signal failures
 (new)

"This is helpful – I expect a train to arrive in a reasonable time"

"It could help you know what to expect"

"They could tell you the maximum so you know the worst case scenario – but I wouldn't want it to be too long"

"I would want it to be positive though, in the morning if I turn up and it says it's 3 minutes to wait, my heart sinks!"

"We hear about signal failures all of the time but is this getting better?"

- These metrics could present reliability in a way that is relatable to actual experience / expected performance
 - Tangible, translatable to day to day experience
 - Fulfil both Operational (turning up on time, running as expected) and Customer Care needs (caring, about the customer, real time information, transparency)
- Can be future focused and show a plan for improvements
- But need to be in line with customer knowledge or 'Ideal' expectations (or they become Negative)
- And need to be developed to a level that can help customers translate to their personal expected experience – particularly peak vs off-peak

Can help customers make the leap from reliability to expected experiences. But customers want positive realism, so there is a need to ensure these are both credible and future facing

Next steps: developing a reliability metric

Big Facts and Future Focused metrics have most potential



Operational

Rejected because information relies on system knowledge for interpretation and doesn't relate to customer experience



Negative

Rejected because they make customers feel insecure by bringing to the fore the most negative potential experiences



Retrospective

Journey data
presented in a way
that discords with
customers' personal
'subjective'
experiences or doesn't
speak in customer
language



Big Facts

These help customers understand the system / how it works and adds to their overall system knowledge



Future Focused

Welcomed as information that can be translated into personal strategies and used to improve future journeys

Easily discredited as customers struggle to relate these to personal experience – making the system feel more disconnected from the customers' needs and wants than it does already

Not working currently due to system averages and language. But Is there potential to reframe these messages? Opportunity for further development as they have the potential to provide information that is both interesting and relevant to customers – adding to their knowledge and helping them stay in control

Five guidelines for development – A reliability metric must...

- 1. Chime with customers' personal data set: feel relatable to customers first hand and personal experiences and relate to future changes that will be felt by them personally
- 2. Use customer language and frameworks: avoid aggregated, average or operational focussed metrics that do not speak directly to the customer (eg overall rather than peak vs off-peak)
- 3. Add new and relevant knowledge: build on existing customer knowledge to help them improve and understand their journey experiences where possible
- **4. Tell a positive and action based story:** customers want to know what is happening in both the short term and mid-term future rather than the past
- 5. Leave no room for confusion or rejection: statistics, averages and system data can be hard for customers to understand it must be easy and automatic

If metrics do not meet customer needs and expectations, the resulting impact is not neutrality or remaining at the status quo, but potentially damaging to LU's reputation

Next steps: taking the customer perspective forward