Intercity Express Train Driver Controlled Operation: Train Safety Check Experimental Assessment

L. Jones, M. Freer & D. Simmons

CCD Design and Ergonomics Ltd, 68-80 Hanbury Street, London, E1 5JL

ABSTRACT

Under Driver Controlled Operation (DCO) passenger safety when boarding a train is the responsibility of the driver alone. Body-side Cameras provide In-Cab CCTV imagery allowing the driver to check all carriage doors are unobstructed and perform the train safety check of the platform train interface before leaving the station.

Current Rail Industry Standards indicate that drivers can reliably view up to 12 CCTV images, one per carriage, to make the train safety check. However, the new Intercity Express Trains (IET) are longer than other UK rolling stock and require x2 opposing cameras per carriage to ensure no blind spots occur. IET Drivers will therefore need to review twice as many images as on other DCO systems; potentially up to 24 images for one 12-car train. CCD was tasked with investigating whether drivers can reliably detect target incidents with up to 24 images and whether task time for reliable performance will significantly increase.

The assessment was conducted as a comparison study between an existing DCO train and the new IET train. Video footage from both trains was captured at Paddington station, with over 200 actors used to simulate a busy platform and target incident scenarios, such as being trapped in the doors, falling over, etc.

Two Train Operating Companies provided 39 drivers to participate in a simulated desktop experiment, across 4 combinations of CCTV imagery; 10, 12, 20 and 24 images. The drivers were presented with a series of CCTV videos and asked to identify if there was a hazard that would prevent them from leaving the station. Responses were recorded on bespoke software, capturing target detection performance and response time. CCD were commissioned to conduct the study on behalf of Hitachi Rail Europe and presented the findings to the stakeholder approvals board in May 2018, including the ORR and RSSB.

KEYWORDS

CCTV, target detection, driver controlled operation, trains, screen layout

1 Introduction and Background

With the continuing increase of technology and Driver aids to improve passenger safety, the use of Driver Controlled Operation (DCO) in passenger trains has been introduced, which allows the driver to independently perform the final check of the platform, via in-cab CCTV monitors, before moving out of the station. The current Rail Industry Standard RIS-2703-RST [Ref 1] allows drivers to operate DCO trains with up to x12 CCTV camera images (i.e. one image per car of a 12-car train).

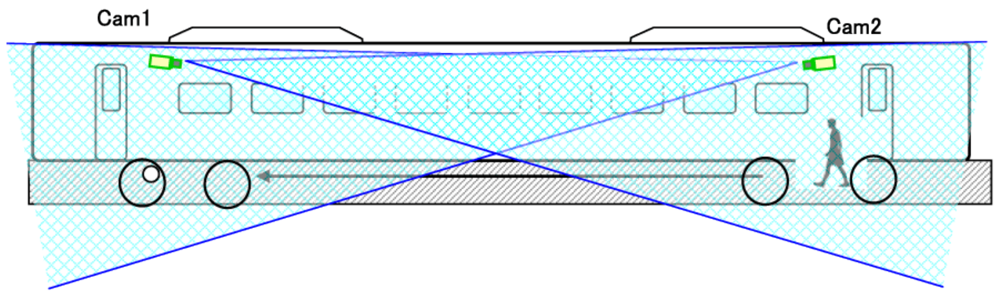


Figure 1 Hitachi IEP train camera set up

However, the Hitachi Rail Group (HRE) Intercity Express Train (IET) carriages are 26m in length and are thus longer than other UK rolling stock. This means that to ensure the Driver does not have any blind spots (especially on curved platforms) each carriage will require two cameras per carriage side. The IET driver will therefore be provided with two images per car, twice as many as on other existing DCO systems, with potentially up to 24 images for one side of a 12[1] car train. The HRE IET train's CCTV system does not therefore comply with the current standard, yet HRE remain responsible for demonstrating to the relevant authorities that the system is fit for purpose.

CCD were asked to carry out a study to examine whether drivers can reliably detect simulated incidents in the platform train interface (PTI) corridor using up to x24 CCTV image pairs, comparing the IEP to an existing DCO train [Ref 1]. The experimental assessment was developed in consultation with the Office of Rail Regulation (ORR) and Rail Safety and Standards Board (RSSB).

This experiment was to expand on previous testing of the IEP train conducted by CCD, carried out solely with static targets (i.e. trap and drag incidents). The inclusion of developing incidents as targets allows an improved evaluation of driver's ability to perform the train safety check.

2 Method

The experiment design was a lab based repeated measures test with driver participants from two Train Operating Companies (TOC) (Virgin Trains East Coast (VTEC) (now London North Eastern Rail (LNER)) and Greater Western Rail (GWR)) being asked to view a series of CCTV images captured from live IET and DCO trains, some of which included Target Scenarios (i.e. incidents / events deemed to constitute an unsafe situation that would prevent dispatch) simulated by actors. The imagery also included varying levels of platform crowding / passenger density and movement. Drivers were required to detect whether a Target Scenario was present in each sequence they viewed. The software control system recorded accuracy of detection and the time taken to do so as the key measures of driver performance. Driver responses were captured as follows:

- Target present and driver detects (true positive)
- Target present but driver fails to detect (false negative)
- No target present but driver detects (false positive)
- No target present and driver correctly gives the all clear (true negative)

[1] Train Operating Companies (TOCs) currently only plan to operate IET train formations of 5, 9 and 10 cars.

In addition, on each occasion where drivers decided not to move the train, they were asked to point out the target they detected or explain their reasons for deciding it was unsafe to move.

The imagery was developed in four sequences to provide the comparison between different train configurations and compliance with RIS-2703-RST. The aim was to provide a comparison between most likely and worst-case train lengths for IET and existing complaint DCO trains. 10 car trains are the most likely IET configuration which the service will run, 12 car configurations are the longest that already meet existing standards.

1. Non-Compliant arrangement
   IET 10-car train with x2 opposing cameras per car = x20 images presented as x10 CCTV image pairs
2. Non-Compliant arrangement
   IET 12-car train with x2 opposing cameras per car = x24 images presented as x12 CCTV image pairs
3. Compliant arrangement
   DCO 10-car train with x1 camera image per car = x10 single images
4. Compliant arrangement
   DCO 12-car train with x1 camera image per car = x12 single images (maximum on compliant DCO trains)

2.1 Incident Scenarios

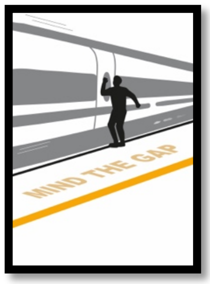
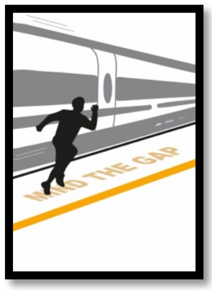
RSSB produced a list of hazards and incidents that can occur within the PTI, the following 10 scenarios were agreed between RSSB, ORR, HRE, GWR, LNER and CCD for the test.

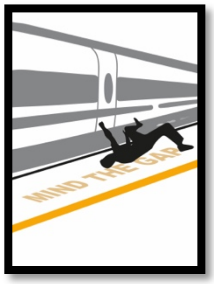
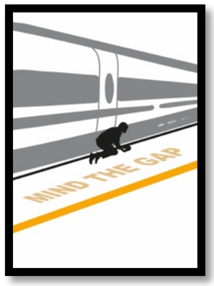
The first four scenarios are hereinafter referred to as **Static Targets**, in that they are present at the start of an image sequence and remain unchanged for the duration.

The remaining six scenarios are types of developing incident, hereinafter referred to as **Emerging Targets**, in that they are introduced sometime after the doors have been closed and during the image sequence (i.e. during or after the driver makes a check of all the doors).

Table 1 Incident Scenarios

	Scenario	Scenario Type	Script
1	**Trapped Aware** (as defined in RIS-2703RST)	Static	Adult passenger caught in the train doors making vigorous efforts (for example, waving arms) to attract attention. 

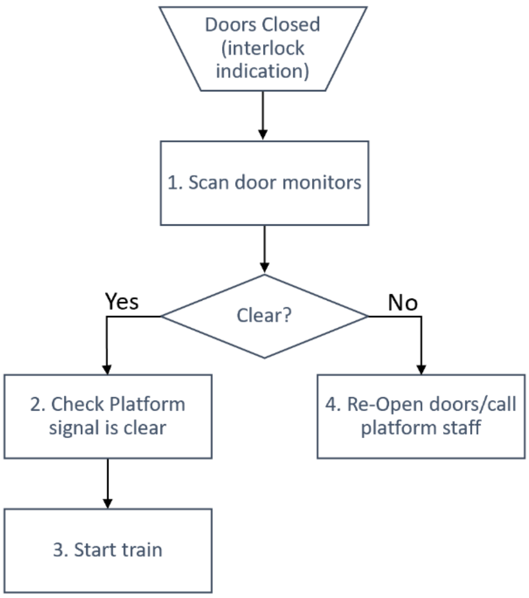
	Scenario	Scenario Type	Script
2	**Trapped Unaware** (as defined in RIS-2703RST)	Static	Adult passenger standing tight against the train but, either as yet unaware they are trapped or otherwise making little effort to attract attention (for example, assuming doors will re-open to release them) 
3	**Pushchair** unattended in train dispatch corridor (as defined in RIS-2703-RST)	Static	Push chair with baby – left on platform close to doorway (abandoned by or stuck in door) 
4	**Small Child** (unaccompanied) within the train dispatch corridor (as defined in RIS-2703-RST)	Static	A 2-year-old child – a model child 825 mm high and 150mm deep placed close to the train doors (i.e. abandoned by, or stuck in door) 
5	**Attempt to Board** person running towards the train and attempting to board (during the train dispatch process	Emerging	Adult runs into view and simulates trying to open doors to get onto the train. 
6	Run Alongside person running alongside the train	Emerging	Adult runs into view (as above) and proceeds to run along PTI (banging on train) 

	Scenario	Scenario Type	Script
7	Collapsed person fallen or lying down within the train dispatch corridor	Emerging	Adult walks into scene and falls and remains lying in the PTI 
8	Retrieving dropped item person leaning towards train having dropped an item and attempting to retrieve it from trackside	Emerging	Adult walks into scene and leans into gap between carriages, then goes to knees / all fours pretending to retrieve something. 
9	Fallen in person has fallen between train and platform edge (action being taken by other passengers to alert driver)	Emerging	Passengers on the platform react, crowd around site of incident and act as if trying to help. It was agreed that there would be no need to simulate the actual fall to the track, merely the reaction of passengers as if a fall had occurred and was seen by them. 
10	**Erratic** passenger behaviour	Emerging	Adult wanders about (as if unwell / drunk), perhaps bouncing off train. 

2.2 DCO Driver Task Simulation

The experiment aimed to simulate the driver experience as closely as possible, however, service provision elements of the task, such as monitoring the platform when the doors are open and /or decisions about when to close the doors were not included.

A practicable experimental method cannot assess the decision / action to close the doors since it would not be possible to extend or contract the lengths of pre-recorded video to match active decisions that drivers might make; an altogether different type of experiment with a live system would need to be conducted in real time to assess this. This experiment seeks to only quantify whether targets can be reliably detected using either the IET or DCO train camera systems for ten and twelve car formations; it is not testing whether DCO is an appropriate / effective method of train dispatch.



For the purpose of testing, the critical portion of the task was simulated, starting from the time the driver receives successful interlocking from the single leaf carriage side doors as illustrated in **Error! Reference source not found.**. This focused the testing to evaluate the final holistic check of the platform before departure.

Figure 2 Experimental DCO Task Breakdown

2.3 Footage capture

CCD conducted the filming of the target scenarios with 200 hired actors to simulate passenger traffic on the platform (Figure 3). The 10 target scenarios detailed in Table **1**, were recorded live, multiple times and at multiple points along both the IET and DCO trains, as well as suitable "clear" footage of passengers waiting on the platform for the purposes of developing the CCTV imagery.

All actors were asked not to cross the yellow line into the PTI unless they were instructed to do so by a member of the CCD team when filming an incident scenario. This was to ensure the incident / target scenarios were not intended as targets, but which might be interpreted as such in testing, did not occur.

Figure 3 Paddington Station Platform 1 hired for filming purposes

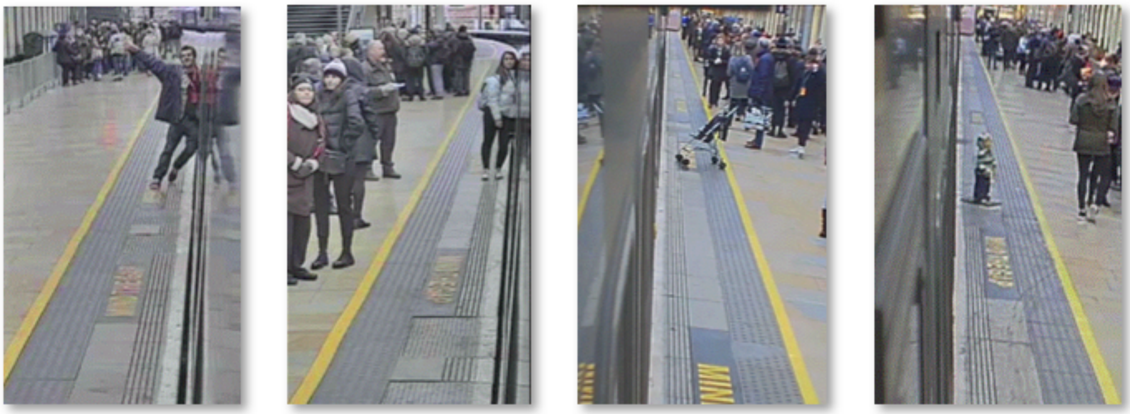


Figure 4 Static Targets – (From left to right) Passenger Trapped Aware, Passenger Trapped Unaware, Unattended Pushchair, Unaccompanied Child

Figure 5 Emerging Targets a) – (From left to right) Passenger attempts to board, Passenger runs alongside, Passenger collapsed

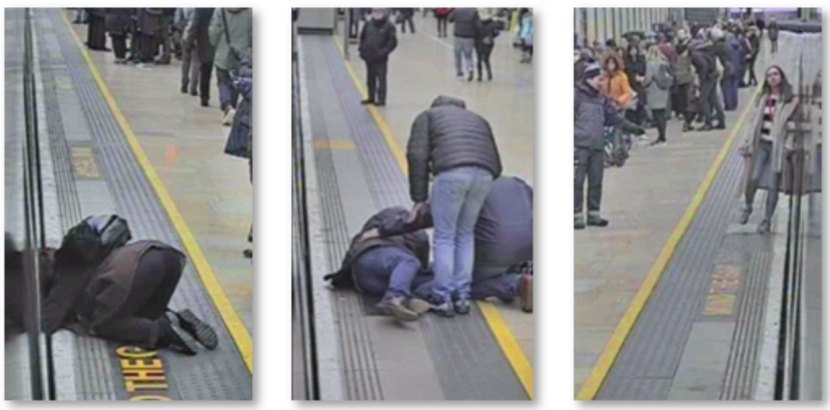


Figure 6 Emerging Targets b) – (From left to right) Passenger retrieving dropped item, Passenger fallen in, Erratic passenger behaviour (e.g. drunk)

Actors were dressed in neutral clothing representative of winter commuter wear, to be no more conspicuous than is likely in everyday life.

Each scenario was filmed several times changing location and using different actors, examples of the footage are shown in Figure 4 to Figure 6. Each CCTV camera exported individual footage enabling flexibility to randomise the target footage across the 12 carriages. These, coupled with "clear" footage in the remaining train CCTV cameras from each film sequence, provided the CCTV imagery to develop the Simulation Test.

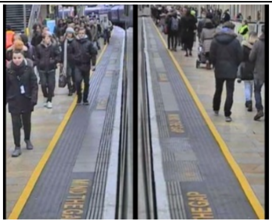
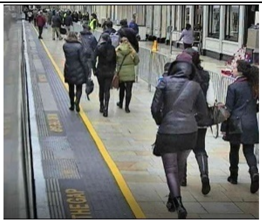
IET Train (double camera)		DCO train (single camera)	
 IET train system		 DCO train system	
Sequence 1A	Sequence 1B	Sequence 2A	Sequence 2B
 10 Car Train (20 images displayed)	 12 Car Train (24 images displayed)	 10 Car Train (10 images displayed)	 12 Car Train (12 images displayed)
40 Video clips played	48 Video clips played	40 Video clips played	48 Video clips played

Figure 7 Structure of CCTV screens for the four different representations of imagery

3.1 Footage Configuration

The CCTV imagery was developed into four sequences, two for each train and length, as illustrated in Figure 7.

Each sequence consisted of either 40 (10 car) or 48 (12 car) sets of video clips for the simulation test; each video clip consisting of 10 or 12 live images, posted to the image slots available on the two monitors.

Target Scenarios were inserted into 25% of the video clips within each of the four sequences. This level of target incidence is more than real life frequency but is entirely consistent with previous DCO studies and is done to give a reasonable prospect of capturing some errors, within a feasible experimental timeframe. Participants were not told how many targets they were about to see nor what sort of frequency to expect. The remaining 75% of video clips in each of the four sequences showed no targets, simulating the safe situation where the driver would expect to dispatch the train.

The frequency and location of different Static and Emerging targets was fully randomised.

The x4 Static targets appeared randomly, since it was not necessary to make sure that all examples appeared in all image positions (as was the case for the previous IET assessment [Ref 2]), merely that enough appear in different screen positions and times in the sequence to induce drivers to perform a sequential check of the doors. It was therefore agreed to use significantly less of these targets (i.e. 8-10% of the video clips).

The Emerging targets, which are the principal focus of this experiment, were the more frequent type, appearing in 15-17% of video clips.

To negate potential order and position effects, the Emerging targets were randomly allocated to the 10 or 12 available image positions on the monitors but were never shown concurrently with any other target scenario and appeared only once in each available image position.

Emerging Target Scenarios started appearing between 8-10 seconds after the sequence started, to give drivers time to get started with the sequential check of the doors. In contrast, Static Target Scenarios, when present, were apparent from the start of the sequence.

Figure 8 Equipment set up

The simulation test was configured so that a participant response could only be made 12 seconds after starting each sequence to ensure that they could not give the all clear before the Emerging Targets appeared.

CCD conducted in-house pilot testing of the system to prove that the software sequencing and control were functioning properly, to verify that imagery was correct and of suitably quality, and to validate the test script and data output.

3.2 Experiment Conditions

The desk-top based exercise was conducted using portable computer-based equipment, consisting of two displays replicating the in-cab display type and size, as well as its position and arrangement, i.e. at the correct visual distance, vertical / horizontal location and orientation (see **Error! Reference source not found.**).

A bespoke software system managed the balancing and randomisation of images as well as recording responses and task time. Drivers inserted responses via a small control pad with three buttons indicating Green if it was safe to move the train, Red if they would not leave the station due to detecting a target, and Blue when prompted to run the next clip.

Each participant viewed the four sequences in a randomised order. If the participant selected the red button, they were told to inform the CCD team of their identified target. If a decision had not been made after 30 seconds, the system automatically assigned a red selection and the participant had to indicate to the lead facilitator why they had not make a decision. The 30 second limit was chosen on basis that this was the 95th percentile task time limit determined in a previous study. Only 5% or less of participant responses could be expected to require more than 30 seconds.

The testing was conducted by two members of the CCD team, one ensuring each participant received the same instructions and conducting the post-test interview capturing participant feedback. The second member was responsible for noting identified targets and data collection.

Each Driver was asked a series of post-test interview questions which consisted of:

1. Did you have a scanning technique?
2. Which of the 4 sequences did you prefer and why?
3. You also saw examples of footage from a 10 car and 12 car trains, how did these compare?
4. Any comments or difficulties?

Each full test lasted for a maximum of 1 hour and 30 minutes, with each of the 4 main sequences lasting approximately 20 minutes. All data collected was anonymous.

3.3 Participants

A total of 39 drivers completed the testing (male n=36; female n=3) covering a range of ages (28-64 years), with a mix of experience of DCO (0 – 28 years). All participants were current drivers, not driver trainers or instructors.

4 Results

4.1 Target Detection Rates

Table 2 shows the overall rates at which participants correctly identified targets from the total number of targets shown for each of the four conditions.

Table 2 Overall rates of correct target detection

Train length	Train Type	
	IET	DCO
10 Car	96.92%	94.87%
12 Car	96.15%	93.59%

Over all conditions the 39 participants failed to detect a total of 80 targets (out of a total of 1716) across both lengths of both trains shown in the experiment. This gives an overall rate of 95.33% for correct detection of targets. The “Retrieving dropped item” target was the one that was most frequently not detected

As the data does not meet the necessary criteria for an Analysis of variance (ANOVA) test, a non-parametric Freidman test can be used to compare the four testing conditions to determine if there is a significant difference between the numbers of undetected targets in each condition (N = 39).

The Friedman test indicated a significant difference (below p = 0.05) in errors between the 10 car IET and 10 car DCO trains, $\chi 2(3)$ = 9.20, p = 0.027. Therefore, a Wilcoxon signed-rank post-hoc test was conducted showing a significant difference in errors between the IEP 10 car and DCO 10 car trains p = 0.021.

The results indicate that performance at target detection is slightly lower for the DCO train for both train lengths, although this is only statistically significant between the two 10 car arrangements (i.e. the performance was worse with the 10 Car DCO compared to the 10 Car IET).

Note that for every positive detection of a target the participants were asked to identify which target they had seen in which image. Post-hoc analysis showed that none of the participants made correct detections for the wrong reason; i.e. made a correct response for a non-existent target in one image while missing an actual target in a different image. There were a small number of incidents where participants correctly identified the right target scenario in the right image but misnamed it (i.e. reported it using the wrong target type – typically running alongside was mixed with running to board).

4.2 Detection Rates by Participant

The number of targets missed by the different participants is illustrated in Figure 9.

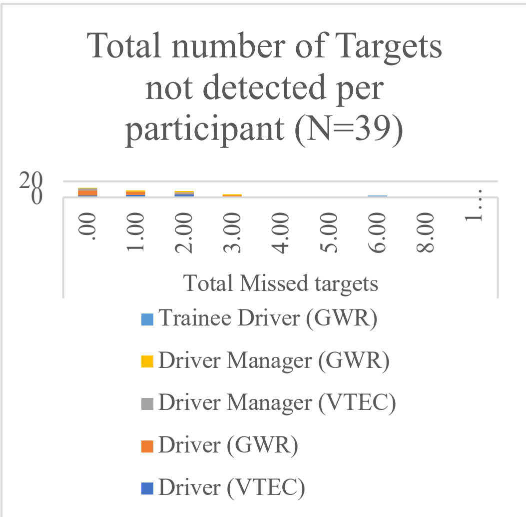


Figure 9 Total number of targets not detected per participant

31% of participants (*n*=12) missed no targets in any of the four conditions. 85% of participants (*n*=33) only failed to detect 3 or less targets.

Over half the failures to detect (43 instances) were recorded by 15% of the participants (*n*=6).

Of the 12 participants who achieved 100% reliability (i.e. made no failures to detect targets in all four conditions), 4 had previous experience of on-board DCO CCTV operation. Of the other 8 participants who achieved 100% reliability, 4 had experience of platform mounted DCO CCTV operation, and 4 had no prior experience of using DCO CCTV at all.

The worst individual performance was 14 targets not detected, which was nearly twice as many as the next worst performing participant (8 targets not detected). On average participants failed to detect 2 targets.

4.3 Emerging Target Detection Rates

The key interest for this experiment was considering how emerging targets influenced driver behaviour.

Table 3 Emerging Target Detection Failures

Target Type	IET 10 Car	DCO 10 Car	IET 12 Car	DCO 12 Car	Total
Attempt to board	2	2	2	3	9
Run alongside	0	3	0	7	10
Collapsed	2	1	1	3	7
Retrieving dropped item	3	5	5	3	16
Fallen in	0	2	1	2	5
Erratic	0	2	3	2	7
Total	7	15	12	20	54
Total No. of Targets Shown	234	234	282	280	1030

Table 3 shows the participants failed to detect a total of 54 Emerging targets out of a total of 1030 shown, equating to an overall rate of 94.76% for correct detection across all conditions.

Table 4 Overall rates of emerging target detection

Train length	Train Type	
	IET	DCO
10 Car	97.01%	93.59%
12 Car	95.74%	92.86%

Table 4 shows the rates of correct detection of Emerging targets for each of the four conditions. Removing the static targets only made a marginal difference; reliability at Emerging target detection shows marginal improvement for the 10 Car IET, and small decreases for the other conditions (the largest change being -1.28% for the 10 Car DCO).

Figure 10 shows 46.15% of participants (*n*=18) detected all Emerging targets and the worst performing individual missed 10 emerging targets.

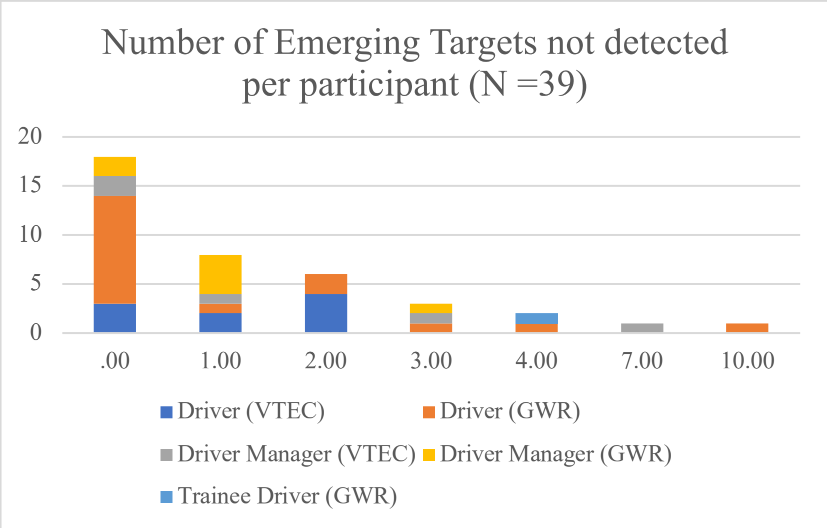


Figure 10 Total number of Emerging targets not detected per participant

4.4 Detection Failures from Timeouts

As explained in Section **Error! Reference source not found.**, providing no response within the 30 second video clip resulted in the system timing out and recording the responses as a "not clear" result.

Table 5 Number of video clip timeouts

Sequence	Total per Session
IET10	74
IET12	75
DCO10	26
DCO12	47
Total	222

Table 5 shows a total of 222 instances of timeout were recorded (3.23% of total responses), mostly for the IET train (both train lengths).

A total of nine of the 222 instances of timeout included a target, three Static and six Emerging, the remaining 213 video clips had no targets in them.

4.5 Results by DCO Experience

As can be seen in Table 6, results show no clear pattern; those with experience of both On-board and On-Platform DCO systems performed best, but those with no experience performed better than those with only On-board or On-Platform DCO experience. The differences were not found to be statistically significant.

Table 6 Target detection rates by DCO Experience

	Targets Detected
On-Board CCTV (*n*=3)	93.94%
On-Platform CCTV (*n*=9)	92.93%
Both (*n*=11)	96.69%
None (*n*=17)	95.02%

4.6 Results by Driving Experience

Drivers experience ranged from 1 month to 40 years.

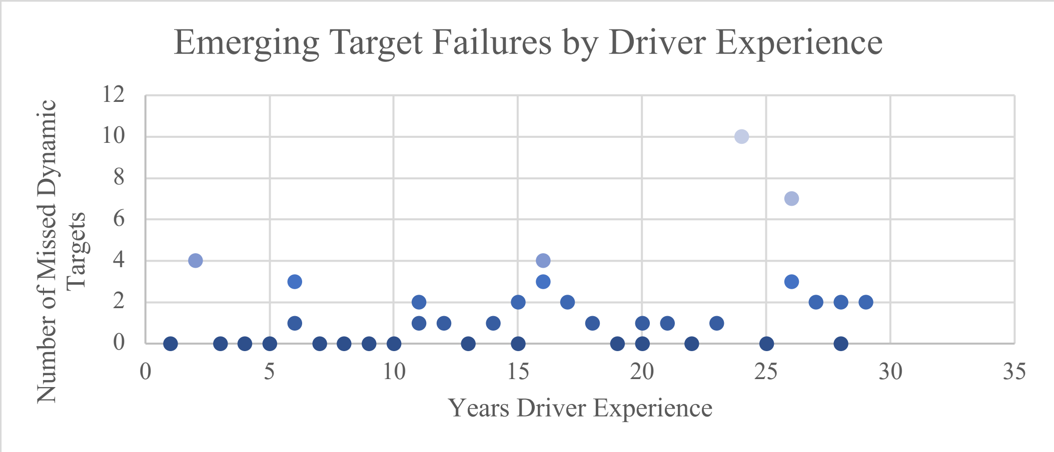


Figure 11 Number Emerging targets not detected compared with driving experience

The graph illustrates an even spread with no statistically significant trend.

4.7 False Target Detection

The percentage rate and total numbers of false positive results (i.e. a target was detected in footage where there was none) are shown in Table 7.

Table 7 False Positives per condition

	Percentage of False Positives	Total Number
IET 10	4.7%	73
DCO 10	3.5%	54
IET 12	10.2%	191
DCO 12	3.9%	73

The results indicate that false positives were more common with the two IET train configurations. A non-parametric Freidman test (N=39) test indicated significant differences between the four train configurations, $\chi2(3) = 51.66$, $p = 0.001$. Post-hoc analysis using the Wilcoxon test showed a significant difference between the IET 10 car and IET 12 car trains (p=0.001), as well as between the IET 12 Car and DCO 12 car trains ($p = 0.001$). The results suggest that drivers tend to make more false target detections with the double image configuration of the IET train in comparison to the DCO train, and comparing the 10 and 12 car IET train, the more images displayed (i.e. the longer the train) the more false detections they are likely to make.

4.8 Response Time

The response times for correctly determining that no targets were present gives the most useful indication of reliable task performance duration for the train safety check task.

Table 8 shows the range of response times in seconds as percentiles for each of the four test conditions. A one-way repeated measures ANOVA test was conducted to evaluate participant's response time in each test comparing (N=39). The results indicate a significant effect on response time between conditions, Wilks' Lambda = .695, F (3,36) =5.267, p = .004. Post-hoc t-tests indicated that there was a significant difference between the 12 car DCO and 12 car IET trains, but not between the other conditions.

The results indicate that participants took marginally longer to correctly determine that the imagery was clear on the 12 car IET train than on the 12 car DCO.

4.9 Post-test Interviews

1. Which of the 4 sequences did you prefer?

Of the 39 participants 74% preferred the DCO single camera view (n=29), 16% (n=6) preference the IET train split camera view and 10% (n=4) had no preference.

2. You also saw examples of footage from a 10 car and 12 car trains, how did these compare?

49% (n=19) of participants had no preference 41% preferred the DCO single camera view (n=16) and 10% (n=4) preference the IET train split camera.

3. Any comments or difficulties?
   - None of the participants reported any difficulties with conducting the experiment or found the task too demanding.
   - Some Drivers commented that the testing as a whole was quite repetitive and demanding but understood the requirement for testing.
   - Drivers with experience in DCO commented that 15 to 30 seconds task timing seemed to be reasonable and quite realistic compared to their own experience.

Table 8 Percentile task times in seconds for correct clear responses (N=39)

Response time Percentiles of all Participants (N=39)							
	5	**10**	**25**	**50**	**75**	**90**	**95**
IET 10	0:00:12	0:00:12	0:00:13	0:00:16	0:00:19	0:00:24	0:00:28
DCO 10	0:00:12	0:00:12	0:00:13	0:00:15	0:00:18	0:00:22	0:00:24
IET 12	0:00:12	0:00:12	0:00:14	0:00:16	0:00:20	0:00:25	0:00:29
DCO12	0:00:12	0:00:12	0:00:13	0:00:15	0:00:18	0:00:22	0:00:25

5 Summary

Overall, participants demonstrated a high level of reliability at detecting targets under all conditions.

The test participants demonstrated similar performance at detecting Emerging targets with both train systems; it can therefore be concluded that there is no significant difference in driver reliability at detection of developing hazards between the IET CCTV configuration and a current compliant DCO CCTV configuration.

It is further concluded that for the particular setups studied, increasing the number of images from 10/12 up to 24 in 12 pairs, does not appear to affect the reliability with which drivers can detect Emerging Target incident scenarios under DCO operation. Therefore, Drivers can be very reliable at making a holistic final check of the PTI to complete the Train Safety Check with both existing DCO systems and with the proposed IET train system; performance rates compare favourably with previous experiments for static targets. Additionally, nearly half of participants were 100% reliable at detecting Emerging targets with the IET train.

Individual differences do not appear to have any systematic impact on performance with the IET train or with the DCO train. Scanning technique was also shown not to have a particularly significant effect, the "Zulu" scanning pattern (left to right row by row) typically used in training being at least as effective as any other.

The IET system was shown to give rise to more false positive responses; this may have been due to driver's being more cautious with an unfamiliar system, but operators should consider and / or monitor the issue during introduction of DCO operation with the IET to determine whether it could lead to a significant number of delays.

Drivers may require up to 30 seconds to achieve a 95% reliability at correctly determining that a train is safe to move. The IET train was found to require longer than the DCO train, typically in the order of about 4 to 5 seconds. Operators should take this into account when determining suitable station dwell times for IET operation.

6 Acknowledgements

CCD and Hitachi would like to thank the GWR and LNER teams for making the experiment possible and for their valuable technical assistance during the trials. Also, CCD would like to thank Hitachi for allowing us to share our results and for being a great client throughout the project.

7 References

[1] RIS-2703-RST - Rail Industry Standard for Driver Only Operated On-train Camera / Monitor Systems