HTO Analysis on Remote Shunting Operations

Final report
within the framework of
SBB Demonstrator Remote Driving

Annika Dreßler,
Niels Brandenburger,
Malte Petersen,
Alexandra König

Oct. 2025

German Aerospace Center (DLR)
Institute of Transportation Systems, Dpt. Rail Transport Systems and Technologies
Rutherfordstraße 2 | 12489 Berlin, Germany

Corresponding author: Dr. Annika Dreßler
annika.dressler@dlr.de

[DLR-TS Rail Human Factors]

Summary

As research and development in the field of ATO progresses, the concept of remote monitoring and control of trains as a fallback level or as a regular solution for defined applications is coming into focus. In order to ensure the efficient and safe operation of such future systems, questions regarding the interaction of human, technology and organisation (HTO) must also be answered during the development phase. As part of an SBB demonstrator on remote train driving, the DLR Institute of Transport Systems Technology was commissioned by SBB to carry out an HTO analysis to determine user requirements and evaluate a prototype workstation for remote locomotive control in the context of shunting. To this end, requirements were researched from the literature and a field test was jointly designed and carried out to test the system.

In the literature research, HTO requirements in the categories of *sensor technology and technical infrastructure*, *workplace design*, *task design*, *individual requirements* and *operational and organisational framework conditions* were excerpted from 36 identified sources relating to remote control in the rail sector and other domains and organised by category. To further specify the requirements with regard to human-technology interaction, a prototype workstation for remote shunting developed by Alstom on behalf of SBB and situated at the Alstom office in Zurich-Oerlikon was tested in the field. From there, an Aem 940 locomotive (Alstom Prima H4) was controlled at the Zurich-Mülligen shunting yard. During the test phase from 5 February to 31 March 2024, a total of 36 test sessions took place with a random sample of 24 SBB train drivers (TD) from the passenger, goods and infrastructure divisions. Tests were carried out during the day and at night to vary the ambient brightness. The participating drivers tested the system by working on 12 different scenarios relevant to shunting operations. These included starting the drive with different positions of the locomotive relative to the relevant dwarf signal, detecting various objects in and at the track, manoeuvring in small spaces and stopping precisely. Aspects of the effectiveness and efficiency of task performance and user satisfaction were recorded as dependent factors. The quantitative analyses were supplemented by comprehensive qualitative surveys, including on the nature of perceived challenges and possible solutions for system design as well as on the perception of possible changes to the job description of the driver.

The majority of the shunting tasks could be carried out effectively with the tested system, with two exceptions: detecting a brake shoe on the track and determining the exact location of the vehicle in the shunting yard. The observed efficiency of task performance was reduced compared to shunting on the locomotive, as both the perceived time required and the subjective effort to achieve the results were higher. This was mainly due to system-related limitations in human perception, primarily at the level of vision, but also at the level of body perception and, to some extent, hearing. Differences in the system with regard to the lever control dynamics that drivers were used to also played a small part in reducing

the efficiency of task execution. On the upside, particularly the extended video display of the Berne area in the system was emphasised as helpful, while the panoramic display of the video stream with wide peripheral areas and the design of the controls, which was largely in line with experience and expectations, were also perceived as positive. User satisfaction with the system design was in the medium range.

In a further developed system for manual remote control in shunting operations, the visible areas should be enlarged compared to the tested system and the recognisability of objects in relevant areas should be increased. The remote operators (RO) require access to other vehicle functions (e.g. lighting) and a system that allows them to easily determine their own exact location in the shunting yard. Feedback on the quality of the data connection should remain part of the system and its functionality can be expanded to provide further support in dealing with latency.

Further research and development is required to answer other design questions in the area of remote train monitoring and control. This concerns, among other things the question of the effects of the size and variability of latency for the purpose of a meaningful definition of critical values, the questions of how system design can support the handling of latency, which areas must be visible and in what quality, and which system design can best combine the coverage of information requirements with the capacities for data transmission, and the question of how the perception of action-relevant parameters such as speed, distances, traction conditions and gradients can be supported at the remote workstation. Further questions concern the effect of longer interaction experience and training with the system on performance as well as the design of the organisation around the workplace, including the description of new job profiles in a highly automated railway system.

Keywords: Automatic Train Operation, Remote Train Operation, Remote Supervision and Control, Rail Human Factors, Human-Technology-Organisation

1 Table of contents

2 List of abbreviations

Abbreviation	Long form
ATO	Automatic Train Operation
DAS	Driving assistance system
DLR	German Aerospace Centre (Deutsches Zentrum für Luft- und Raumfahrt e.V.)
DS	Dwarf signal
HF	Human factors
HTO	Human - technology - organisation
RO	Remote operator; the person controlling a railway vehicle from a distance
RTC	Rail traffic controller
RV	Rail vehicle
SBB	Swiss Federal Railways (Schweizerische Bundesbahnen AG)
TD	Train driver
v	Velocity

3 Background and research questions

3.1 Background and motivation

The HTO analysis decribed here was carried out as part of the SBB demonstrator on remote train operation within the project R2DATO (Digital & Automated up to Autonomous Train Operations) in Flagship Area 2 of Europe's Rail (cf. https://projects.rail-research.europa.eu/eurail-fp2). An ultimate aim of the project is to examine and validate the applicability of Swiss and European standardisation for automation in the rail industry with regard to regulations and technologies. The immediate aim of the project was to demonstrate remote train operation (RTO) in shunting operations and to gain insights into the associated human factors requirements (human-technology organisation, HTO). The relevant HTO analysis was carried out by the Rail Human Factors Group of the DLR Institute of Transportation Systems Technology (DLR-TS) together with SBB.

3.2 Framework and focus of the HTO analysis

A prototype workplace to be tested for remote shunting was developed by Alstom on behalf of SBB (see also section 6.1.4). The system was tested with an Aem 940 locomotive (Alstom Prima H4) at the

Mülligen shunting yard (Zurich). The remote shunting workstation was located at the Alstom office in Oerlikon. During the test phase from 5 February to 31 March, a total of 36 test sessions took place with a random sample of 24 SBB engine drivers from the three business areas of *passengers*, *goods* and *infrastructure*. In order to be able to vary the brightness conditions, the tests were carried out in two time windows: from 9.30 a.m. to 3.30 p.m. and from 7 p.m. to midnight.

In the course of the joint development of the research concept for the HTO analysis in accordance with the human factors test criteria of SBB (Thomas-Friedrich & Grippenkoven, 2021), it was first necessary to make some fundamental decisions regarding the scope of the research question. The options were discussed between SBB and DLR. The decisive criteria for the decision were the development and test status of the remote shunting system to be tested (prototype, initial investigation with HTO focus) and the available resources. The following decisions were made in consultation with the SBB project team:

- **Cross-sectional study.** The focus is on the extent to which it is principally possible to perform manoeuvring tasks remotely with the technology used, for example with regard to the design of the sensors, the transmission of information, the presentation of information at the workstation and the design of the control elements. Basic findings on this can be obtained in a cross-sectional study. Conducting a long-term study, for example on the performance that can be achieved with the system over the duration of a shift or longer periods of time, becomes more practicable once the fundamental questions on system design have been answered.
- **Focus on human-technology interaction.** With the same background, the focus is placed on human-technology interaction and its fundamental design. This corresponds to the system analysis as defined by SBB's HF criteria. Organisational factors can only be taken into account within the possibilities of a cross-sectional study with a prototype that exists as a unique product. Investigating the impact of organisational factors such as operating procedures and task sharing, shift organisation, training or safety culture would appear to make more sense at a later date.
- **Descriptive analysis and evaluation of the new system.** In the human-centred evaluation of a new technology, it is generally advisable to compare the system with the status quo or another reference in order to find out under which aspects it performs better, equally well or worse. In the current case, a sensible reference would be performing shunting tasks on site on the locomotive. Nevertheless, in consultation with SBB and in view of the development status of the prototype, the HTO analysis was designed to initially focus on qualitative research into the new system and the identification of fundamental problems and ideas for solutions, which is also possible through testing without reference measurements. The principles of good design of human-technology interaction, the use of standardised questionnaires and the practical experience of those involved can provide a benchmark for classifying the results.

3.3 Research questions

The HTO analysis carried out focussed on answering the following questions:

1. **How suitable is the prototype system in principle for carrying out the shunting tasks?**
   The question is aimed at determining the usability of the system in terms of effectiveness, efficiency and user satisfaction (EN ISO 9241). Effectiveness is determined by considering the fulfilment or non-fulfilment of selected tasks in the context of shunting. Descriptive references can be determined to assess the time component of efficiency, i.e. the time required to perform the task with the system (e.g. speed travelled), but no quantitative comparison can be made with manual shunting. However, subjective assessments of the test participants regarding the efficiency and precision achieved compared to shunting on site are collected. Another aspect of efficiency is the mental strain caused by completing the task. User satisfaction refers to the extent to which users rate the user experience positively and their expectations regarding use are met. The latter two aspects are determined using standardised survey instruments.
2. **Where do problems still occur and how do they relate to the system design?**

   The aim here is to identify and qualitatively describe the causes of usability problems. As part of the HTO analysis, this information is obtained on the one hand by systematically observing task execution and on the other hand by surveying users after each task has been completed.
3. **How can the system be optimised for high suitability for shunting tasks?**
   This question relates to the formulation of user requirements for the system and the derivation of recommendations for its future design. On the one hand, this information can be inferred from the observed progress of task execution in conjunction with the user evaluations, and on the other hand, it can be formulated by the users themselves. The surveys conducted as part of the field test also included this aspect.

4 Work structure

The work was divided into three packages (WP). The aim of WP1 was the joint development of a concept for the HTO analysis between DLR and SBB. This included defining the overarching questions regarding the suitability of the remote control panel for SBB's shunting tasks and the joint development of a suitable investigation concept for the field test. The procedure and the resulting concept are presented in chapter 6 of this report.

The aim of WP2 was to develop an overview of the requirements from an HTO perspective for a remote shunting workstation and the basic processes in shunting operations. To this end, findings from current research were compiled in a literature review, documented in tabular form and assessed with regard to their significance and transferability to remote shunting operations. The procedure and the results are reported in the following section 5.

WP3 comprised the preparation, implementation and evaluation of the field test and aimed to gain specific insights into the suitability of the remote control workplace for SBB shunting tasks and the HTO requirements for the design of a remote shunting locomotive driver's workplace. The relevant content is presented from chapter 6.2 .

5 Literature review: HTO requirements regarding remote control workplaces

In order to obtain an overview of human factors requirements for a remote train control workplace, a literature search was carried out as a first step. The sources used were English and German-language scientific journals, technical journals, books, specific human factors guidelines and reports from previous DLR projects on the topic of remote train control. The online research was carried out using the scientific research portals Google Scolar and Scopus. The search term *remote train operation* and corresponding German and English synonyms (e.g. remote safety and control, remote supervision, Zugfernsteuerung und Zugfernüberwachung) were combined with terms relating to user requirements (e.g. *user requirements, user-centred, design*, etc.). As the topic of remote control in the railway sector has been relatively little researched in comparison to other domains, sources on the topic of remote control in other domains from which findings could potentially be transferred were also included in the selection, e.g. aviation, robotics and road transport.

The works found were recorded in a tabular overview that summarises the aim of the study, the methodology and the results. The findings from each source were categorised according to 5 thematic pillars. These included the areas of *sensor technology and technical infrastructure*, *workplace design*, *task design*, *individual requirements* and *operational and organisational framework conditions*, whereby a given source could also be assigned to more than one category. To categorise the HTO requirements in the five categories more precisely, they were subdivided again into subcategories. Table 1 shows an overview of the content categories and subcategories.

Table 1. Overview of the subject areas and explanation of the subcategories

Subject area	Subcategory	Explanation
Sensors and technical infrastructure	Transmission and latency	Requirements for stable data transmission and guarantee of sufficiently high image quality
	Safety	Technical requirements for ensuring safety in RTO operation
	Standardisation	Formulation of standards for large-scale use
	Automation	Requirements for autonomy in order to achieve the required GoA level

Subject area	Subcategory	Explanation
	Human-technology interaction	Technical requirements for optimal human-technology interaction
Workplace design	Control elements	Requirements for the design of the control elements
	Information	Information requirements of the train operator
	Communication	Required communication channels of the train operator
	Safety	Requirements for the workplace to ensure safety in RTO operation
	Assistance	Requirements for support of the train operator by the system
Task design	Role definition	Clarification of roles, activities and areas of responsibility
	Tasks	Task definition of the train operator in RTO operation
	Takeover / handover	Requirements for ensuring a safe handover/takeover in RTO operation
Individual requirements	Workload	Requirements for achieving the optimum workload for the train operator
	Capabilities	Required capabilities of the train operator (e.g. physical requirements)
	Skills	Required skills of the train operator (skills, knowledge)
	Cooperation	Requirements for cooperation between different players in the system
Operational and organisational framework conditions	Further training	Requirements for further training for people already working in traditional railway operations
	Realisation concepts	Creation of implementation concepts for switching to RTO operation
	Stakeholders	Integration and consideration of the views of various stakeholders

A total of 36 relevant sources were identified. The categories and subcategories that were used to sort the content are listed in Table 1. Table 2 shows how much information (number of statements) on HTO requirements could be extracted from the sources in each category and subcategory.

Table 2. Number of HTO requirements extracted per category and subcategory.

Subject area	Total number	Subcategory	Number per subcategory*
Sensors and technical infrastructure	23	Transmission and latency	6
		Security	9
		Standardisation	4
		Automation	7
		Human-technology interaction	10
Workplace design	13	Control elements	7
		Information on	8
		Communication	1
		Security	4
		Assistance	4
Task design	5	Role definition	1
		Areas of responsibility	4
		Takeover / handover	1
Individual requirements	7	Workload	2
		Skills	2
		Skills	3
		Co-operation	1
Operational and organisational framework conditions	4	Further training	1
		Realisation concepts	2
		Stakeholders	2

To make it easier to work with the results, the identified sources and the excerpted HTO requirements were prepared in a sortable and filterable format in the document *Anhang.MTO-Anforderungen.Literatur.xlsx* (in German language), which is part of this report. Figure 1 shows an excerpt to explain the structure.

Informationen zur Quelle							MTO-Anforderungen				
Quellen-angabe	**Ursprüngliche Anwendungs-domäne**	**Ziel der Arbeit**	**Methode**	**Ergebnis**	**Relevanz** [1 = hoch, 2 = mittel, 3 = geringfügig]	**MTO-Themengebiet** [Sensorik und technische Infrastruktur, Arbeitsplatzgestaltung, Aufgabengestaltung, Individuelle Voraussetzungen, betriebliche und organisationale Rahmenbedingungen]	**Sensorik und technische Infrastruktur** [Übertragung und Latenz, Sicherheit, Standardisierung, Automatisierung, Mensch-Technik-Interaktion]*	**Arbeitsplatzgestaltung** [Bedienelemente, Informationen, Kommunikation, Sicherheit, Assistenz]*	**Aufgabengestaltung** [Rollendefinition, Aufgabenbereiche, Übernahme/-gabe]*	**Indivduelle Voraussetzungen** [Workload, Anforderungen: Fähigkeiten, Anforderungen: Fertigkeiten, Kooperation]*	**Betriebliche und organisationale Rahmenbedingungen** [Weiterbildung, Umsetzungskonzepte, Stakeholder]*
Brandenburger et al. (2017) - Der Train Operator, Situative Fernsteuerung von automatisierten Zügen	Bahn	Zusammenfassungs-/ Übersichtsartikel zum Remote Train Operator	Summary	Anforderungen	1	**Arbeitsplatzgestaltung** **Aufgabengestaltung** **Individuelle Voraussetzungen**		Grundkonzept: eine Betriebszentrale mit mehreren Train-Operator-Arbeitsplätzen, einem FDL-Arbeitsplatz sowie einem Shared-Knowledge-Space an einer Wand, an dem der komplette Betrieb eines Bereichs abgebildet ist **Bedienelemente** - manuelles Layout sollte der üblichen Anordnung der	**Rollendefnition** - Verantwortung und Handlungsausführung sollten nicht getrennt werden, um ein höchstmögliches Maß an Sicherheit zu gewährleisten (z.B. könnten auch Bahnhofseinfahrten oder Durchfahrten zu einer Anfrage an einen TO führen und eine manuelle Zielbremsung bzw.	**Anforderungen: Fähigkeiten** - Integration von Informationen in eine mentale Repräsentation des gesamten Betriebsgeschehens wird zur zentralen Aufgabe	
Brandenburger et al. (2019) - Task-induced fatigue when implementing high grades of railway automation	Bahn	Auswirkungen von GoA2 und 3 auf die Müdigkeit (und Übernahmegeschwindigkeit) während einer längeren Fahrt.	Tfs sollten in einer simulierten Teleoperationsumgebung (entweder GoA2 oder GoA3) auf Störungen reagieren. Die Fahrt dauerte 2h, dabei wurde die Müdigkeit mit der KSS und der Workload mit dem NASA TLX erfasst	Müdigkeitseffekte waren weniger stark bei GoA3 als bei 2; dennoch war die Zunahme der Müdigkeit in beiden Gruppen zu beobachten. Der Workload in GoA3 war zwar immer noch im Unterforderungsbereich aber aufgrund der vielfältigeren Aufgaben sig. höher als bei GoA2; die zunehmende Müdigkeit	1	**Individuelle Voraussetzungen**				**Workload** - Der Workload sollte im optimalen Bereich liegen, um Müdigkeitseffekte zu begrenzen; - GoA2 scheint insgesamt eher negative Effekte zu haben und sollte evtl. übersprungen werden (so weit wie möglich)	

Figure 1. Structure of the literature overview of HTO requirements in the context of remote control.

The information on the left-hand side (headings highlighted in grey) is used to describe the source. The columns list from left to right: (1) the source citation with authors, year and title and (2) the original application domain in which the respective HTO requirements were formulated. This is followed by keywords on (3) the aim of the work, (4) the method and (5) the result, and (6) an assessment of the relevance of the results for the area of remote train control. Column (7) lists the categories for which the source contains statements on HTO requirements. The corresponding content is listed separately by HTO requirement category on the right-hand side (white headings). The subcategories are labelled in bold within the text in the cells.

An examination of the HTO requirements found in the literature shows that even in the case of highly relevant sources, the requirements are mostly formulated at a relatively abstract level in terms of basic necessities with regard to required information, functions or interfaces, or - if they are named - requirements for the design of these elements. To a certain extent, these properties are inherent in the formulation of requirements for workstations for remote train monitoring and control, just as they are for any system at an early stage of development. The more concretely the context of use under consideration and the available technical framework are specified, the more precise a description of the requirements can be. This is a basic idea of an iterative user-centred design process in which the system design and user feedback inform each other over several cycles with increasing concretisation (Goodwin & Cooper, 2011; ISO 9241-210; Norman & Draper, 1986).

With the availability of a prototype, the performance of tests in a real environment and the specification of tasks typical for the field, the HTO analysis carried out in the SBB demonstrator project offers an advanced opportunity to test the suitability of a system in a defined context and to sharpen the formulation of the relevant requirements.

6 HTO analysis of the prototypical remote shunting workplace

6.1 Methods

6.1.1 Selection of test scenarios

In order to meaningfully evaluate the usability of a system, it is important to know the context of use and to represent it accordingly in the user tests (EN ISO 9241). Therefore, the first joint step was to select suitable use scenarios for the user-centred evaluation – tasks and operational situations that could be used to test the suitability of the sensors, information transfer, information display and functions of the prototype workstation in a meaningful way. The most important criterion here is the relevance of the scenarios in real SBB shunting operations. On the one hand, scenarios should be tested that are typically to be expected in shunting operations with the system. On the other hand, scenarios that are associated with risks for safety and operation are also highly relevant. In addition to both aspects, the scenarios must also appear to be realisable in remote shunting operations. These three components span the scope of desirability when selecting the scenarios. In addition, there is the aspect of feasibility in the sense of realisability in testing, which must also be taken into account.

An extensive list of potential application scenarios that SBB had already collected served as the starting point for the selection process. In a joint iterative process, questions about the scenarios were first clarified. Subsequently, a multi-criteria evaluation of all potential scenarios was carried out for the purpose of prioritisation according to relevance and feasibility. The criteria of frequency, criticality and strategic relevance of the respective task or scenario were included in the relevance assessment. The feasibility assessment focussed on the feasibility of the requirements for remote control on the one hand and the feasibility of the scenario in the field test on the other. Based on the resulting assessments and the time options for the tests, the twelve scenarios shown in were selected.

Table 3. Scenarios selected for the field test to perform with the remote shunting workstation.

Scenario	Explanation
Dwarf signal in front of vehicle tip	RO takes control of the stationary vehicle, with the front of the vehicle behind the relevant dwarf signal so that the dwarf signal is visible.
Dwarf signal directly at vehicle tip	RO takes over control of the stationary vehicle, with the front of the vehicle directly at the height of the relevant dwarf signal so that the dwarf signal is not visible.
Start before crossing switch	RO takes control of the stationary vehicle in front of a crossing switch where two tracks merge into one; RO sees a dwarf signal behind the switch, but cannot be sure that this applies to his/her own track. The relevant dwarf signal is behind the virtual operator position.
No special features	Drive according to signals without any particular incidents.

Scenario	Explanation
Stop before person (orange vest) on the track	During the drive, there is a person on the track in front of the vehicle. At the Mülligen shunting yard, the scenario was implemented by placing a high-visibility orange vest on a stake on the track.
Stop before brake shoe	During the drive, there is a brake shoe on the track in front of the vehicle.
Signal unexpectedly reverts to stop	In this unusual case, a route is cancelled during the run and a dwarf signal reverts to *stop*. For this purpose, the corresponding dwarf signal is set to *stop* by the RTC immediately after the vehicle has passed the preceding dwarf signal showing *proceed*. In this way, the RO cannot see the change in the camera image of the preceding dwarf signal from *proceed* to *proceed with caution*, so that RO does not expect the signal indicating to stop.
De-energised section	The vehicle encounters a pantograph-down signal. As an added difficulty, the signal is positioned to the right of the track so that RO has to recognise the arrow on the signal, which assigns the signal to their own track.
Stop before signal	The drive ends at the end of the set route in front of a dwarf signal indicating *stop*.
Stop before red stop signal board on the track	The vehicle encounters a red stop signal board on the track.
Approaching a railway vehicle	At the end of the drive, there is a railway vehicle on the track that RO should approach as for coupling.
Stop in front of buffer stop	At the end of the drive there is a buffer stop in the track, which RO should approach within 1 m.

6.1.2 Selection of the influencing factors to be analysed

The selection of factors to be considered in the study design (independent variables or control variables), which are assumed to have an influence on the performance achieved with the remote shunting system, was also carried out iteratively in consultation with the SBB. Here, too, a collection of interesting factors was first subjected to an evaluation. The relevance of the factor from the HTO perspective was assessed by DLR human factors experts, the feasibility of the respective factor in the context of the planned field test was assessed by SBB railway operations experts. The results of the assessments were discussed in a joint workshop. A number of factors were excluded due to the intended focus of the HTO analysis described above, for example, aspects of task design such as monotony due to repetitive tasks, which are nevertheless highly relevant for the long-term development of remote shunting workstations. For many factors that are also highly relevant, targeted manipulation in the field test appeared impossible or could only be realised with disproportionate effort, such as weather conditions or the targeted variation of transmission latency. Other paradigms, such as experiments in a driving simulator or computer-based tests, are more suitable for investigating such factors in system development.

Finally, two experimental factors were implemented in the field test. The first is the ambient brightness due to the time of day. By carrying out tests during the day and at night, it was possible to investigate one of the most important influences on vision. A second experimental variation concerns the availability of a possible additional aid for orientation in the shunting yard away from the workplace. This is SBB's ILTIS system, which shows a track plan of the station with markings of the pre-set routes. The system is not currently intended for use by train drivers. However, if the reliability of the data could be ensured, a visualisation like this could provide a remote shunter with useful information, for example to determine their own location and the orientation of the vehicle when taking over.

Other factors that could not be specifically varied experimentally were included in the study as control variables. For these variables, the characteristics that occured during the survey were also recorded and can thus be taken into account in the analysis. On the side of the individual, these are the age, professional experience and local knowledge of the participant as well as alertness at the start of the test; on the technical side, the transmission latency and any image distortions that may occur; and on the environmental side, snow, rain, fog and wet conditions as opposed to clear visibility and dry surroundings.

6.1.3 Dependent measures

The dependent measures for answering the research questions (3.3) were collected at two levels of observation: on the one hand with regard to the entire interaction with the system and the experiences made with it, and on the other hand separately for each of the up to five blocks in which different tasks were completed (see 6.1.5). All the measures collected are presented in an overview in Table 4.

Table 4. Dependent measures considered in the HTO analysis and their operationalisation.

Level	Measure	Instrument / Method
Overall	Subjective usability	SUS (System Usability Scale)
Overall	User satisfaction / acceptance	van der Laan Scale
Overall	Subjective sleepiness - comparison of sleepiness before and after the test	KSS (Karolinska Sleepiness Scale)
Overall	Qualitative statements on the general evaluation of the system by users, safety considerations, comparison with on-site driving in terms of efficiency and precision	Items of the structured final interview

Level	Measure	Instrument / Method
Overall	Expected effects on the job profile, perceived potential fields of application for remote shunting	Items of the structured final interview
Per block	Performance	Successful execution of the task, detection distance for obstacles
Per block	Subjective relative task difficulty compared to on-site shunting	Analogue scale (easier – equal - more difficult)
Per block	Qualitative data on subjective relative task difficulty compared to on-site shunting	Item of the block-related follow-up surveys
Per block	Subjective workload	NASA-TLX (Task Load Index)
Per block	Subjective workload with comparison to the subjective optimum	DLR-WAT (Workload Assessment Tool)
Per block	Aspects of situation awareness: arousal, division of attention, information quantity, familiarity with the situation	Items from SART (Situation Awareness Rating Technique)
Per block	Ideas for improving the workplace for the tasks	Item of the block-related follow-up surveys

6.1.4 The prototypical remote shunting system

The field tests were carried out at the Mülligen shunting yard using an Aem 940 locomotive (Alstom Prima H4). The locomotive was equipped with additional sensors and technology for transmission of information so that vehicle and additional sensor data could be transmitted to the remote workplace and inputs from there could be transmitted to the vehicle.

Four cameras were used to make visual information from the environment available. Three were located in the driver's cab and each provided one of the images of the left, centre and right monitor (UHD, 55 inches; see Figure 2) of the panoramic display. The fourth camera was located in the Berne area, with the angle of coverage extending further forwards into the track than in usual Berne area applications. Its image was displayed on a smaller monitor below the centre large monitor. Vehicle displays such as the speedometer, traction and braking force display and the status of the brake system were shown on the display to the right. The speed was shown both as an analogue display and as a digital value with one decimal place. A touch display to the left of the RO featured interactive visualisations of switches and controls such as the main switch, pantograph, sand, signal horn and direction selector switch, parking brake and target speed setting. In the primary reach zone, a control lever for the direct brake was located on the left and a drive/brake lever on the right.

Figure 2. Prototype of a remote locomotive control system tested in the field. The controlled locomotive Aem 940 was located in the Mülligen shunting yard. The RO was located in an office building in Oerlikon.

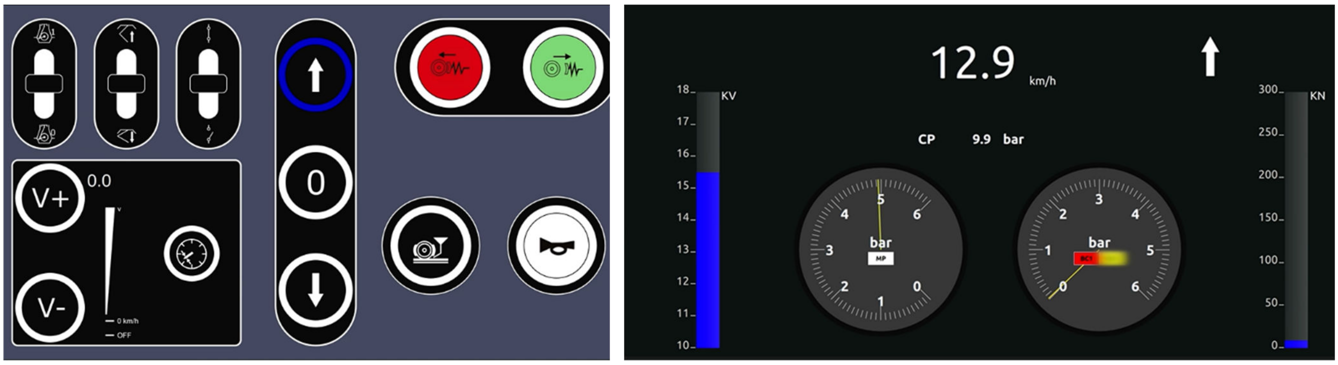

Figure 3. Left: Touch display integrated into the workplace (schematic). Right: Integrated displays.

The three cameras that provided the panoramic image had a resolution of 1920 x 1080 active pixels with a pixel size of 6.5 µm x 6.5 µm. The resolution of the Berne-area camera was 1920 x 1080 active pixels with a pixel size of 2.9 µm x 2.9 µm. Beyond the video displays shown, there was no way for the RO to view the space around the vehicle. Audio transmission from the driver's cab was initially planned, but was not activated for the tests due to disturbing noises. In the driver's cab of the locomotive, a safety driver was ready to intervene at any time if the situation required it. Communication between the RO and the staff on the locomotive could take place at any time via a conference call.

6.1.5 Task blocks, order of presentation, assignment of participants to conditions

The selected scenarios (6.1.1) were summarised into task blocks for the field test in such a way that meaningful action sequences were created. Each of the scenarios was assigned to one of the three driving phases *(1) start*, *(2) progression* or *(3) end*. In each phase, there is a reference scenario with a theoretically lowest level of difficulty that can be used to compare the performance with the performance in the other scenarios of that phase. The reference scenarios are the respective phases of a drive by signals without special incidents: (1) *signal in front of vehicle tip*, (2) *drive by signals without any special features* and (3) *stop before signal*. When setting up the blocks, presumably more difficult scenarios were distributed in such a way that there was no accumulation in any block. In addition to the task blocks for the test, a training block was created so that the drivers could familiarise themselves with the system and the test procedures before the test was carried out. The training block contained a sequence of the described reference versions of the scenarios. The composition of the test blocks is shown in Table 5 for an overview.

Table 5. Composition of the five test task blocks (A to E) regarding the included scenarios

Block ID	Scenario 1: Start	Scenario 2: Progression	Scenario 3: Conclusion
A	Dwarf signal directly at vehicle tip	Stop before person (orange vest) on the track	De-energised section
B	Signal in front of vehicle tip	Stop before brake shoe	Stop before red stop signal board in the track
C	Start before crossing switch	No special features	Stop in front of buffer stop
D	Signal in front of vehicle tip	No special features	Approaching a railway vehicle
E	Signal in front of vehicle tip	Signal unexpectedly reverts to stop	Stop before signal

Before the start of Block C, there was an additional task involving the driver's awareness of the vehicle's position. The drivers were asked where they were in the Mülligen shunting yard and in which direction they were looking.

Over the course of a test session, participants' individual performance preconditions change, e.g. their experience with the remote control system and the test course increases, and state variables such as motivation and fatigue can change over time. In order to be able to make statements about the difficulty of the individual test blocks, it therefore makes sense to balance the order in which they are carried out across the participants. At the same time, the balancing of the sequence must meet the requirements of a field test taking place in a shunting yard with otherwise normal operations and should be as manageable as possible. For this reason, three different sequences were defined for the blocks, which are characterised by the fact that each block occurs once in the beginning, once in the middle and once in the end part of the entire test run. In addition, certain operational requirements of SBB regarding the sequences had to be taken into account. The following three sequences were implemented: ACBDE, BDECA and ECADB. For the last 14 of the total of 36 test sessions, in each block sequence containing block C, this block had to be carried out first within the sequence for operational reasons, with the rest of the sequence remaining unchanged. The participants were randomly assigned to the block sequence conditions.

The experimental factor *ambient brightness* was varied within the participants. There were always two days between the two sessions by day and night. It should be noted that no balancing of the order in which a participant completed the test under the two conditions was implemented. This was due to safety considerations on the part of SBB: the participants should never complete the test under night conditions first, due to greater suspected difficulty. This should be taken into account when interpreting the results, as the results from the night condition are also always the results of the participant's second interaction with the system. In addition to the ambient brightness, the circadian conditions on the part of the participant also inevitably differ between the day and night sessions.

The experimental factor *availability of the ILTIS system* was varied across the sessions and participants in such a way that, for participants with a total of only one session (see below), the system was randomly either available or not. For participants with two sessions, in half of the cases ILTIS was available in the first session and not in the second, and vice versa in the other half of the cases. With regard to the effect of the ILTIS factor, it was not the overall performance that was of interest, but the completion of a specific task before the start of block C, in which the participants were asked to determine their exact location and indicate their viewing direction. Independent of the ILTIS condition, all drivers could use a conventional track plan and, if applicable, information from the video image to solve the task.

In the period from 5 February 2024 to 14 March 2024, the time slots from 8 a.m. to 4 p.m. were available on Mondays and Tuesdays for the tests in light conditions, while the time slots from 6 p.m. to 10 p.m. were available on Wednesdays and Thursdays for the tests in dark conditions. Unlike in the dark, the time windows in bright light offered the opportunity to carry out more than one test session in one day. The daytime sessions that had a nighttime counterpart took all place between 8:00 and 12:00. In order to make optimum use of the available resources, additional daytime test sessions were carried out in the time slot from 1 p.m. to 4 p.m. Due to the shorter time available, a shortened block sequence was tested here. According to the initial planning, only the first three blocks of the relevant sequence were to be tested. This was extended to four blocks from test session 14 onwards. A detailed overview of the sessions and their features according to participants, conditions and block sequences can be found in the appendix to this report (see 10.1).

6.1.6 Procedure

The recruitment of the participating drivers and the scheduling of appointments was organised by SBB. The test sessions took place in the six weeks from 5 February 2024 to 14 March 2024 within the time windows described above in a room reserved for the remote shunting workplace in an Alstom building in Zurich. In addition to the respective participant, at least one responsible person from each DLR, SBB and Alstom was present to perform various tasks as part of the test procedure. Those responsible of SBB coordinated the overall implementation, including the operational aspects, which included coordination with other SBB staff on the vehicle and on the track in Mülligen. The involved Alstom staff was responsible for the technical support of the remote workplace. DLR's task was to manage the experimental test and ensure the collection of data as part of the HTO analysis. A test protocol was used to structure the execution of the test and record the observations. The protocol document contained guidelines for conducting the test as well as standardised items to be completed by the HTO test manager on the circumstances of the respective measurement (e.g. weather, visibility conditions) and on the observations during the execution of the action (e.g. checklists for completing tasks and sub-steps). For the detailed evaluation, e.g. in order to determine the detection distances in the scenarios with obstacles in the track, video recordings of the execution of the task blocks were also made. After each task block, qualitative interviews were made and audio-recorded. The method of thinking aloud (Ericsson & Simon, 1999) was used to make the driver's mental processes during the execution of the tasks transparent. Participants were asked to say aloud everything that came to their mind during the execution of the tasks, no matter how unimportant it might seem or whether it appeared to be a good or bad thought.

An overview of the procedure within the test sessions is shown in Figure 4. The participants were received by the test team in a meeting room. The SBB test manager briefly provided some background information on the context and purpose of the study, data protection and anonymisation in the evaluation, and the participants confirmed their informed consent in writing if this had not already been provided in advance. The DLR test manager gave the participants an overview of the course of the session and the various types of data collection for the HTO analysis. The method of thinking aloud was explained and its importance in the context of the study was emphasised. It was emphasised that the research question was focused on the suitability of the system for different tasks and not on the evaluation of the participant's individual performance.

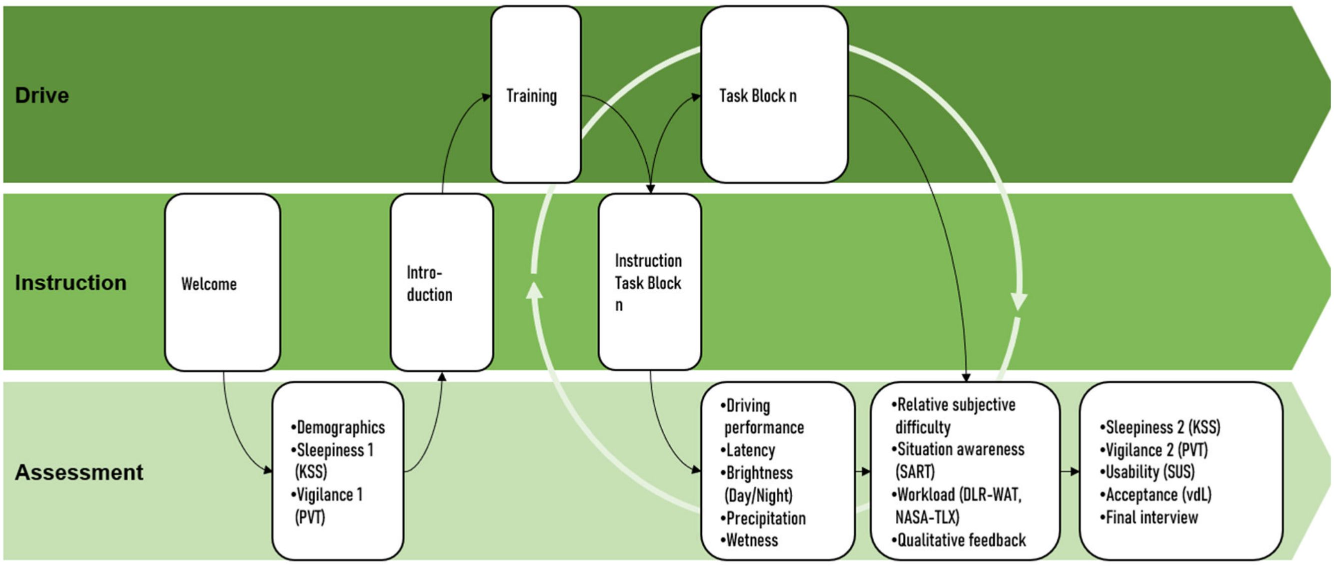

Figure 4. Procedure including times of data collection during the test.

The participant then completed a short questionnaire on their demographic background (age, work experience, local knowledge of the Mülligen shunting yard, experience with the Aem 940 locomotive), assessed their own current state on the Karolinska Sleepiness Scale (KSS, cf. 6.2.3.2) and completed the psychomotor vigilance task (PVT), all on a tablet computer. The participants then moved to the room with the shunting workplace. There, the participant received a briefing on the functions and operation of the system from the person responsible at Alstom. The participants then familiarised themselves with the use of the system by completing a training drive (all phases of the drive without any special features). The training drive could be repeated if desired. Before each drive, a brake test was carried out at standstill and for function, i.e. with the vehicle moving. After the training, the test phase began, in which each participant completed between three and five task blocks (see 6.1.5). The procedure was the same within each block. First, the DLR test manager gave the participant some information and instructions regarding the task at hand, which were formulated as generally as possible, e.g. “pay attention to the relevant signals and carry out the necessary actions”, “pay attention to the speed display and that obstacles may occur on the track.” More precise instructions were given for the two blocks that required precise positioning of the vehicle: “At the end of the drive you will find a buffer stop in the track. Please drive to within 1 metre of this” in block C and “At the end of the drive you will find a rail vehicle on the track. Please approach it with the locomotive as you would for coupling” in block D. Before each block, the participant was also reminded and asked to think aloud. At the beginning of each block, the DLR test manager noted the environmental conditions according to the protocol (see Figure 4). While the participant was working on the task block, the execution of the task was systematically observed and recorded using checklists. Following each task block, the participant completed the items from the surveys on subjective difficulty, situational awareness and workload presented in 6.1.3. This was followed by an interview on difficulties experienced and wishes or ideas for improving the workplace. Between the task blocks, the video transmission from the vehicle was switched off. While the participant worked on the follow-up surveys and interview, the test assistants at the Mülligen shunting yard brought the vehicle to the starting position for the subsequent block. Depending on the speed of the two processes, there were still waiting times for the start of the next block in some cases after the post-block survey. This must be taken into account when analysing the measures of subjective experience.

6.2 Results

6.2.1 Sample

6.2.1.1 Participants

A total of 24 SBB train drivers took part in the trials. The participants were aged between 20 and 59 (*M* = 41.4, *SD* = 12.1), predominantly male (*n* = 23) and had between one year and 33 years of professional experience (*M* = 13.4, *SD* = 10.8). The majority of participants (*n* = 14) stated that they were very familiar with the Mülligen shunting yard; five participants stated that they were moderately familiar and five that they were not very familiar. The sample comprised a wide range of different backgrounds from the SBB sectors of passenger and freight transport as well as infrastructure, including some locomotive test drivers.

6.2.1.2 Contextual conditions

The field trials took place in February and March. The hope of being able to make observations under frost and snow conditions during this time was not realised. The weather was predominantly dry, with rain only occurring during three sessions, in all cases at night. In one daytime session, the track was wet due to previous rain. Visual impairments due to fog or misted-up windows were not observed. The occurrence of glare or reflections was noted five times, twice during the day and three times at night. During the day, oncoming sunlight only caused a slightly greyish appearance of the video image from the centre camera. In seven sessions, image distortions or artefacts were recorded in the transmitted video images, which varied in nature: Some were sporadic short-lived image dropouts on individual cameras or flickering, others were artefacts in the playback, e.g. a coloured appearance of the lights of the dwarf signals, which are white in reality, at night at certain distances (including an appearance in red or green) or the appearance of phantom lights. In three session, it was reported that drilling noises were to be heard sporadically from another room.

6.2.2 Analysis by task blocks

6.2.2.1 Performance

6.2.2.1.1 Successful execution of tasks (effectiveness)

The success of task execution is a key parameter for answering the research question regarding the fundamental suitability of the remote shunting system for handling shunting tasks. For evaluation purposes, a success criterion was defined for each scenario, with the fulfilment of which the task was understood to be completely successful (see Table 6).

Table 6. Criteria for successful execution of actions by task block and scenario.

Block	Scenario	Criterion for successful execution
A	1 - Dwarf signal directly at vehicle tip	RO recognises signal position or contacts the RTC to obtain authorisation
	2 - Stop before person (orange vest) on the track	RO recognises the dummy and brings the vehicle to a standstill in front of the dummy
	3 - De-energised section	RO recognises the pantograph-down signal and stops before it
B	1 - Signal in front of vehicle tip	The course of events shows that RO has correctly recognised the signal position
	2 - Stop before brake shoe	RO recognises brake shoe, brakes and comes to a standstill before the brake shoe
	3 - Stop before red stop signal board	RO recognises the stop signal board and brings the vehicle to a standstill before the board
C	1 - Start before crossing switch	RO contacts RTC to enquire about the status of the relevant signal and can correctly indicate the position of their own vehicle
	2 - No special features	RO does not make any driving errors and does not come to a standstill without an operational reason
	3 - Stop in front of buffer stop	RO recognises the buffer stop and brings the vehicle to a standstill in front of it
D	1 - Signal in front of vehicle tip	The course of events shows that RO has correctly recognised the signal position
	2 - No special features	RO does not make any driving errors and does not come to a standstill without an operational reason
	3 - Approaching a railway vehicle (as for coupling)	RO brings the vehicle to a standstill at a distance = 0 or a short distance in front of the other vehicle. In case of contact, the speed is < v_{max} (no impact)

Block	Scenario	Criterion for successful execution
E	1 - Signal in front of vehicle tip	The course of events shows that RO has correctly recognised the signal position
	2 - Signal unexpectedly reverts to stop	RO recognises the stop case and brings the vehicle to a standstill in front of the DS
	3 - Stop before signal	RO does not commit a SPAD (Signal passed at danger)

Figure 5 shows the percentage of tasks successfully completed according to these definitions for all available observations from the field test. The scenarios at the first position within the block, which dealt with the correct perception of the driving release at the start of the drive, could be solved correctly with the system for the most part. A significant exception here is the scenario *Start before crossing switch*, which could only be completed correctly in two thirds of the cases. In this scenario, the relevant dwarf signal was always located far behind the head of the vehicle and could not be seen by the RO, so that they had to contact the RTC to enquire about the signal position or clearance of the route, respectively. For this enquiry, it was also necessary for the RO to be able to precisely state the location of their own vehicle. Although the ROs recognised the need to contact the RTC in all cases, the vehicle location could only be determined with sufficient accuracy and communicated to the RTC using the information available in the system in 65% of cases. Furthermore, not all ROs ensured that the relevant switch was set in the correct direction for the shunting movement when setting off. This aspect was not part of the criterion for task completion here, but represents an additional safeguard for action in order to recognise a possible own error or error of the RTC with regard to the position, which could have led to a flank movement. The reference scenario of a dwarf *signal in front of vehicle tip* was correctly solved in all cases. Under the more difficult condition of a dwarf *signal directly at the tip of the vehicle*, the relevant dwarf signal was not visible from any camera perspective in most cases , so that the RO had to request clearance from the RTC. In one of these cases, the RO relied on the instructions of the test supervisor to start the block and did not check the signal position with the RTC. In eight cases, the relevant dwarf signal was recognisable in at least one camera perspective and was correctly perceived by the RO. These cases are also included in the instances of correct task completion shown.

The scenarios in the centre part of the task blocks show the greatest differences in the success of completion. In the *person on the track* scenario, the RO recognised the orange warning vest placed on the track in all cases and reacted appropriately. The dwarf signal showing an unexpected stop was also recognised correctly in all cases observed. On the other hand, the early detection of the brake shoe on the track proved to be an almost impossible task with the tested system design, being successfully completed in around one fifth of the cases only. In most cases, the safety driver on the vehicle initiated a braking manoeuvre in front of the brake shoe before the RO was able to detect the brake shoe; in some cases, the detection occurred at around the same time as this braking manoeuvre was initiated. The scenarios in which *no special features* occurred in the middle section of the drive could be completed largely without errors. In a few cases, the maximum speed of 30 km/h was temporarily exceeded. These cases were not categorised as correctly solved.

The scenarios at the end of the drive were completed successfully for the most part. The *stop signal board*, the pantograph-down signal (*de-energized section*) and the dwarf *signal indicating to stop* were recognised 100%. In all cases, the ROs brought the vehicle to a standstill in front of the respective object, with the exception of one case in which the RO did not stop in front of the pantograph-down signal, but lowered the pantograph in good time, which was also included in the cases of successful completion in Figure 6. The scenarios in which the system had to be used to manoeuvre in a small space were also successfully solved for the most part. The approach to the required stopping distance was often carried out in partial steps, i.e. the RO stopped and corrected again. In three cases, the driver on the locomotive intervened in front of the *buffer stop* because the speed set from the remote shunting workplace seemed too high. The same occurred in one case when *approaching a railway vehicle (as for coupling)*.

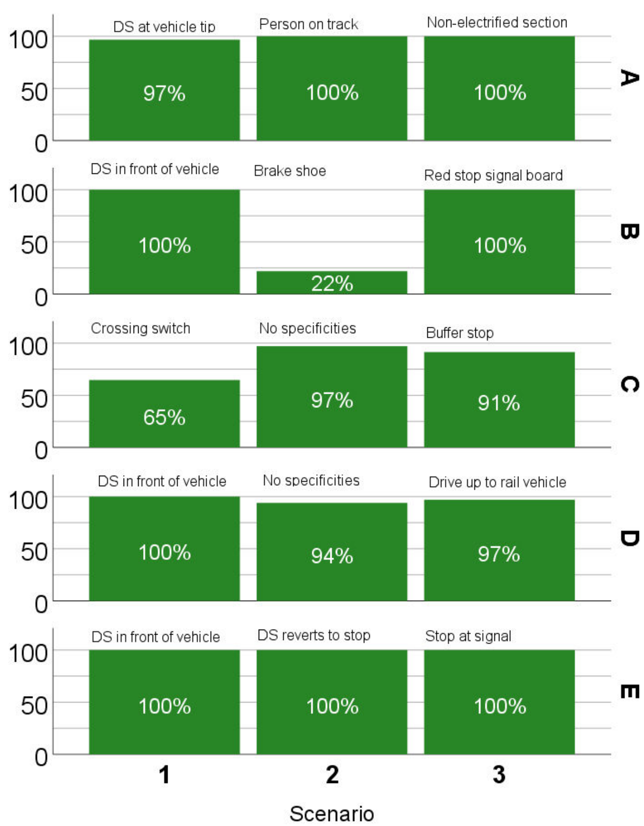

Figure 5. Proportion of tasks successfully completed by block and scenario

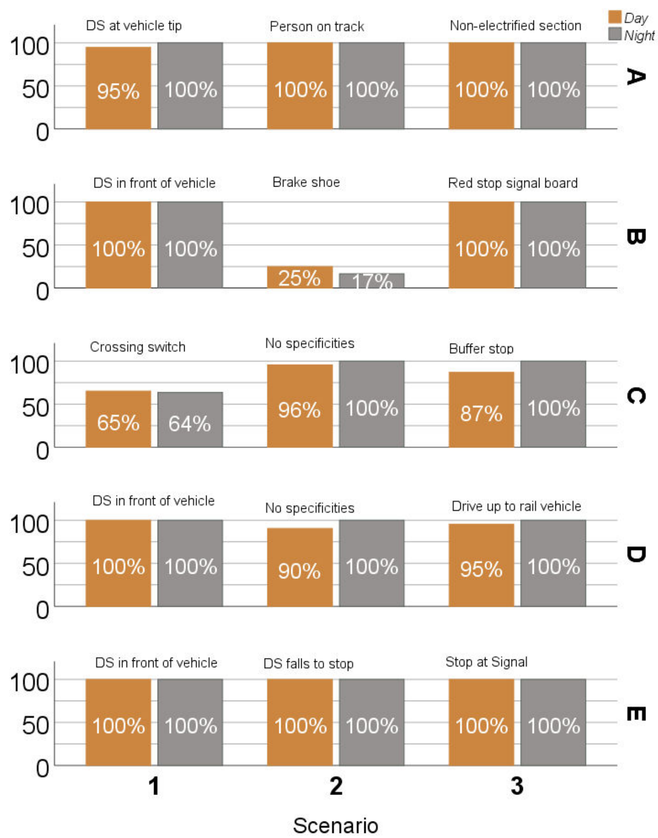

Figure 6: Proportion of tasks successfully completed by day (first session) vs. night (second session)

In Figure 6, the success of the task execution is broken down by day and night, whereby only the data of the 24 drivers who completed both conditions are included. It should be noted that the day session was the first interaction with the system in all cases and the night session was the second (see 6.1.5). In the night or second session, almost all tasks were successfully completed by the participating drivers, with the exception of starting the drive in front of the *crossing switch* and detecting the *brake shoe*. The detection performance was slightly further deteriorated at night compared to during the day in the brake shoe task.

The additional task before the start of Block C, which involved determining the vehicle location on the Mülligen shunting yard, was rated as correct if the RO was able to correctly state the location at the level of a track number or signal number. This was achieved in 65% of the observed cases (n = 34). The availability of the ILTIS display made no difference to the correct solution (72 % correct without ILTIS, 57 % correct with ILTIS, $X^2(1) = 0.95$, $p = .33$). In addition to the ILTIS system, the drivers had a conventional track plan and information from the video image at their disposal to solve the task.

To describe the perception performance achieved with the remote control system more precisely, the detection distances for the respective objects were estimated for four scenarios in which the visual perception of obstacles or events in or on the track was decisive. For this purpose, the time at which the RO reported detecting the respective object or circumstance was determined for all cases with successful task execution in the videos. The distance to the respective object of perception was estimated from the video images corrected for recording angle using triangulation. The results are shown in Figure 7. In a comparison of the objects, the orange warning vest was the most easily recognised by the RO; the mean distance at the moment of reported recognition was 122.3 metres (SD = 35.8). The red stop signal board was in second place in terms of detectability (M = 105.5, SD = 32.9). The dwarf signal demanding an unexpected stop was detected on average at a distance of 89.4 metres (SD = 32.9). In the few cases in which the brake shoe was detected before the safety driver intervened, the average detection distance was 43.2 metres (SD = 22.7). The distributions of the detection distances for the various objects are visualised as box plots in Figure 8. For the warning vest and the dwarf signal, the distributions of the individual values are relatively symmetrical around the mean value. For the stop sign and the warning vest, the median tends to be below the mean value, i.e. the detection distance was slightly smaller than the mean value in more than half of the cases.

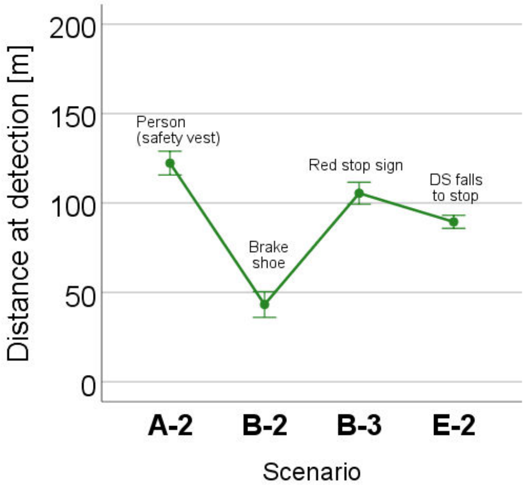

Figure 7. Mean distance to the object at the moment of detection (error bars: +/-1 SE).

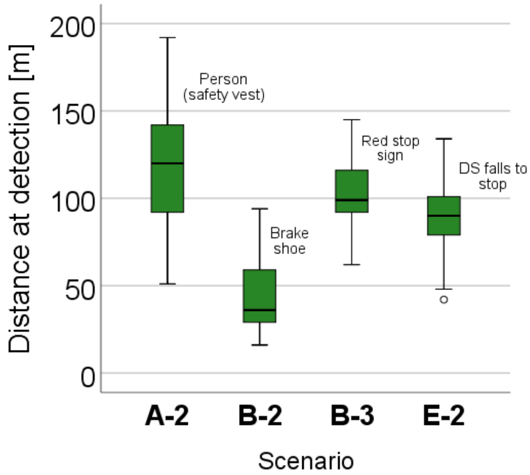

Figure 8. Distribution of the distance to the object at the moment of detection (boxplots).

Figure 9 shows the mean detection distances for the four objects in a comparison of day and night. Depending on the scenario, the mean difference is between three and 11 metres and is small compared to the variability of the values. Descriptively, the objects examined were detected slightly later at night than during the day, with the exception of the red stop signal board, which was detected slightly earlier at night.

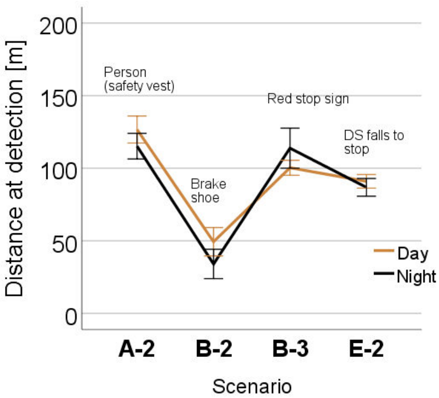

Figure 9. Mean distance to the object at the moment of detection by day (first session) vs. night (second session) (error bars: +/-1 SE).

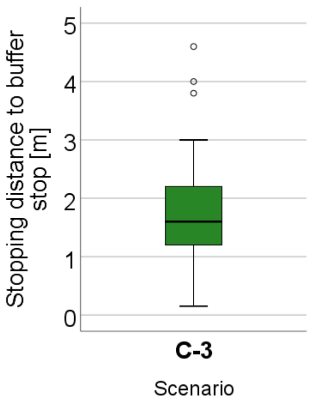

Figure 10. Distribution of the stopping distance in front of the buffer stop (box plot).

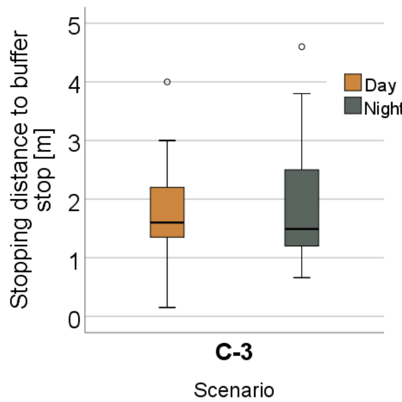

Figure 11. Distribution of the stopping distance in front of the buffer by day (first session) vs. night (second session).

To examine the precision achieved with the system when manoeuvring in a small space, the distance between the buffer and the buffer stop was measured for the scenario *Stop in front of buffer stop* after the final stop. The task instruction was to bring the vehicle to a standstill one metre before the buffer stop. The mean stopping distance produced by the ROs was 1.8 metres *(SD* = 1.0, *Min* = 0.15, *Max* = 4.6). The execution during the day and at night did not differ substantially, as shown in Figure 11.

It should be noted that the graph only includes the successful task executions in which the safety driver on the locomotive did not intervene by braking (87% in the first or day session, 100% in the second or night session, see Figure 6).

6.2.2.2 Relative perceived task difficulty

After each block, the drivers were asked how difficult they found completing the respective block compared to shunting on the locomotive. The assessment was made on an analogue scale with the minimum *much easier* and the maximum *much more difficult* by moving a slider, which was positioned at the centre of the scale – *just as difficult* – at the beginning. For the analysis, the centre of the scale was assigned the value 0 and the end points were assigned the values -50 and 50.

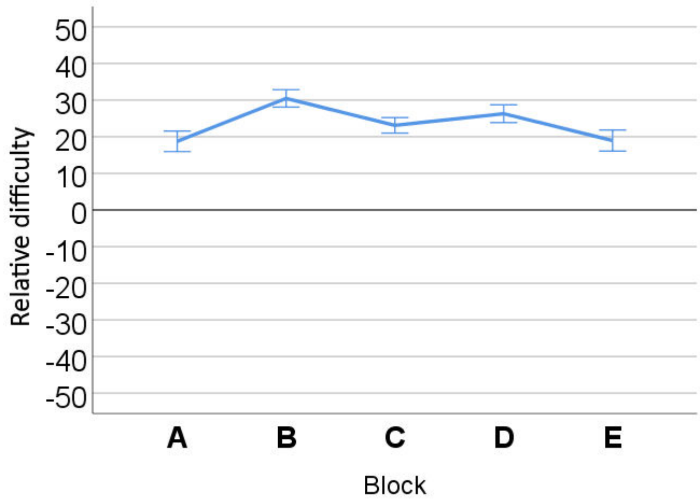

Figure 12. Mean value of the subjective relative difficulty of remote shunting compared to shunting on the locomotive, by task block. (Error bars: +/-1 SE).

The results are shown in Figure 12. Overall, the drivers found it more difficult to carry out the shunting tasks with the remote shunting system than on the locomotive. In comparison, the tasks in blocks A and E, which involved the detection of relevant signals and larger, relatively high-contrast objects on and along the track, were the easiest to perform (M_A = 18.8, SD_A = 15.8; M_E = 19.0, SD_E = 15.3).
A further increase in difficulty was perceived in block C, in which the vehicle had to be brought to a halt at a certain distance in front of the buffer stop (M_C = 23.1, SD_C = 12.4). The drivers found block D, with the task of approaching another railway vehicle as if to couple to it (M_D = 26.3, SD_D = 13.5), and block B, where the brake shoe had to be detected (M_B = 30.5, SD_B = 13.0), the most difficult.

6.2.2.3 Perceived workload

Two instruments were used to measure the perceived workload. The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) comprises six dimensions: *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort*, and *frustration*. Each dimension is rated on a 21-point Likert scale with the end points 0 - *very low* - to 20 - *very high*. In favour of performance efficiency, the contribution of each dimension to the overall workload experienced was not additionally weighted. In order to obtain an overall value for the workload experienced, the values of the scales are added together and standardised to a value range of 0 to 100.

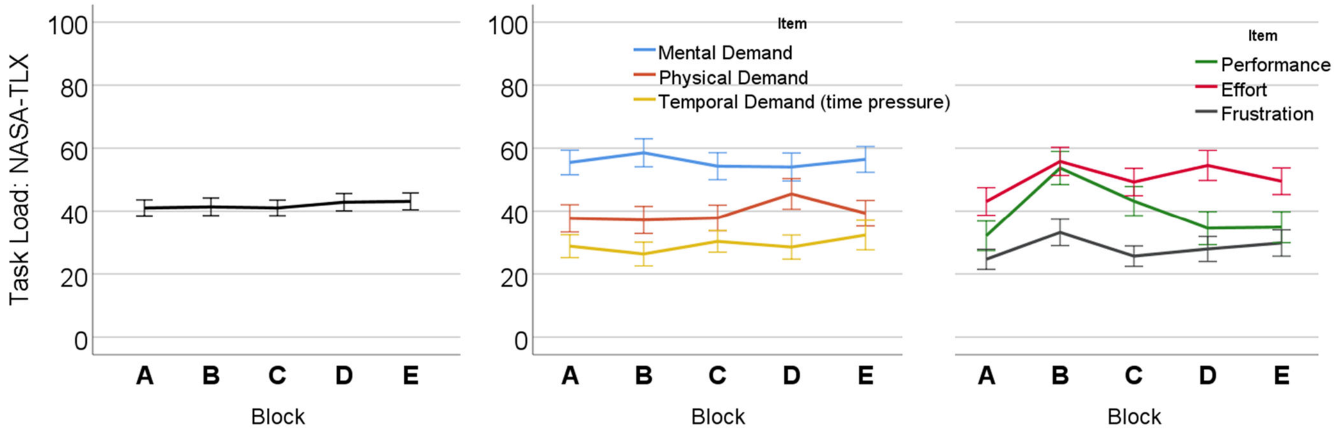

Figure *13* shows the mean values of the total workload per block on the left-hand side; the values are broken down by workload dimension in the centre and on the right-hand side.

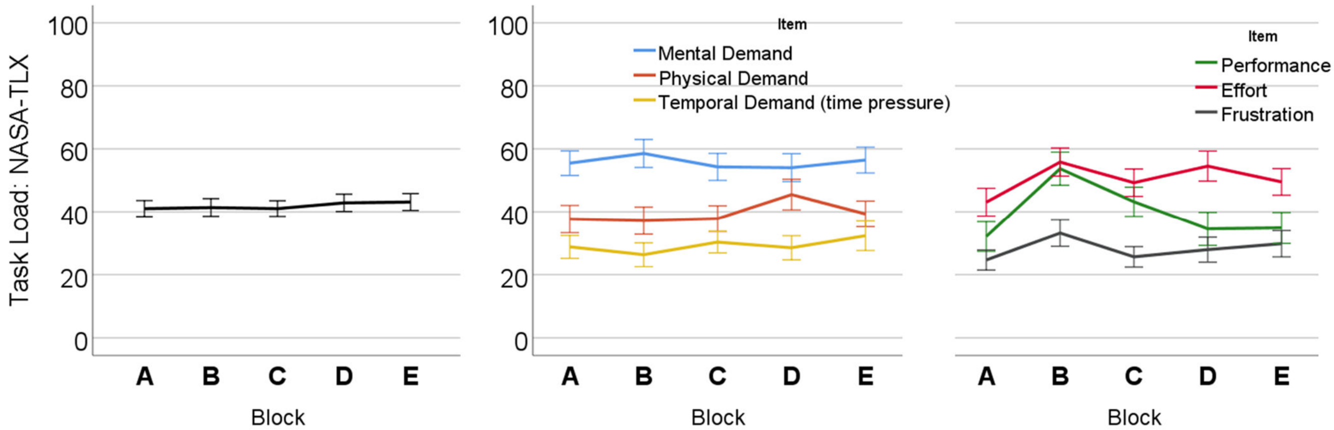

Figure 13. Left: Mean value of total subjective workload (NASA-TLX), centre and right: mean subjective workload on the TLX subscales, by task block. (Error bars: +/-1 SE).

The overall workload reported on the TLX is relatively constant in the medium range across the task blocks. A look at the various dimensions shows that the workload is primarily experienced on a mental level. The perceived time pressure in the experiment was relatively low. Differences between the blocks are particularly apparent at the levels of performance, effort and frustration. The pattern of effort required matches the results of the perceived task difficulty compared to shunting on the locomotive.

When designing workplaces, it is important to achieve an optimum workload level that does not lead to either underload and monotony or overload (Dunn & Williamson, 2012; Warm et al., 2008; Yerkes & Dodson, 1908; Young et al., 2006) by means of distributing and dimensioning tasks and designing work equipment. The DLR Workload Assessment Tool (DLR-WAT; Brandenburger et al., 2023; Grippenkoven et al., 2023) was therefore used as a second instrument to assess the workload experienced by participants. Like the TLX, it measures experienced workload on various dimensions, but on a scale that assesses the experienced workload level in comparison to the subjective optimum. In addition, a further distinction is made in the area of mental demand between demands in connection with *information acquisition*, *knowledge retrieval* and *decision-making*. The mental and physical demand are recorded on an analogue scale with the end points *extreme underload* and *extreme overload,* which are assigned the values 0 and 200 in the evaluation. The midpoint of the scale is labelled *optimal* and corresponds to the value 100. The temporal demand scale has the labels *too much time* (value: 0), *optimal* (100), and *too little time* (200); the effort scale has the labels *too little effort* (0), *optimal* (100), and *too much effort* (200). The frustration experienced is recorded on a scale from *not frustrated* (optimal) with the value 100 to *extremely frustrated* with the value 200, the perceived task performance on a scale from *very poor* with the value 0 to *very good (optimal)* with the value 100. The perceived optimum is therefore 100 on all dimensions.

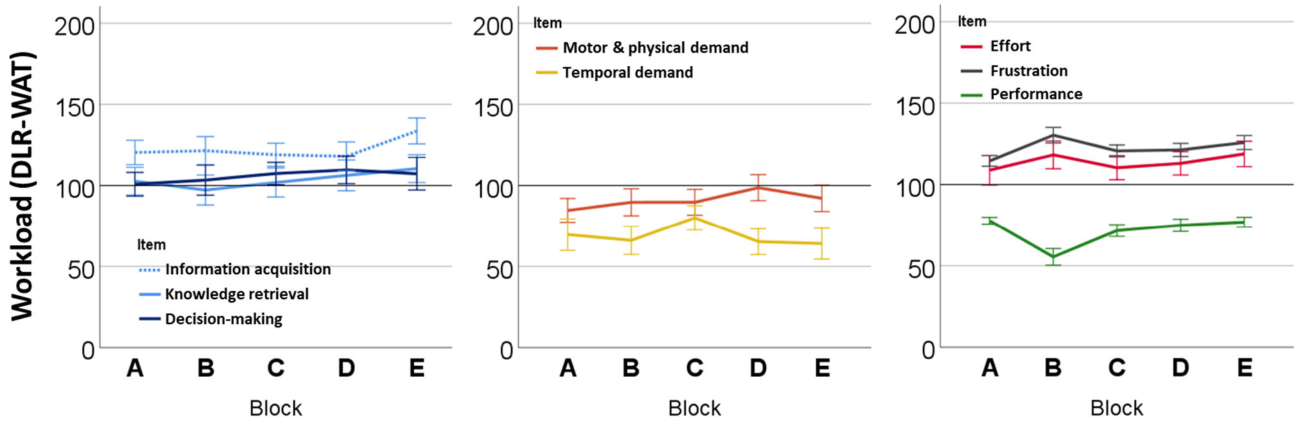

Figure 14. Mean value of subjective workload on the DLR-WAT subscales, by task block. (Error bars: +/-1 SE).

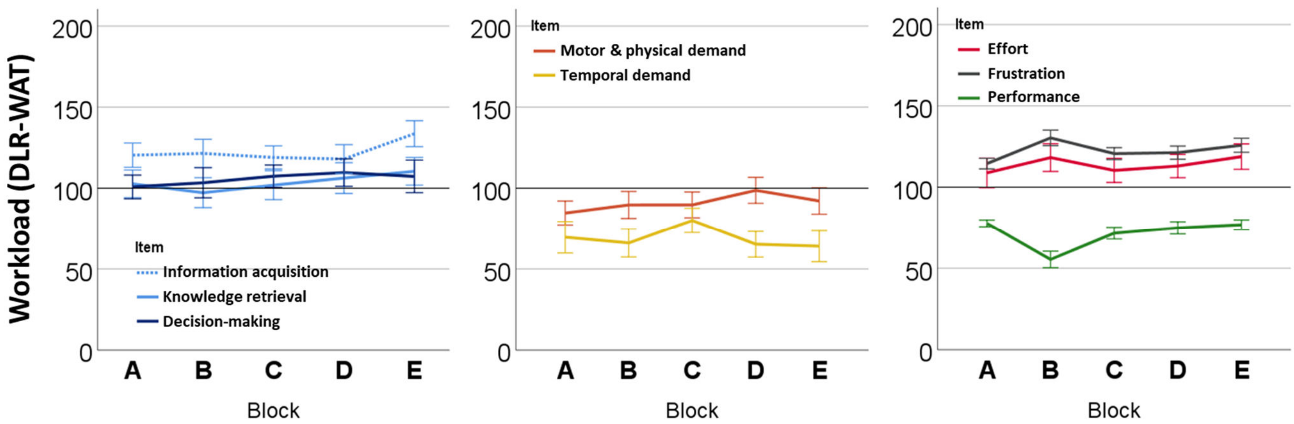

Figure 14 shows the results broken down according to the dimensions of workload for all task blocks. It becomes clear here that the higher mental effort recorded in the TLX is primarily due to greater demands in the area of information acquisition. The timing of the tasks in the test tended to be less demanding than the personal optimum. It should be remembered that this aspect was explicitly not in the focus of the study design (3.3). It cannot be ruled out that this experience is at least partly due to the waiting times between and within blocks that repeatedly occurred due to the necessary organisation of the test conditions in coordination with the normal operations in the shunting yard. The drivers were not completely satisfied with the task fulfilment achieved. This is reflected in the values in the dimensions of performance and frustration, to a greater extent in block B, which contained the brake shoe scenario.

6.2.2.4 Aspects of situational awareness

The term situational awareness (SA) refers to the perception and understanding of all relevant information in the environment that is necessary for making well-founded decisions. Endsley (1995) shows various approaches to assessing SA. In terms of implementation efficiency, the SART (Situation Awareness Rating Technique; Selcon & Taylor, 1990; Taylor, 1989) was selected for the present HTO analysis. The procedure comprises a total of 12 items in the three categories of attentional demands, attentional supply and understanding of the situation. As many of the items are similar to aspects that have already been surveyed, in particular workload, only four items were selected for the field test that are suitable for qualitatively enriching the picture of the perception and understanding of the TD: (1) *arousal*, the degree of alertness or readiness for action, (2) *division of attention*, attention to many vs. few aspects of the situation (3) *information quantity*, the amount of knowledge received and understood, and (4) *familiarity with the situation*, the degree of prior experience and knowledge, with the remark that respondents were to take into account both perceived familiarity with the driving situation in general and with handling the situation in remote-controlled shunting. In the SART, the values are recorded on a scale from 1 to 7 with the end points *low* and *high*.

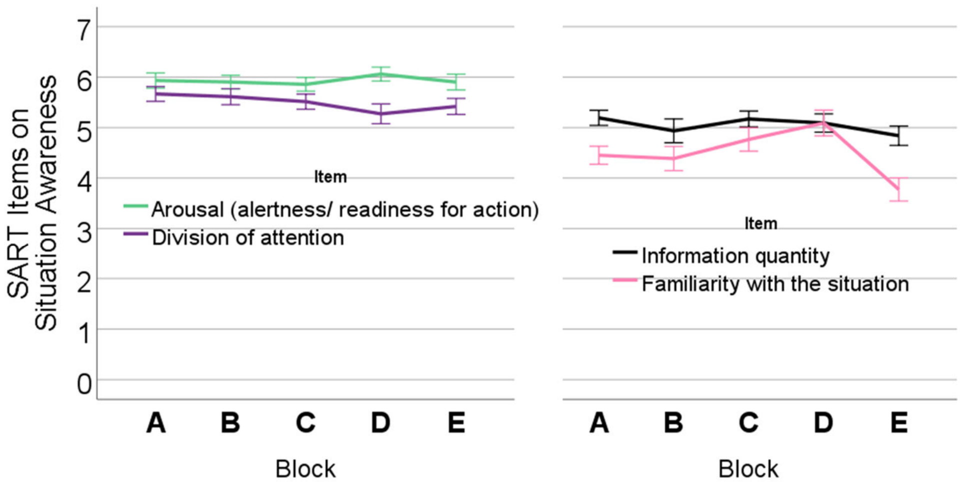

Figure 15. Mean value of the SART items selected, by task block (error bars: +/-1 SE).

The results are shown by task block in Figure 15. Both arousal (alertness) and the division of attention, i.e. paying attention to many aspects of the situation at the same time, were experienced as very high. This can be found equally in all task blocks. The perceived gain in information (information quantity) is in the medium to high range in all blocks. The perceived familiarity with the situation is in the medium range and is similar for blocks A, B and C. The highest familiarity was experienced in working at the remote shunting yard when approaching a rail vehicle. The situation in block E was less familiar, which is most likely due to the scenario of the dwarf signal unexpectedly falling to a stop.

6.2.2.5 Type of perceived challenges in task performance and solution approaches

After each block, the participants were asked about the following aspects: *(1) What challenges did you face when completing the tasks compared to driving on the locomotive? (2) What suggestions for improvement can you think of for this prototypical workplace that would have helped you to complete the driving tasks in this block? (e.g. with regard to available information, control options, etc.).* The responses were analysed using the audio recordings and noted down in keywords.

To further structure the content, the statements per block were grouped into subject areas, e.g. *control relatively coarse in the lower speed range* and *here I always have to hold the lever, otherwise it drops to zero - on the locomotive I can let go of it* to the category *lever control dynamics / traction control is different.* The topics were then categorised in descending order of frequency of mention. For the presentation in this chapter, those topic areas that were mentioned more than once per task block were compiled for each block on the challenges side in Table 7. All of the categories mentioned are listed for each solution approach.

Table 7. Type of perceived challenges and solutions by task block (overview). In the "Challenges" column, other feedback is marked as follows: [+] - helpful aspect of the design, [o] - other feedback.

Block	Challenges	Solution approaches
A	• Visibility is worse • Missing feel of the vehicle movement • [o] pantograph lowering signal and high-visibility west are easier to recognise at night than during the day • Pantograph lowering signal arrow can only be seen from a short distance • Braking distance estimation more difficult • More strenuous for the eyes • Operation is different / system behaves differently - you have to get used to it	• higher video quality • Make additional areas visible • DAS (driver assistance system) Obstacle detection • Position adverts differently • Front fog lights • Adjust lever control dynamics • Windows with fresh air • Display on curved screen • Visualisation of your own position at the station • Redesign of signals and equipment • DAS Obstacle detection
B	• Visibility is worse • Missing feel of the vehicle movement • Distance perception more difficult • Stop signal can only be seen late • Speed perception more difficult	• Higher video quality • Customise displays • Variable lighting on the vehicle can be controlled by the RO • Zoom function

Block	Challenges	Solution approaches
	• Position of switching points can only been seen late • Operation is different / system behaves differently - you have to get used to it	• Additional sensors • DAS signal detection • Adapt switch behaviour • Redesign of signals and equipment • Support the perception of movement more strongly • AP how to organise the stand • Lighting on site
C	• Signal position cannot be seen (behind vehicle tip) • Location determination more difficult • Missing feel of the vehicle movement • Distance perception more difficult • Speed perception more difficult • [+] BR camera helpful • Visibility is worse • No Hearing: Brake; no possibility to communicate with people on the track • More control of displays needed • Lever control dynamics / traction control is different • Operation is different / system behaves differently - you have to get used to it	• Make additional areas visible • Visualisation of your own position at the station • Additional sensors • DAS distance estimation • Design AP exactly like driver's cab • Higher video quality • DAS Obstacle detection • Customise displays • Front fog lights / cornering lights • Adjust lever control dynamics • Variable lighting on the vehicle can be controlled by the RO • Length scale on the rail • Display of LEA information on MDU or DDU • Vehicle should have a parking brake • Audio-transmission • DAS distance estimation
D	• Lever control dynamics / traction control is different • Speed perception more difficult • Missing feel of the vehicle movement • Distance perception more difficult • Slope/ gradient more difficult to perceive • It is not possible to check from outside whether the vehicle's wagons are braked • Frequent changes of glance between video and speed display necessary • [+] BR camera helpful • Visibility is worse • More control of displays needed	• Enable v-target setting to 0 km/h (results in coupling drive mode with 1 km/h) with monitoring of the actual v • higher video quality • Adjust lever control dynamics • Support the perception of movement more strongly • Audio-transmission • Zoom function • Customise displays • Adapt switch behaviour • Customise vehicle design • Make additional areas visible

Block	Challenges	Solution approaches
	• Braking distance estimation more difficult • Operation is different / system behaves differently - you have to get used to it	
E	• Visibility is worse • Stop pointing DS late recognisable • Recognisable course of action • Missing feel of the vehicle movement • Change of view between video above and below necessary; more strenuous	• DAS signal detection • higher video quality • Support the perception of movement more strongly • Variable lighting on the vehicle can be controlled by the RO • Replay function (to be able to watch previous sequences in the video again) • DAS route detection • DAS Obstacle detection • Make additional areas visible • Audio-transmission

Overall, the challenges named by the ROs are similar across the task blocks, but the importance shifts depending on the task. For example, in the three blocks A, B and E, in which the detection of objects and events in particular was required, the feedback of poor visibility is the top priority. In block D, with the approach to a rail vehicle, deviations between expectations and system behaviour when operating the levers for traction control came to the fore, and in block C, in which the train was started in front of a crossing switch, the limited field of vision was the number one challenge mentioned. Some of the topics in which various subordinate aspects are represented are explained in more detail below.

The aspect of *visibility being poorer* was mentioned very frequently overall and mostly related to the resolution achieved on the large monitors in the distance. On the extended Berne area video, the resolution was perceived as very good, but even this screen did not help the ROs to recognise distant objects at an early stage, as the covered range was too short. The ROs also often reported that they had noticed an object or “something” from a distance or that it appeared strange or in need of inspection, but then it took a long time until it was so close and clearly recognisable that they were able to identify it without any doubt (e.g. the dwarf signal in block E that had fallen to a stop, the additional arrow on the B-rail signal in block A). Recognising the route, i.e. the current setting of the points, was reported to be particularly difficult in the dark. The ROs reported increased effort and mental strain from trying to recognise everything in good time despite the more difficult conditions, as well as from the strategy of adjusting the driving speed downwards and slowly feeling their way forward in order to be able to stop in time before a possible obstacle that could be recognised late. Estimating distances was also occasionally cited as being more difficult.

With regard to the aspect of the *signal position not being visible (behind the front of the vehicle)*, it can be added that the ROs often reported on the strategies they use on the locomotive in such a situation (e.g. using the rear-view mirror, looking out of the window, getting out of the vehicle) and pointed out the time lost in the test due to the need to call the dispatcher.

Challenges in connection with the lack of feeling for vehicle movement were named in all task blocks, although mentions were somewhat more frequent in the blocks in which manoeuvring had to be coordinated in a small space. The lack of driving feeling refers to the absence of vestibular, tactile and proprioceptive feedback[1] at the remote control station. This made it difficult for the RO to perceive the starting and stopping or, in general, the acceleration and deceleration of the vehicle, partly also when controlling the speed to avoid exceeding the prescribed maximum speed. The occurrence of environmental factors such as wetness or gradient, which also influence vehicle behaviour and whose perception is made more difficult by the lack of somatosensory feedback, further increased the level of difficulty. Some of the ROs reported that they looked at the speed display much more frequently to compensate for the lack of driving feeling. The frequent change in fixation between the environment and the display was experienced as more strenuous.

Some ROs also mentioned the *lack of auditory feedback* as a challenge. In the test, this related to the perception of the brake being released and the contact made when approaching a railway vehicle. In one case, the lack of opportunity to communicate verbally with the surroundings (people on the track) was discussed. In this case, people on the track in the task block had shown no reaction to the approaching locomotive, and the RO stated that he had missed the opportunity to look out of the window and call out.

Perceived deviations from the usual in the *lever control dynamics / traction control* were frequently cited as a challenge by the ROs in Block D (approaching a rail vehicle as for coupling) and occasionally in the other task blocks. In most cases, this involved the behaviour of the drive-brake lever, which, unlike what many drivers were used to, fell to zero when released and did not remain in its previous position, so that the ROs were unable to apply a familiar strategy when manoeuvring in a small space (keeping the tractive force on the driving brake lever constant and regulating the movement with the direct brake). Additional feedback in individual cases concerned what was perceived as relatively coarse control in the lower speed range and a perceived delay in the effect of lever operation.

---

[1]Vestibular feedback refers to the processing of information from the vestibular organ in the inner ear, which provides information about the position and movement of the head in space; tactile feedback in this case mainly concerns the perception of vibrations; proprioceptive feedback refers to the perception of the position and movement of one's own body in space, which is mediated by receptors in muscles, joints and tendons.

Further challenges were reported by the ROs in the task of determining the location of their own vehicle. This was primarily due to the limitation of the area covered by the cameras and, in some cases, the video quality: features that were used for orientation on site, such as boards with the track number, were either not shown in any video image in the difficult cases or were too difficult to recognise, especially at night. As in the cases with the dwarf signal that could not be seen, the ROs missed the opportunity to obtain additional information about the surroundings, which is available to them on site, e.g. by looking to the rear or getting out and walking around the vehicle.

The solutions proposed for the further development of the system for improved task performance reflect the difficulties experienced in each case. At this point, the approaches should only be named in general terms and labelled according to the frequency with which they were mentioned. In first place was the desire for easier perception of obstacles, set routes, signals and other relevant information through higher video quality and / or the use of a DAS for obstacle and road detection. The approach of making additional areas around the vehicle visible was mentioned almost as frequently. In blocks C and D in particular, a DAS was often suggested to facilitate distance estimation. Occasionally, adjustments to the displays were suggested, mostly relating to a more central positioning of the speed display. The following options were also occasionally mentioned: a system to display the exact position of the driver's own vehicle in the shunting yard, a DAS for signal recognition, the design of the workplace as an exact replica of the driver's cab (particularly relating to the design and behaviour of the levers and switches), greater support for motion perception, supplementing the video information by visualising information from additional sensors, transmitting audio signals, integrating control elements for controlling variable lighting on the vehicle, adapting the lever control dynamics and supplementing a target velocity setting for the low speed range (e.g. travel at 1 km/h for coupling), whereby the actual speed travelled should be monitored to prevent overshooting of the setpoint. Other sporadically mentioned ideas included adapting the design of signals and other equipment in the shunting yard to the requirements of remote control (e.g. larger arrows on the pantograph-down signal, bright colour design of the brake shoes, length scale on the rail in front of the buffer stop, area-covering illumination) and the integration of a replay function (e.g. to make it easier to determine the location by seeing how the vehicle got to the current location). More detailed implementation options for the approaches mentioned are discussed in section 7.1.3.

6.2.3 Cross-trial evaluation

6.2.3.1 Perceived usability and user acceptance

The SUS (System Usability Scale; Brooke, 1996) was used to measure the perceived usability of the prototype remote shunting system. The instrument consists of 10 statements that are rated on a five-point Likert scale from *strongly disagree* to *strongly agree*. Examples of items are *I thought the system was easy to use* and *I needed to learn a lot of things before I could get going with this system*. An overall

SUS value is calculated from the individual values, which lies between 0 and 100, with higher values indicating better usability. An average SUS value is around 68.

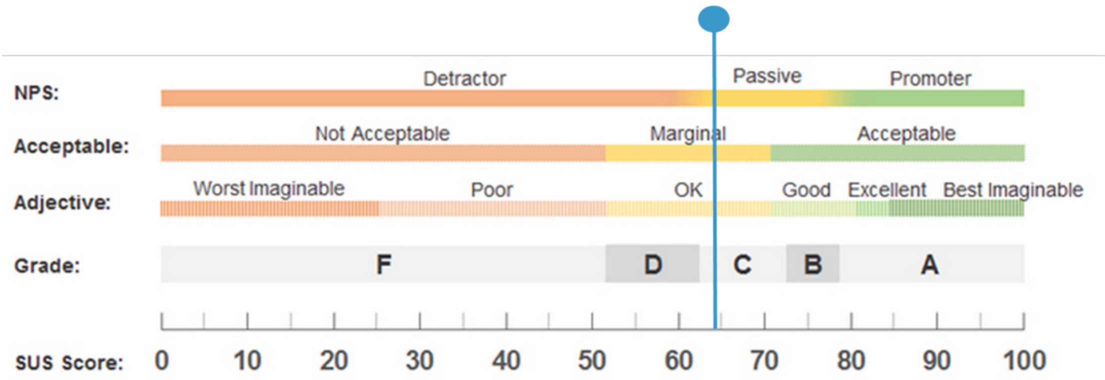

Figure 16. Illustration of the SUS value determined for the prototypical workplace using common classifications (see Sauro, 2018).

The mean overall SUS score for the prototypical workplace was 64.3 (*SD* = 12.6). The overall SUS scores of the individual TD ranged from 35.0 to 87.5. Several classification systems can be used to categorise the score, which are visualised in Figure 16 . The abbreviation *NPS* refers to the so-called *Net Promoter Score*, a key figure for measuring customer satisfaction, which is determined using the question *"How likely is it that you will recommend our product to others?"*; the terms *detractor*, *passive* and *promoter* refer to *dissatisfied customers* versus *satisfied but not enthusiastic customers* and *enthusiastic customers*. In the context of the HTO analysis, the categorisation with adjectives in the sense of acceptability or goodness (lines 2 and 3) appears most suitable. In both contexts, the SUS value achieved corresponds approximately to being mediocre or sufficient.

User acceptance was assessed using the van der Laan scale (1997). It consists of nine items with pairs of opposing word marks (semantic differential) as endpoints of a five-point scale. Values from -2 to 2 are assigned to the answers and the individual values are calculated into two overall values after appropriate polarisation: The perceived *usefulness* results from the scores on the word pairs *useful-useless, good-bad, effective-superfluous, assisting-worthless* and *raising alertness-sleep-inducing;*the *satisfaction* with the system from the pairs *pleasant-unpleasant, nice-annoying, likeable-irritating* and *desirable-undesirable*. Both parameters are just in the positive range (Figure 17). The individual assessments of the users with regard to usefulness ranged from -1.2 to 1.6 (*M* = 0.5, *SD* = 0.6), the satisfaction values between -0.8 and 0.8 (*M* = 0.2, *SD* = 0.4). The system has therefore not yet achieved a high level of acceptance at the stage of development tested.

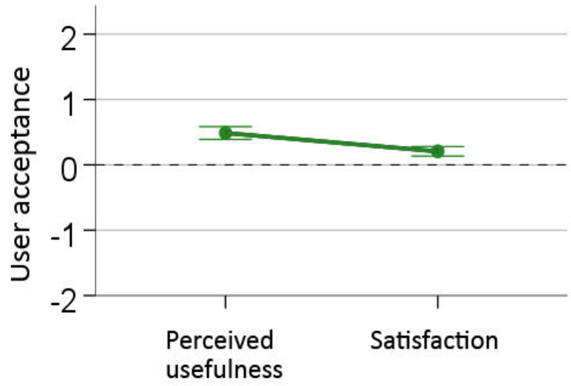

Figure 17: : Mean values of user acceptance of the remote shunting system (van der Laan scale; error bars: +/-1 SE).

6.2.3.2 Tiredness

At the beginning and end of each session, the participants were asked how alert they felt. The Karolinska Sleepiness Scale (KSS, Akerstedt & Gillberg, 1990) used for this purpose is a nine-point scale that ranges from *extremely alert* (value = 1) to *extremely sleepy – fighting sleep* (value = 9). On average across all sessions, the subjective tiredness after carrying out the tests (*M* = 3.6, *SD* = 1.4) was around one scale point higher than before (*M* = 2.8, *SD* = 1.0), but at a low level overall (scale value: 3 - *alert*).

6.3 Qualitative feedback (final interview)

The final interview was conducted orally with the participants at the end of the test session. The interview comprised seven substantive questions with an open response format from five areas:

General impressions of the system
- How would you describe your general impressions of the remote-controlled shunting system?
- Would the use of such a system add value to your work?

Safety
- Are there any specific safety concerns or improvements regarding safety that you would like to share?

Effects on the execution of actions
- Have you noticed a difference in terms of efficiency and precision when shunting compared to the conventional approach?
- Are there situations in which you would prefer traditional manual shunting?

Changes to the job profile
- How do you assess the impact of such remote-controlled systems on the working culture and self-image of train drivers?

Transmission options
- Are there other areas of railway operations where you see potential for the use of remote-controlled systems?

Participants also had the opportunity to share further thoughts and comments in an open debriefing. The results of the final survey are presented below along the five question areas.

6.3.1 General impressions

This section asked: *What general impressions do you have of the remote control system for shunting? Would the use of such a system add value to your work?* The responses were each assigned to one of three categories: (1) positive, (2) neutral and (3) negative impressions. An overview of the results is presented in Table 8. The *neutral* category does not have its own column there due to the small number of corresponding statements; the results are summarised in the text.

Most of the positive impressions relate to admiration for what is already technically possible, as the following quote illustrates: “Positively surprised, thought it would be harder and slower and more sluggish to operate. The usability was already very sophisticated”. In addition, the comments show that the respondents could well imagine using the remote control system in the future. However, the negative impressions also make it clear that some interviewees do not yet perceive the current status as usable: “The current status is unusable for me, i.e. for carrying out work. So I find this situation difficult”. Some interviewees expressed concerns about changes to their workplace and the job in general: “I'm also a bit worried if the workplace looks like this in the future. It's scary”. Another point of criticism in the general impressions relates to the perceived safety, which is considered to be too low when working with remote control in shunting operations: “It doesn't get any safer when you work with the system”. The safety concerns also related, for example, to the possible effects of a technical fault, such as a power failure or contact breakage. With regard to night-time driving, some respondents emphasised that obstacles became visible too late. In the *neutral* category, there was one mention relating to the perceived immersion: *It is very special to switch that you only sit in front of one screen*.

Table 8. General impressions of the RO with regard to the remote control system

Positive	negative
• Impressive to see what is technically possible • It's something new, a new experience. I can well imagine using this in the future. It's amazing how much you can see. • I can already imagine that it could work well in the future, i.e. the basic idea • I am in favour of this. It will come at some point and you should slowly familiarise yourself with these processes • Positively surprised, thought it would be harder, slower and more sluggish to use. The usability was already very sophisticated • Overall impression is positive.	• It's more strenuous and more intense than riding the locomotive properly. I don't feel so confident either. Because you don't know where you are, it's more difficult to see obstacles. The strain is mainly on the eyes and the uncertainty. • Driving behaviour not available, so you react differently from a distance • The position of the switch points is visible later when travelling at night • Operation is perceived as more difficult • Kinetic feedback on the movement of the train is missing

Positive	negative
• Operation is not very difficult, even easier than a driving simulator. • The image is very realistic. • positively of the possibilities and the presentation	• From a distance you lose the feeling for the conditions, the machine reacts completely differently when it is wet • The current situation is unusable for me, i.e. I can't do it at work. So I find this situation difficult. • It does not become safer if you manage with the system • I'm also a bit worried if the workplace looks like this in the future. It's scary • It still needs improvement, definitely. • the picture quality is somewhat blurred and there is distortion • screens are very large and very close, which is very tiring for the eye • Orientation is difficult when you are suddenly transferred. Otherwise you are already sitting in the locomotive and are orientated • Fear that an entire profession will disappear • It's exhausting to sit there for 8 or 9 hours. It's exhausting for your eyes and your concentration • It is difficult to see some obstacles, like the scotch block • In my current position, there is no added value for me because I don't do my job to work on a screen

The respondents also commented on the perceived added value for their work. Some of the interviewees see added value for their own work. This is based on perceptions and expectations regarding various factors:

- **Less cognitive strain**: “It could be an advantage for concentration that certain things no longer occur that you no longer notice, such as other trains passing by”
- **Stronger local networking with colleagues**: “The positive thing would be that you could take breaks more flexibly and also get coffee. In future, teams would be spread across Switzerland, you wouldn't be alone but with several people in one room, you would have more contact with colleagues”
- **Independence from the weather**: “Positive would be that you are not so exposed to the weather”
- **Lower risk of accidents at work**: “… and fewer accidents at work are possible” and “the safety aspect, as walking on the track can be dangerous”

- **Deployment on an inaccessible route**: "I think it could be added value for certain situations that would otherwise require you to be on site, e.g. in the event of a breakdown, if a vehicle has broken down"
- **Time efficiency**: "It is useful because you save resources as you don't have to walk to the vehicle, especially in bad weather" and "yes, for operations, as you can be much more efficient. Waiting and travelling times in the staging area can be reduced"
- **Additional information**: Bernese clearance zone camera brings added value

However, another part of the interviewees also sees no added value or no current added value, but only future added value: "Not at the moment, but in the future with automated partial operation, then yes" and "… but it still needs a lot before it is developed to the point where it can be used for everyday life". Some interviewees noted that the added value is limited as long as people still have to be on site: "Only when you can operate it on your own and no longer need people on site does it have added value".

Other respondents see added value for other areas – "Added value? No. Not for infra, as every construction site is different. Perhaps in the area of passenger transport, because there is not so much variation there" - or for the overall system, but not for myself: "I would not see any added value for myself directly. I see more added value at an overall level".

6.3.2 Safety

In the analysis of the interviews regarding the respondents' safety concerns, the information was categorised into eight categories: (1) *new stress and insufficient challenge*, (2) *distancing*, (3) *increased concern*, (4) *data- quality and reliability*, (5) *latency*, (6) *lack of information*, (7) *orientation problems* and (8) *responsibilities*. The evaluation of the security concerns makes it clear that the data quality of the videos in particular plays a decisive role in the reliability and security of the system. In addition, the interviews show that the interviewees have safety concerns due to a lack of information, such as a lack of all-round visibility. The latency of the system is also perceived as safety-critical. Table 9 provides an overview of the categories and exemplary quotes.

Table 9. Perceived sources of potential safety challenges

Category	Exemplary quotes
Changed stress or insufficient challenge	• "The challenge is to stay focussed, as it is easier to wander off" • "Sitting in front of the PC and staying focussed is difficult" • "You get tired faster when you're not outside"
Distancing	• "The direct cooperation between man and machine in the locomotive makes the system safe"
Changed concern	• "The feeling of safety is different because you are not directly affected by an accident, you are interested, but you are not directly involved on site"

Category	Exemplary quotes
Data quality and reliability	• “Image quality is very important. For example, you could overlook the inhibitor. Image quality must be 100% and not 99%“ • “Technical challenges are even greater at night, so better video quality is all the more important“ • “It is always possible that there is an obstacle on the track. You only see some things at the last moment. It is not technically possible to make it as good as it is in reality“ • “Vision is a key aspect of safety. Visibility can be restricted if the camera is dirty“ • “If there are people on the track, you notice this later and less and obstacles in the track are also more difficult to recognise. • “Scotch block was not visible, which is a safety risk“ • “If the Internet connection is interrupted, this is critical“
Latency	• “When I move levers and how this affects the locomotive is a matter of judgement, but the 1 or 2 seconds of transmission time can be crucial if someone runs in front of the locomotive“ • “You can roll back later with“ • “The severity of braking and accelerating is higher“
Missing information	• “The blind spot cannot be seen, in the worst case I get out of the locomotive and walk around. • “You have to rely entirely on the screen, which is problematic because there are no auditory signals from the surroundings, e.g. horns from other trains. • “You don't feel anything, you don't get any kinesthetic feedback. It would certainly be very different with a pendant on it“ • “There are too many inadequacies that can only be ensured by visual inspection, such as an overall view, looking left and right or to the rear“ • “Critical at roadworks, because there may be people on the track and you do not have full all-round visibility“
Orientation problems	• “Orientation is difficult when you are suddenly transferred. Otherwise you are already sitting in the locomotive and are orientated“
Responsibilities	• “The problem is who checks that the vehicle is parked correctly and that there is no damage to the vehicle?“ • “What happens in the event of an accident? Does a person have to travel to repair it? “ • “There will still be people working around it“

6.3.3 Perceived effects on the precision and efficiency of the execution of actions

The perceived impact on working practices was assessed in terms of precision and efficiency. The results for the two areas are presented below.

Precision

- When shunting manually, you can see out of the side windows and may be able to see and hear more (e.g. whether the scotch block has been removed or wagons are coupled)
- On site I can walk two metres to the right and look out of the window to the rear to see the dwarf signal or get off the locomotive to see if the dwarf signal is open
- the latency is good, but when driving with a trailer, inertia is certainly higher
- is more of a stretch and will probably also be high when shunting remotely
- As you have no sense of movement, it is difficult to assess movements, e.g. whether you are rolling backwards
- Seeing obstacles is more difficult and uncertainty is greater because you recognise details later, for example whether the dwarf is open or closed
- Delay is not as bad as I had thought
- Accuracy is very good at shorter distances, but worse at longer distances
- Control is very sensitive and you can control the locomotive very well

Efficiency

- Converted speed is lower than when travelling on the locomotive, some shoot up to 50% slower
- Slower speed is mainly due to more defensive driving behaviour due to lack of precision
- Only slightly longer, as haptics are missing
- When driving at night, speed is even lower due to limited visibility
- Efficiency is reduced because you have to call the traffic controller if you have not seen the dwarf signal.

6.3.4 Changes to the job profile and work culture

The majority of interviewees rated the impact of remote control on changes to the work culture and the self-image of locomotive drivers as strong. One interviewee summarises this succinctly: "This will be a jolt." Another interviewee is more critical: "A completely different way of working. There could be someone sitting in the office who has no idea about the track." The answers also show that the interviewees are aware that their work is currently undergoing a major transformation and that change is inevitable: "The locomotive driver's job will change anyway, we have to be honest about that." Some interviewees were also sceptical about remote control because they wouldn't want to miss the activities on the locomotive: "It would be a step back for me as a locomotive driver because I like driving". Another interviewee expressed a similar opinion: "I think the locomotive driving profession will suffer quite a bit".

One interviewee also commented on the job title: "A different job title is needed instead of locomotive driver, similar to drone pilot.

One interviewee was critical of the prospect of jobs being lost as a result of remote control: "It's a massive number of jobs that will be lost. And not everyone can become a computer scientist or something like that.

Many interviewees estimate that older colleagues will find it more difficult to cope with the changes, as the following quotes show: “Older calibres with monopoly training would find it more of a burden than people who have done various jobs“ and “that’s not for me anymore“. This image is also emphasised by the following quote: “With time, it's all right, but now there are still older generations who had the childhood dream of driving a steam locomotive. But the enthusiasm for the railway is waning, so perhaps it will be easier to introduce it“. Another interviewee was more positive about the job prospects for younger generations: “If you look to the future, then yes. Generation X and older will retire and the labour force will shrink. There are already too few employees. This will be particularly interesting for the "Playstation generation", who prefer to sit in front of their computers rather than run around at night with an iron bar“.

The respondents also expressed some concerns about the specific organisation of their workplace. For example, the survey shows that working from home is rather undesirable: “I can't imagine working from home because of safety concerns“. In addition, some interviewees expressed strong concerns about the reduced physical activity that remote control entails compared to manual shunting, as the following quote illustrates: “The negative would be that I would be inside all day“. On the other hand, interviewees also mention the advantage that remote control would reduce the physical strain of the job.

One interviewee suggested that an alternating model would be advisable, which would provide for shifts with remote control and shifts with manual shunting on the locomotive: “I can well imagine working one week like this and another like this as a change“.

The interviewees see particular advantages of the change in work culture in the increased local collaboration with colleagues, which also makes it easier to take breaks together. On the other hand, other respondents tend to see a loss of social contact as a result of remote working: less communication and social interaction between employees and a lack of social interaction.

6.3.5 Perceived potential use cases for remote control

The interviewees see several areas of potential for remote control. The analysis clearly shows that the transmission possibilities differ greatly and that some respondents even contradict each other in their views. The potential areas mentioned are

- for self-contained systems and branch lines
- on long monotonous routes, such as the ICE from Zurich to Berne without stopping
- Shunting trips to remove defective trains from the track and towing broken-down trains
- to put away trains in the evening when they are parked for the next morning

- In passenger transport, as less variable than when used on construction sites
- mainly for shunting, not for passenger transport and long distances
- for emergency situations when there is disruption
- for cargo transport
- for empty drives
- Automatic commissioning during shunting in stations equipped with ETCS (AIRA project) can save a lot of time if trains commission themselves
- monotonous shunting tasks, such as loading sugar beets, then you can remotely control other locomotives while waiting for the next wagon to arrive

6.3.6 Suggestions for improvement

The interviews resulted in numerous suggestions for improvement that can be categorised into different subject areas.

Camera position/ and image resolution

- Image resolution must be higher
- Locomotives with a larger front end should be fitted with larger camera angles
- Depending on the vehicle type, cameras should be dynamically customisable
- Additional camera for rear view, which shows, for example, whether all wagons are attached
- Coherent image for all-round visibility
- It would be more realistic if you could see the engine block at the front
- Adaptive camera position to get a view from a standing position over the edge
- Cameras should be installed at the front of the stem, especially for night riding, so that the camera is close to the lighting (closer to the light source)

Display of further information

- Power-off areas should be displayed in an overview of your own position on the ground
- Enable audio transmission (e.g. feedback as to whether the brake is applied)
- you should be able to hear the noise and signals from outside
- System that recognises and warns of small objects in the camera image
- Warning system would be helpful if, for example, you can't see the scotch block

Representation

- the speed display should be positioned further to the centre

Structure of the workstation

- Use of a dynamic driving simulator that simulates locomotive movements
- Would prefer to sit a little further away from screens, as screens are too big and you have to move your head too much to see everything

Co-operation

- Create redundancy through teamwork. If you always work in pairs as a team, it's not too strenuous to sit in front of the screen for so long
- A personal fallback level using a camera in the driver's cab so that a second remote driver from the control centre can help the driver in the locomotive in the event of a shock or similar.

Education

- New locomotive drivers must be able to gain experience with materials in order to develop an understanding of technology

7 Discussion

7.1 Summarising the answers to the research questions

7.1.1 How suitable is the prototype system in principle for carrying out the shunting tasks?

The HTO analysis provided the opportunity to test the usability of the prototype remote control system in terms of effectiveness, efficiency and user satisfaction with a comprehensive sample of locomotive drivers using a wide range of typical shunting tasks under day and night conditions.

The effectiveness achieved, i.e. the basic ability to successfully complete the tasks set, was good overall. One exception is the task of detecting an obstacle on the track, which was only successfully completed with the system in less than a quarter of cases. A second exception is the task of determining the position, which could only be solved accurately in two thirds of the cases. The position of all dwarf signals shown on the transmitted video images was recognised correctly without exception and successfully implemented in the driving behaviour. This included a scenario in which a dwarf signal unexpectedly fell to a stop. Manoeuvring over a short distance, as required when stopping in front of a buffer stop and when approaching a rail vehicle for coupling, was successful in most cases. In a few cases, all of which occurred in the first session of the respective participant, there were unintentionally high speeds in the approach, which led to the safety driver intervening. When travelling after signals without any other special features, the only incident that prevented the scenarios in question from being processed completely correctly was the occasional exceeding of the maximum speed of 30 km/h by a small number of participants. Other scenarios with special features such as detecting and reacting correctly to a person (dummy with warning waistcoat) on the track, a stop signal board and a section without power were successfully solved with the system without exception.

In addition to effectiveness, efficiency as a quality indicator describes how much effort is required to provide the service with the system. In the analysis, aspects of the time required and aspects of mental strain were recorded. The information provided by the participating drivers can be used to compare the time required to complete the tasks with the remote shunting system compared to the reference system. Around half of the drivers stated in the final interview that they had used the system to implement a slower driving speed than on site on the locomotive. Among other things, the speed adjustment was made to compensate for a poorer perception of the route, the driver's own speed and small obstacles in the route. Comparatively large efficiency losses also resulted from the limited all-round visibility and the resulting need to call the dispatcher to find out the position of signals at or behind the front of the vehicle. The drivers rated the mental strain in terms of the difficulty they experienced with the system as higher than when driving the locomotive. Overall, the NASA-TLX values for mental stress were in the medium range, with the values in the dimensions of mental demands and effort being slightly above the centre of the scale and above the other dimensions. The DLR-WAT values also showed that the increased mental demands resulted primarily from increased demands in the area of information intake. These demands, as well as the values on the dimensions of effort and frustration, were slightly above the subjective optimum, whereas the values on the dimensions of motor-physical and time demands as well as task mastery remained slightly below the subjective optimum.

User satisfaction, i.e. the extent to which the users rated the user experience positively and felt that their expectations regarding use had been met, was in the medium range. The SUS value determined corresponds to a rating of “medium moderate“ or “sufficient“. User acceptance of the system, as determined using the van der Laan scale, was just positive, although there is still plenty of room for improvement.

To summarise, it can be stated that the tested system proved to be fundamentally suitable for the majority of the shunting tasks tested, although the efficiency of the execution of actions was impaired primarily due to perception deficits and partly due to other aspects. The nature of these deficits is described in more detail below.

7.1.2 Where do problems still occur and how are they related to the system design?

When summarising the aspects identified as needing improvement in the following section, we are guided by the frequency with which they were named by the test subjects or observers, starting with the most frequent. Overall, the majority of the problems named can be attributed to the lack or reduced quality of various sensory information available to a test subject on the locomotive. Somatosensory and visual aspects were named with roughly equal frequency.

The lack of vestibular, tactile and proprioceptive feedback reduces awareness of the vehicle's movement compared to shunting on the locomotive and makes control more difficult. This is particularly true when manoeuvring over short distances with the associated processes of slowing down and accelerating. Any additional factors that influence the vehicle's movement - e.g. wetness or bumps - make the situation even more difficult. The feedback from the drivers also indicates that the perception and control of speed while driving is also altered by the lack of a “popometer”. To control speed, drivers had access to information from the optical flow (Gibson, 2025; Lee, 1980; Wolfe et al., 2017) of the video streams on the one hand and the speed display on the other. If the lack of information leads to increased gaze control on a head-down speed display, correspondingly less time can be used for route observation (Brandenburger et al, 2017; Wickens, 2008). With regard to the optical flow, it was found that the resulting speed impression differed greatly between the two monitor categories: due to the finer resolution of small structures, the speed on the Bernese area image was perceived as higher than on the large monitors (see 6.4.1).

Based on the video images available, the participating drivers found it much more difficult to recognise the position of the switch points and the resulting routes than when shunting on the locomotive. This applies in particular to night-time operation, but was also reported in some cases during the day. Recognising the position of dwarf signals was also reported to be more difficult in some cases. Timely detection of a relevant small object such as the scotch block on the tracks was almost impossible with the system. The ballast lowering signal as a whole was easy to recognise, but it was only possible to detect the small arrow that indicates the assignment to its own track[2] by carefully approaching and inspecting it from a short distance. It can be assumed that the causal parameter of the video quality in the tested system design was the image resolution achieved in the preview. To maintain safety, the drivers often compensated for the difficulty in perception by selecting a correspondingly reduced driving speed. When using a remote shunting system in the present design, this form of efficiency loss or safety reduction would have to be taken into account if the RO had to make a different choice when balancing speed and accuracy. Other visual perceptual performances reported as more difficult included the estimation of distances and the perception of sensations. Just like the lack of somatosensory information, these aspects have an impact on the control of vehicle propulsion for situation-appropriate acceleration and deceleration.

In addition to the resolution of the video images, the choice of environmental sections visible in the video was also identified as an aspect complicating the execution of the task. Among other things, this concerned orientation with regard to the position of the vehicle within the shunting yard. The latter is made even more difficult by the fact that, unlike when shunting on the locomotive, the RO can find himself transferred to a new, unknown location without any transition when taking over a vehicle. In a third of the cases, the drivers were unable to develop a precise situational awareness of where the vehicle was located. In combination with the inability to see a dwarf signal[3] located directly at or behind the head of the vehicle, which is also caused by the limited visibility of the environmental sections, the problem of difficult orientation is exacerbated. In the corresponding scenarios, the drivers had to contact the dispatcher to enquire about the status of the route. Apart from the major delays in the course of action that this caused, the inadequate orientation increases the probability of errors due to incorrect location information. With regard to other perceptual abilities, some drivers also felt that the possibilities for obtaining information from the vehicle's surroundings were inadequate, for example when building up situational awareness with regard to people or objects around the vehicle or in lateral proximity.

---

[2] In the test, the articulated signal was positioned on the right, which can occur in certain situations and requires the additional arrow on the signal to be recognised in these cases. Signals in Switzerland are generally positioned to the left of the track.

[3]Note: The design of the signals also allows the signal position to be recognised on their rear side.

A perceived delay in vehicle reactions was only occasionally mentioned as a difficulty. The rare mentions of this aspect occur predominantly in the two tasks for manoeuvring in small spaces, i.e. stopping precisely in front of a buffer stop and approaching a rail vehicle. On the one hand, a perceived delayed braking effect in the *dwarf signalling* scenario was *fall to stop*. In the given field test, it is not possible to assess the extent to which these perceptions are due to the transmission latency of the remote control system and / or to a reaction characteristic specific to the test locomotive. It should be noted that the tasks in question were successfully completed in most cases, albeit with increased effort and time, but that the safety driver also intervened in some cases. Further research is required to determine critical latency values (see also section 7.3).

Another, but less frequently mentioned, aspect concerns the differences between local and distant shunting with regard to the accommodation (distance adaptation) of the eyes. In the real driving environment, there is a change between near and far accommodation. In comparison, there is less looking into the distance at the VDU workstation. This was reported by some drivers as more strenuous and fatiguing in terms of concentration. This perception is consistent with the physiological processes involved in accommodation. Added to this is the generally increased effort in the perception of objects, which was used by many TD as a reaction to the perceived poorer vision.

The lack of auditory feedback was occasionally mentioned as a challenge; the problems mentioned related to the fact that the ROs did not hear how the brake was released or how the vehicle touched the other vehicle when travelling in a coupling. One RO commented that without sound transmission, the events felt less real, like a computer game. In one case, the possibility of being able to make verbal contact with people in the vicinity was missed.

Perceived deviations from the usual lever control dynamics were occasionally described by the drivers as challenging. This mostly concerned the behaviour of the driving brake lever, with which the drivers concerned were unable to apply their usual strategy when manoeuvring in small spaces. In individual cases, the control was felt to be too rough at low speeds or a delay was perceived in the vehicle's reaction to the control action on the lever.

Another aspect reported by the test managers and TD concerns the occasional occurrence of interference and visual artefacts in the video images. During the test drives, there were occasional brief image failures on individual cameras or flickering. In this specific case, none of these disturbances were associated with a faulty performance, but this should not lead to the conclusion that such an association is not possible. Visual artefacts in the video images were occasionally observed, particularly at night. These were, on the one hand, points of light that were presumably caused by reflections and, on the other hand, a colour rendering of lights that was not true to reality. Again, none of these phenomena were associated with a malfunction in the experiment. However, such effects potentially lead to the risk of confusion. Points of light could be mistaken for a dwarf signal, for example. A coloured appearance of white lights, especially in the colours red or green, leads to the risk of confusion with differently designed signals that indicate stop or go. A different appearance of de facto coloured lights can also lead to errors. In connection with image resolution, a visual clumping of individual lights was also reported, which appeared more pronounced at night.

It was not possible to check the lighting on the vehicle from the remote control station. Although it was possible for participants to request a different light switch setting from the driver on the vehicle at any time, some ROs reported that they would have been more flexible with the lights if they had been in control themselves.

Finally, design aspects should also be mentioned here that the ROs found helpful and positive in the completion of tasks compared to travelling on the locomotive. Many emphasized the depiction of the extended Bernese area (cf. 6.1.4) as a good support for tasks that required the processing of information from this area, such as approaching a rail vehicle or stopping in front of a buffer stop with pinpoint accuracy. The basic arrangement of the monitors and control elements was also positively mentioned. Among other things, the presence of a wide peripheral area on the monitors was found to be helpful. The basic consistency of the control elements and the connection between operating actions and vehicle behaviour with what is familiar was also perceived positively and should be retained in the design of the remote control station, especially when using drivers who are trained to drive on site.

7.1.3 How can the system be optimised for high suitability for shunting tasks?

The test clearly showed that, based on the available video images, manual remote shunting with the system was possible, but with limited efficiency. The extent to which this limitation is acceptable depends on the scenario in which the remote control is used. If efficiency is to be increased, a clear starting point from the tests is to provide the RO with visual information. Finer image resolution would improve the execution of tasks. At the same time, however, this also increases the demands on data transmission. The challenge for a sensible system design is to find a good compromise between information supply and data transfer requirements. Concepts in which the entire image is not permanently transmitted in high resolution, but only certain important areas, could be helpful here. In addition, there could be switchable cameras or cameras that can be aligned and focussed for the RO itself in order to make previously invisible areas visible or to focus areas. Applications from the remote tower dome can serve as inspiration for the more precise design of such systems (Van Schaik et al., 2022), but must be adapted for the railway context. The experiment has already shown that at least the view to the rear should also be available for the tested scenarios.

In a further developed remote control system, RO should have access to additional vehicle functions. This includes controlling the lighting on the vehicle. In addition, adaptive functions can be added to the lighting. Both measures will contribute to improved perception at the RO workplace in the dark.

An option that can be used as an alternative or supplement to increasing the resolution for a more usable system design is to support perception on the basis of information from other sensors. To improve visibility at night, NIR-sensitive cameras with additional illumination of the surroundings by NIR spotlights (Rivera-Velázquez et al., 2022) and thermal imaging cameras, whose image can be combined with that of other cameras to form a false colour image (Hwang et al. 2015), can be considered. With NIR systems, recognisability depends on the range of the NIR illuminator, the reflectivity of the objects (the range is shorter for black objects) and the image resolution of the sensor and lens (cited in Thomas-Friedrich et al., 2024). Compared to images from other sensors such as lidar or radar scans, human operators can interpret camera images more easily. Additional data processing steps are required to make the information from sensors that generate point clouds usable for RO: Based on usually fused sensor data, objects are automatically recognised and then highlighted in the displays for the RO or presented in other ways (Kyatsandra et al. 2022; Ristic-Durrant et al., 2022, von Einem et al., 2023). One advantage of such systems is that they can provide information as the basis for assistance functions, such as those suggested by the drivers in the test (obstacle detection, route detection, signal detection), even in poor visibility conditions; disadvantages are the high level of complexity and the high development costs.

The results also show that the perception of vehicle acceleration, deceleration and speed should be supported even more by the system design. A range of measures can be considered for this purpose. One idea, which was also supported by the drivers, is a more central arrangement of the speed display in the field of vision. In the specific case of the Aem 940, this could be achieved, for example, by superimposing the locomotive body in the central field of vision without superimposing ambient information (cf. Wickens, 2013). While the complete transmission of vestibular, tactile and proprioceptive feedback by a dynamic driving simulator at the remote control station seems impracticable, partial aspects of this perception could be implemented with reasonable effort. In the context of road vehicles, it has already been shown that additional haptic feedback in the form of seat vibration can be used to better estimate the speed in teleoperation than purely visual feedback. The combination with auditory feedback (motor noise) further improved speed perception (Chen, 2015). How such and similar measures can be adapted for the rail transport sector is an open research and development question. The same applies to supporting the perception of inclines and gradients and the unintended driving dynamics caused by them, which also appears to make sense. The basis for an assistance function geared towards this could be sensor data from the vehicle, which could be combined with map information about the terrain. On this basis, various representations of this information are conceivable, for example as a visual, auditory or haptic display. To date, there are no findings as to which type of information display best supports perception.

An advanced remote control system should support the RO more than the tested variant in determining the exact location of their own vehicle. In the best case scenario, a reliable display is available to the

RO, from which its own location can be taken together with the information required for orientation and communication with other personnel (e.g. station, track, neighbouring signals and points, orientation of the locomotive and current line of sight). The representation in the form of a digital map would be useful, but the pure information could also be used in combination with an additional track plan. If it is not possible to automatically communicate location information, the orientation of the RO can be at least partially facilitated by implementing the above-mentioned improvements to the representation of the surroundings, which increase the visibility of the surroundings and the recognisability of the information contained therein.

One finding of previous research is that ROs require feedback on the quality of the signal and the data connection (Masson et al., 2019). This applies in particular to latency. In the present system, a basic function for this was implemented in the form of a colour display on the left monitor of the bottom row. A bar at the top of this monitor changed colour from green to red when a set value was exceeded. To date, there are hardly any specifications for the railway sector with regard to limit values and their display. Masson et al. (2019) merely state that low latencies are necessary for the remote control of trains. DIN EN 16186-3 also provides possible starting points, which stipulates that the time between the receipt of new data and the updating of the displayed data must not exceed 200 ms when data is shown on the display of a full-track vehicle (DIN, 2022). Further findings on the effect of latency are available from other research areas such as the teleoperation of road vehicles (Bodell & Gulliksson, 2016; Diermeyer et al., 2011; Held & Durlach, 1991; Neumeier et al., 2019). At this point, it should be noted that a feedback signal for data transmission should be available in every remote train control system. In addition to pure feedback, other functions that make it easier for the RO to deal with latency are desirable. Various forms of predictive displays have been tested for teleoperation on the road (Chucholowski et al., 2013; Dybvik et al., 2020; Graf et al., 2020; cited in Thomas-Friedrich et al., 2024), which can serve as inspiration for further research and development in the rail transport sector.

The infrastructure and equipment at the shunting yard are currently tailored to the requirements of on-site vehicle control. One possible approach to improving perception at the remote control station is to adapt the design and use of relevant equipment to the requirements of remote control. Examples in the context of the test include the colour-coded, reflective or luminous design of restraining shoes, the stipulation that signals in a certain area of operation are always positioned to the left of the track or area-wide illumination of the area of operation.

In the prototype system tested, the switches for vehicle operation were implemented in the form of interactive buttons on a touch display. When designing switching elements in future systems for long-term use, care should be taken to ensure that the design protects against accidental incorrect operation of switches.

7.2 Scope and limitations of the study design

The test of a prototype system for remote locomotive control in the field has provided many insights into its suitability for shunting tasks and starting points for further development. For a meaningful interpretation and further use of the results, the investigated area of application and the associated limitations should be reflected at this point.

A system prototype for manual remote control was tested with an Aem-940 locomotive in shunting operations under day and night conditions at a Zurich shunting yard. The participants were SBB drivers with between one and 33 years of professional experience and came from the areas of passenger and freight transport and infrastructure. Interaction with the system took place over one or two appointments lasting a few hours. The tests took place in February and March under frost-free and predominantly dry ambient conditions (see also 6.1.4 – 6.2.1).

When comparing the remote control system with the reference system of driving on the locomotive, the different experience of the users with the respective system must be taken into account. Some of the perceived efficiency losses when working with the system can probably be compensated for by increased user experience. This relates, for example, to the handling of the levers, some of which differed from what the drivers were used to. In the case of other suspected causes of reduced execution efficiency in the specific system design, it seems less plausible that they can be fully compensated for by increased experience alone, such as the visual restrictions or the difficulties in perceiving speed, gradients and distances.

The performance of the system could not be observed under certain conditions, or only to a limited extent. These included snow and frost, but also heavy rain, fog or extreme heat. The latency and transmission quality were within a certain range in the test, which did not require any interruption, and could not be systematically varied any more than other system parameters.

The area of application investigated was the direct manual remote control of a single vehicle of a specific type at a single railway station. Therefore, no conclusions can be drawn about the possible effects of changing the vehicle or the operating environment within the remote control system. The same applies to overarching HTO design aspects such as the embedding of the workplace in the overall organisation with the questions of task distribution between different roles in the railway system and technical systems, quantitative task assessment, working time, communication requirements, necessary training and many more. The organisation of these aspects is the subject of further research and development.

7.3 Outlook: Need for further research

There are still many questions to be answered in the next stages of development at the various design levels with regard to humans, technology and organisation relating to remote train monitoring and control as an ATO fallback level or as a standard operating solution for defined application cases, some of which are outlined here in conclusion.

The study carried out focussed on the level of human-technology interaction in the direct manual remote control of a rail vehicle. As outlined above, research here must focus in particular on the development of solutions with which the required information can be transmitted as economically as possible between the place of use and the remote control station. An important starting point here is to investigate the required quality of perception for different scenarios. Exemplary research questions are: In which areas of the visual field is which resolution required to efficiently recognise relevant objects? What influence do other video quality parameters such as the refresh rate or the use of mono vs. stereo cameras have on performance and stress? Which areas should be permanently visible and which should be available on demand? At what resolution in which visual areas does the RO have an appropriate perception of speed? Which other feedback signals can best support the perception of speed, gradients and distances in a practicable way? How do latency parameters affect performance in different application scenarios? To what extent can greater image resolution be compensated for by assistance systems for the perception of objects, signals and driving routes? There are also questions that address the long-term use of the system, such as what increases in efficiency can be expected through repeated interaction with the system through learning. Research approaches in which various parameters can be systematically varied in an efficient manner, such as computer or simulator-based tests, are suitable for investigating this type of question.

Prospective concepts must also be developed and evaluated for the level of organisational design. The results have an impact on the requirements for the design of human-technology interaction. One example is the question of how the tasks of the dispatcher, remote train monitoring and remote train control are divided up. Are remote train monitoring and control separate areas of responsibility that are handled by different operators, or are monitoring and control carried out together at the same workstation, or which variant is better suited to which application scenarios? How are monitoring tasks divided between dispatchers and train teleoperators? How extensive are the respective areas of responsibility and how are they organised? What demands are placed on those involved as a result of different organisational variants, and how can the work be made attractive? What education and training is required to carry out the new tasks? In the interests of a favourable design, it is important to involve those involved in the development process from the outset. Participatory design methods are suitable for answering these questions by making the effects of the different concepts tangible.

Regardless of how the automation of rail transport is organised in the future, humans will remain an indispensable part of the system. Consideration of the overall system of human, technology and organisation is a prerequisite for the safety and efficiency of operations. The HTO analysis carried out has helped to specify the requirements for the manual remote control of rail vehicles and to develop solution concepts.

8 Funding Disclaimer

Funded by the European Union. Views and opinion expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them. The project R2DATO is supported by the Europe's Rail Joint Undertaking and its members.

9 References

Akerstedt, T. & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. The International Journal of Neuroscience, 52(1-2), 29–37. https://doi.org/10.3109/00207459008994241

Bodell, O., & Gulliksson, E. (2016). Teleoperation of autonomous vehicle with 360° camera feedback. Department of Signals and Systems, Chalmers University of Technology

Brandenburger, N., Dreßler, A. & Grippenkoven, J. (2023). DLR Workload Assessment Tool (DLR-WAT) - Official English Version. Verfügbar unter: https://elib.dlr.de/201450/

Brooke, J. (1996). A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (Hrsg.), Usability evaluation in industry (First edition, S. 189–194). Bristol, PA: Taylor and Francis.

Chen, T. (2015). Methods for improving the control of teleoperated vehicles. Dissertation, TU München, München.

Chucholowski, F., Buechner, S., Reichender, J., & Lienkamp, M. (2013). Prediction Methods for Teleoperated Road Vehicles. Paper presented at the Conference on Future Automotive Technology.

Diermeyer, F., Gnatzig, S., Chucholowski, F., Tang, T., & Lienkamp, M. (2011). Der Mensch als Sensor-Der Weg zum teleoperierten Fahren. Paper presented at the AAET-Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel

DIN Deutsches Institut für Normung e. V. (2018). DIN EN ISO 9241-11 Ergonomie der Mensch-System-Interaktion –Teil 11: Gebrauchstauglichkeit: Begriffe und Konzepte (ISO/DIS 9241-11:2018).

DIN Deutsches Institut für Normung e. V. (2019). DIN EN ISO 9241-210 Ergonomie der Mensch-System-Interaktion -Teil 210: Menschzentrierte Gestaltung interaktiver Systeme (ISO 9241-210:2019).

DIN Deutsches Institut für Normung e. V. (2022). DIN EN 16186-3:2022-05 Bahnanwendungen - Führerraum - Teil 3: Displaygestaltung für Vollbahnfahrzeuge; Deutsche Fassung EN 16186-3:2022

Dunn, N. & Williamson, A. (2012). Driving monotonous routes in a train simulator: the effect of task demand on driving performance and subjective experience. Ergonomics, 55(9), 997–1008. https://doi.org/10.1080/00140139.2012.691994

Dybvik, H., Løland, M., Gerstenberg, A., Slåttsveen, K., Steinert, . (2020). A low-cost predictive display for teleoperation: Investigating effects on human performance and workload. International Journal of Human-Computer Studies. doi: 10.1016/j.ijhcs.2020.102536.

Endsley, M. R. (1995). Measurement of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1), 65–84. https://doi.org/10.1518/001872095779049499

Ericsson, K. A. & Simon, H. A. (1999). Protocol analysis. Verbal reports as data (A Bradford book, Rev. ed., 3. print). Cambridge, Mass.: The MIT Press.

Gibson, J. J. (2015). The ecological approach to visual perception (Psychology Press classic editions, Classic edition). London, New York: Psychology Press Taylor & Francis Group. https://doi.org/10.4324/9781315740218

Goodwin, K. & Cooper, A. (2011). Designing for the Digital Age. How to Create Human-Centered Products and Services (1., Auflage). New York, NY: John Wiley & Sons.

Graf, G., Abdelrahman, Y., Xu, H., Abdrabou, Y., Schitz, D., Hußmann, H., Alt, F. (2020). The Predictive Corridor: A Virtual Augmented Driving Assistance System for Teleoperated Autonomous Vehicles. In International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments. https://doi.org/10.2312/egve.20201260

Grippenkoven, J., Rodd, J. & Brandenburger, N. (2018). DLR-WAT: Ein Instrument zur Untersuchung des optimalen Beanspruchungsniveaus in hochautomatisierten Mensch-Maschine-Systemen. In *AAET - Automatisiertes und vernetztes Fahren. Beiträge zum gleichnamigen 19. Braunschweiger Symposium am 14. und 15. März 2018, Stadthalle, Braunschweig.* Braunschweig: ITS mobility e.V.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload* (Advances in Psychology, Bd. 52, S. 139–183). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

Held, R., & Durlach, N. (1991). Telepresence, time delay and adaptation. Pictorial communication in virtual and real environments, 232-246.

Hwang, S., Park, J., Kim, N., Choi, Y., & So Kweon, I. (2015). Multispectral pedestrian detection: Benchmark dataset and baseline. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1037-1045).

ISO. (2019). Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems (ISO 9241-210:2019). International Organization for Standardization.

Kyatsandra, A. K., Saket, R. K., Kumar, S., Sarita, K., Vardhan, A. S. S. [Aanchal Singh S.] & Vardhan, A. S. S. [Akanksha Singh S.]. (2022). Development of TRINETRA: A Sensor Based Vision Enhancement System for Obstacle Detection on Railway Tracks. *IEEE Sensors Journal*, *22*(4), 3147–3156. https://doi.org/10.1109/JSEN.2021.3140032

Lee, D. N. (1980). The optic flow field: the foundation of vision. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 290(1038), 169–179. https://doi.org/10.1098/rstb.1980.0089

Masson, É., Richard, P., Garcia-Guillen, S., & Adell, G. M. (2019). TC-Rail: Railways remote driving. In Proceedings of the 12th World Congress on Railway Research, Tokyo, Japan (Vol. 28).

Neumeier, S., Wintersberger, P., Frison, A.-K., Becher, A., Facchi, C., & Riener, A. (2019). Teleoperation: The holy grail to solve problems of automated driving? Sure, but latency matters. Paper presented at the Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications

Norman, D. A., & Draper, S. W. (1986). User Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Erlbaum Associates.

Ristić-Durrant, D., Haseeb, M. A., Banić, M., Stamenković, D., Simonović, M. & Nikolić, D. (2022). SMART on-board multi-sensor obstacle detection system for improvement of rail transport safety. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *236*(6), 623–636. https://doi.org/10.1177/09544097211032738

Rivera Velázquez, J. M., Khoudour, L., Saint Pierre, G., Duthon, P., Liandrat, S., Bernardin, F., ... & Peleg, R. (2022). Analysis of thermal imaging performance under extreme foggy conditions: applications to autonomous driving. Journal of imaging, 8(11), 306.

Sauro, J. (2018). 5 Ways to Interpret a SUS Score. Verfügbar unter: https://measuringu.com/interpret-sus-score/

Selcon, S. J. & Taylor, R. (1990). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. Proc. AGARD, Situational Awareness Aerosp. Oper., 5/1–5/8.

Taylor, R. (1989). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. Proc. AGARD, Situational Awareness Aerosp. Operations., 3/1–3/17.

Thomas-Friedrich, B., Dreßler, A. & Leich, A. (2024). Teleoperation als Ergänzung für den automatisierten Bahnbetrieb (ATO) - Rahmenbedingungen und menschzentrierte Gestaltung. Zwischenbericht: Ergebnisse des dritten Arbeitspakets. DLR-interner Bericht; Veröffentlichung durch DZSF nach Projektabschluss 2025.

Thomas-Friedrich, B. & Grippenkoven, J. (2021). Prüfkriterien für den Nachweis einer integrierten Gestaltung von Mensch, Technik und Organisation. Abschlussbericht (Interner Projektbericht).

Van der Laan, J. D., Heino, A. & Waard, D. de. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research Part C: Emerging Technologies, 5(1), 1–10. https://doi.org/10.1016/S0968-090X(96)00025-3

Van Schaik, F. J., Roessingh, J. J. M., Bengtsson, J., Lindqvist, G., & Fält, K. (2022). The advanced remote tower system and its validation. In Virtual and Remote Control Tower: Research, Design, Development, Validation, and Implementation (pp. 429-444). Cham: Springer International Publishing.

Von Einem, C., Cramariuc, A. & Cadena, C., Siegwart, R. (2023). LROD: Long-range obstacle detection for railway driver assistance systems. Verfügbar unter: https://ethz.ch/content/dam/ethz/special-interest/mavt/csfm-dam/events/2023/symposium/posters/mi/voneinem-asl-lrod.pdf

Warm, J. S., Matthews, G. & Finomore, V. S. (2008). Vigilance, workload, and stress. In P. A. Hancock & J.L. Szalma (Eds.): Performance Under Stress (S. 115–141). Aldershot: Ashgate Publishing Limited.

Wickens, C. D. (2008). Multiple resources and mental workload. Human Factors: the Journal of the Human Factors and Ergonomics Society, 50(3), 449–455. https://doi.org/10.1518/001872008X288394

Wickens, C. D. (2013). Engineering psychology and human performance (Fourth edition). London, England, New York, New York: Routledge.

Wolfe, B., Dobres, J., Rosenholtz, R. & Reimer, B. (2017). More than the Useful Field: Considering peripheral vision in driving. Applied Ergonomics, 65, 316–325. https://doi.org/10.1016/j.apergo.2017.07.009

Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology, 18(5), 459–482. https://doi.org/10.1002/cne.920180503

Young, M. S., Stanton, N. A., & Walker, G. H. (2006). In loco intellegentia: human factors for the future European train driver. International journal of industrial and systems engineering, 1(4), 485-501.

10 Annexes

10.1 Overview: Test conditions

Table 10. Overview of all test sessions and their assignment to participants, block sequences and conditions

Meeting	TN-ID	Test time	Blocks	Sequence					Brightness	ILTIS
1	1	Morning	4	A	C	D	B		bright	Yes
2	2	Afternoon	3	A	C	B			bright	Yes
3	3	Morning	5	B	D	E	C	A	bright	no
4	4	Afternoon	3	B	D	E			bright	no
5	1	Evening	5	E	C	A	D	B	dark	no
6	3	Evening	5	A	C	B	D	E	dark	Yes
7	5	Morning	5	B	D	E	C	A	bright	Yes
8	6	Afternoon	3	E	C	A			bright	Yes
9	7	Morning	5	E	C	A	D	B	bright	no
10	8	Afternoon	3	A	C	B			bright	no
11	5	Evening	5	A	C	B	D	E	dark	no
12	7	Evening	5	B	D	E	C	A	dark	no
13	9	Morning	5	E	C	A	D	B	bright	Yes
14	10	Afternoon	4	B	D	E	C		bright	Yes
15	11	Morning	5	A	C	B	D	E	bright	no
16	12	Afternoon	4	E	C	A	D		bright	no
17	9	Evening	5	B	D	E	C	A	dark	no
18	11	Evening	5	E	A	D	B	C	dark	Yes
19	13	Morning	5	A	C	B	D	E	bright	Yes
20	14	Afternoon	4	A	C	B	D		bright	Yes
21	15	Morning	5	B	D	E	C	A	bright	no
22	16	Afternoon	4	B	D	E	C		bright	no
23	13	Evening	5	C	E	A	D	B	dark	no
24	15	Evening	5	C	A	B	D	E	dark	Yes
25	17	Morning	5	B	D	E	C	A	bright	Yes
26	18	Afternoon	4	E	C	A	D		bright	Yes
27	19	Morning	5	E	C	A	D	B	bright	no
28	20	Afternoon	4	A	C	B	D		bright	no
29	17	Evening	5	C	A	B	D	E	dark	no
30	19	Evening	5	C	B	D	E	A	dark	Yes
31	21	Morning	5	E	C	A	D	B	bright	Yes
32	22	Afternoon	4	B	D	E	C		bright	Yes
33	23	Morning	5	A	C	B	D	E	bright	no

34	24	Afternoon	4	E	C	A	D		bright	no
35	21	Evening	5	C	B	D	E	A	dark	no
36	23	Evening	5	C	E	A	D	B	dark	Yes

10.2 Overview: Perceived challenges in task execution and suggestions for improvement

Table 11. Overview of all categories of challenges and suggestions for improvement mentioned, sorted by task block and frequency of naming

	Block A - DS at peak - person on the track - de-energised section	Number of entries
Challenges	Visibility is worse	13
	I would otherwise have seen DS myself	3
	Chassis missing	3
	Bow lowering signal easier to recognise at night than during the day	3
	High visibility waistcoat easier to recognise at night than during the day	3
	[+] BR camera helpful	2
	Bow lowering signal	2
	Additional arrow only recognisable at a short distance	
	Braking distance estimation more difficult	2
	Tiring for the eyes	2
	Operation is different / system behaves differently - you have to get used to it	2
	Driven slower	1
	More control of adverts needed	1
	High visibility safety waistcoat	1
	Brake is sluggish	1
	Traction control is different	1

	[+] Less leverage	1
	Bearing lowering signal harder to recognise at night than during the day	1
	Camera position better than in real life	1
	brake with right lever instead of left	1
	Brake reacts more slowly	1
	Distance perception more difficult	1
	Hinge switch should be protected against accidental operation	1
Suggestions for improvement	higher video quality	5
	Make additional areas visible	3
	DAS Hindenis detection	1
	Position adverts differently	1
	Fog lights	1
	Adjust lever control dynamics	1
	Windows with fresh air	1
	Display on curved screen	1
	Presentation of your own position at the station	1
	Redesign of signals and equipment	1

Block B - DS in front of tip - Inhibition shoe - Stop signal board		Number of entries
Challenges	Visibility is worse	17
	Chassis missing	6

	Distance perception more difficult	4
	Stop signal board can only be recognised late	3
	Speed perception more difficult	3
	Recognisable change of course	2
	Operation is different / system behaves differently - you have to get used to it	2
	I would have been more flexible with the lighting on the locomotive	1
	[+] BR camera helpful	1
	Orientation more difficult at the beginning	1
	Image quality is sufficient	1
	More control of adverts needed	1
	Lever control dynamics / traction control is different	1
	brake with right lever instead of left	1
	More tiring for eyes at night	1
	[+] Handling as usual	1
Suggestions for improvement	DAS obstacle detection	9
	higher video quality	6
	Customise displays	3
	Variable lighting on the vehicle can be controlled by the RO	2
	Zoom function	1
	Additional sensors	1
	DAS signal detection	1

	Adapt switch behaviour	1
	Redesign of signals and equipment	1
	Support movement perception more strongly	1
	AP how to organise a teacher's stand	1
	Lighting on site	1

Block C - in front of unoccupied points - no special features - Buffer stop		Number of entries
Challenges	Signal position not visible (behind vehicle head)	12
	Positioning more difficult	8
	Chassis missing	7
	Distance perception more difficult	7
	Speed perception more difficult	6
	[+] BR camera helpful	5
	Visibility is worse	4
	Do not brake: Brake; no possibility to communicate with people on the track	3
	More control of adverts needed	2
	Lever control dynamics / traction control is different	2
	Operation is different / system behaves differently - you have to get used to it	2
	You are protected from bad weather	1
	Recognisable change of course	1
	Signal position more difficult to recognise	1

	Latency	1
	Brake is sluggish	1
	brake with right lever instead of left	1
	Braking distance estimation more difficult	1
	Tiring for the eyes	1
	Maintaining focus difficult	1
Suggestions for improvement	Make additional areas visible	9
	Presentation of your own position at the station	5
	Additional sensors	4
	DAS distance estimation	4
	Design AP exactly like a teacher's stand	4
	higher video quality	3
	DAS obstacle detection	2
	Customise displays	1
	Front fog lights / cornering lights	1
	Adjust lever control dynamics	1
	Variable lighting on the vehicle can be controlled by the RO	1
	Length scale on the rail	1
	Display of LEA information on MDU or DDU	1
	Vehicle should have a parking brake	1
	Audio transmission	1

	Block D - DS before tip - no special features - Approach to RV	Number of entries
Challenges	Lever control dynamics / traction control is different	7
	Speed perception more difficult	6
	Chassis missing	5
	Distance perception more difficult	5
	Slope/ gradient more difficult to perceive	3
	No external check as to whether the vehicle' s wagons are braked	2
	Frequent changes of view between video and speed display necessary	2
	[+] BR camera helpful	2
	Visibility is worse	2
	More control of adverts needed	2
	Braking distance estimation more difficult	2
	Operation is different / system behaves differently - you have to get used to it	2
	Audio info missing	1
	Recognisable change of course	1
	Latency	1
	Brake reacts slower / weaker	1
Suggestions for	DAS distance estimation	7
	Enable v-target setting to 0 km/h (results in coupling drive mode with 1 km/h) with monitoring of the actual v	4
	higher video quality	2

	Adjust lever control dynamics	2
	Support movement perception more strongly	2
	Audio transmission	2
	Zoom function	1
	Customise displays	1
	Adapt switch behaviour	1
	Customise vehicle design	1
	Make additional areas visible	1

Block E - DS in front of tip - DS falls to stop - stop in front of DS		Number of entries
Challenges	Visibility is worse	9
	Stop pointing DS recognisable	6
	Setting the course too late recognisable	3
	Chassis missing	3
	Changing view between video above and below necessary; more strenuous	2
	Audio info missing	1
	Latency	1
	Positioning more difficult	1
	Driven slower	1
	Speed perception more difficult	1
	Brake reacts slower / weaker	1

	Operation is different / system behaves differently - you have to get used to it	1
Suggestions for improvement	DAS signal detection	5
	higher video quality	2
	Support movement perception more strongly	2
	Variable lighting on the vehicle can be controlled by the RO	2
	Replay function (to be able to watch previous sequences in the video again)	1
	DAS route detection	1
	DAS obstacle detection	1
	Make additional areas visible	1
	Audio transmission	1