"The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723205"

Deliverable D2.5

Human factor methodological framework

Due date of deliverable: 29/02/2020
Actual submission date: 28/02/2020

Authors:
Grigore M. Havârneanu, UIC
Anne Silla & Esko Lehtonen, VTT
Eduardo Prieto & Aida Herranz, FFE
Annika Dreßler, DLR
Annie Kortsari, Neofytos Boufidis, Josep Maria Salanova Grau &
Georgia Aifadopoulou, CERTH-HIT

Project details

Project acronym	SAFER-LC
Project full title	SAFER Level Crossing by integrating and optimizing road-rail infrastructure management and design
Grant Agreement no.	723205
Call ID and Topic	H2020-MG-2016-2017, Topic MG-3.4-2016
Project Timeframe	01/05/2017–30/04/2020
Duration	36 Months
Coordinator	UIC – Marie-Hélène Bonneau (bonneau@uic.org)

Document details

Title	Human Factor methodological framework
Workpackage	WP2
Date of the document	28/02/2020
Version of the document	v04
Responsible partner	UIC
Reviewing partner	VTT, DLR, FFE, CERTH
Status of the document	Final
Dissemination level	Public

Document history:

Revision	Date	Description
01	07/11/2019	First draft (with preliminary inputs from task partners)
02	11/02/2020	Second draft (integrated version)
03	27/02/2020	Third draft (after partner review)
04	28/02/2020	Ethical and final review

Consortium - List of partners

Partner No	Short name	Name	Country
1	UIC	International Union of Railways	France
2	VTT	Teknologian tutkimuskeskus VTT Oy	Finland
3	NTNU	Norwegian University of Science and Technology	Norway
4	IFSTTAR	French institute of science and technology for transport, development and networks	France
5	FFE	Fundación Ferrocarriles Españoles	Spain
6	CERTH-HIT	Centre for Research and Technology Hellas - Hellenic Institute of Transport	Greece
7	TRAINOSE	Trainose Transport – Passenger and Freight Transportation Services SA	Greece
8	INTADER	Intermodal Transportation and Logistics Research Association	Turkey
9	CEREMA	Centre for Studies and Expertise on Risks, Environment, Mobility, and Urban and Country planning	France
10	GLS	Geoloc Systems	France
11	RWTH	Rheinisch-Westfaelische Technische Hochschule Aachen University	Germany
12	UNIROMA3	University of Roma Tre	Italy
13	COMM	Commsignia Ltd	Hungary
14	IRU	International Road Transport Union - Projects ASBL	Belgium
15	SNCF	SNCF	France
16	DLR	German Aerospace Center	Germany
17	UTBM	University of Technology of Belfort-Montbéliard	France

Executive summary

This deliverable presents the revised version of the Human Factors (HF) methodological framework which has been developed in the SAFER-LC project as part of Work Package 2 (WP2).

The objective of Task 2.2 of WP2 is to develop a Human Factors methodological framework to evaluate the effectiveness of selected safety measures in terms of making level crossings (LCs) more self-explaining and forgiving, and hence increasing their safety. The methodological framework includes a practical Human Factors Assessment Tool (HFAT) accompanied by an implementation guide which presents how the HFAT can be used in a real case study.

The purpose of this deliverable is to summarise the theoretical background of the Human Factors methodological framework and the development process of the first version of the Human Factors Assessment Tool. In addition, this deliverable aims to explain how the HFAT was adjusted and updated in the second part of the project based on feedback obtained during the HFAT testing phase in four of the project's pilot tests, covering 14 measures.

The overall objectives and structure of this deliverable is described in Chapter 1. Chapter 2 reviews and summarises the most important theoretical aspects of the Human Factors methodological framework in the LC context. The framework was developed in line with the principles *self-explaining* and *forgiving* infrastructure and by considering LCs as socio-technical systems, where individual road users and the technical infrastructure interact. Models on human information processing and human behaviour in terms of *errors* and *violations* at LCs have also been considered. These theoretical aspects represent the theoretical backbone of the HFAT, and were presented in detail in deliverable D2.2 (Havârneanu et al., 2018).

Further, Chapter 3 shows how the HFAT was applied in the SAFER-LC pilot tests and presents the feedback received from the pilot test leaders. The two-step evaluation of the HFAT by the pilot test leaders was a useful and productive exercise. It allowed collecting valuable inputs, suggestions and ideas on how to improve specific parts of the tool. While most of the evaluation feedback was taken into account during the HFAT revision process, not all received suggestions could be implemented within the SAFER-LC timeframe and resources. Other suggestions were subject to group discussion during the project meetings and were implemented only partially, following the collective decision.

Chapter 4 explains the differences between the first version of the tool and the revised version. Based on the received feedback, changes concerned only the classification criteria (orange form) and the criteria to assess the behavioural safety effects (green forms). Major changes involved the revision of effect mechanism list in the classification criteria table and the regrouping of areas of psychological function in assessment of behavioural safety effects.

Chapter 5 provides an overall discussion of the HFAT, its strengths and limitations, its current utility as a stand-alone methodology, and possible directions in its further development. For example, the HFAT could be used in the future as a checklist to support the consideration of human factors perspective in the evaluation of LC safety measures. The HFAT will also be included in the SAFER-LC toolbox, accessible through a user-friendly interface.

Table of content

Abbreviations

Short name	Meaning
CBA	Cost-Benefit Analysis
D	Deliverable
DSS	Decision Support System
HF	Human Factor(s)
HFAT	Human Factors Assessment Tool
KPI	Key Performance Indicator
LC	Level Crossing
MRU	Motorised Road Users
TTC	Time to collision
VRU	Vulnerable Road Users
WP	Work Package

Definitions of main concepts

Concept	Definition
Human Factors	The application of psychological and physiological principles to the design of products, processes, and systems
Passive LC	An unmanned level crossing that has no crossing barriers, gates or road traffic signals. It is typically protected by road signs such as the crossbuck and sometimes additional 'Give Way' signs on each road approach.
Active LC	A level crossing which is equipped with an active protection system such as automatic half-barrier or full barrier, warning lights, or sound

1. INTRODUCTION

1.1. Objectives of SAFER-LC project

The SAFER-LC project (*Safer level crossing by integrating and optimizing road-rail infrastructure management and design*) aims to improve safety of level crossings (LCs) by minimising the risk of LC accidents. This is to be done by developing a fully integrated cross-modal set of innovative solutions and tools for the proactive management of LC safety and by developing alternatives for the future design of LC infrastructure.

The solutions and tools developed in the SAFER-LC project will enable road and rail stakeholders to find more effective ways to: (1) detect potentially dangerous situations leading to collisions at LCs, (2) prevent incidents by innovative user-centred design, and (3) mitigate the consequences of disruptions due to accidents or other critical events. The main output of the SAFER-LC project is a toolbox which will be accessible through a user-friendly interface which will integrate the project's practical results, tools and recommendations to help both rail and road stakeholders to improve safety at LCs.

The project focuses both on technical solutions, such as smart detection systems and advanced infrastructure-to-vehicle communication systems, and on human processes to adapt infrastructure designs to road user needs. Furthermore, the aim is to enhance coordination and cooperation between different stakeholders from different land transportation modes. The challenge is also to demonstrate the acceptance of the proposed solutions by both rail and road users and to implement the solutions cost-efficiently.

1.2. Purpose of this deliverable

Within the project, the objective of Work Package 2 (WP2) was to enhance the safety performance of LC infrastructures from a Human Factor (HF) perspective, making them more self-explaining and forgiving. More specifically, the objective of Task 2.2 of WP2 was to develop a HF methodological framework that evaluates the effectiveness of selected safety measures in terms of making LCs more self-explaining and forgiving and increasing their safety. The methodological framework includes a practical Human Factors Assessment Tool (HFAT) accompanied by an implementation guide which shows how the HFAT can be used in a real case study.

The main purpose of this deliverable is to present the improvement and refinement of the HFAT developed earlier in this project (Havârneanu et al., 2018) that was done using the results from the SAFER-LC pilot tests. More specifically, this deliverable explains how the HFAT was adjusted and updated based on feedback obtained during the HFAT testing phase in four SAFER-LC pilot tests. In order to do this, this deliverable also summarises the theoretical background of the HF methodological framework and the development process of the first version of the HFAT.

1.3. Structure of the document

This deliverable represents the final version of the framework that has been tested during the SAFER-LC pilot tests and refined based on the received feedback. This deliverable consists of the following chapters:

- Chapter 1: Introduction. Describes the overall objectives and structure of this deliverable.
- Chapter 2: Development of the Human Factors methodological framework. Reviews and summarises the most important theoretical aspects of the HF methodological framework in the LC context which represents the theoretical backbone of the HFAT (presented in detail in Havârneanu et al., 2018).
- Chapter 3: Piloting the Human Factors Assessment Tool. Presents how the HFAT was applied in the SAFER-LC pilot tests and the feedback received from the pilot test leaders.
- Chapter 4: Further Development of the Human Factors Assessment Tool. Highlights the differences between the first version of the tool and the revised version.
- Chapter 5: Discussion and Conclusions. Provides an overall discussion of the HFAT, its strengths and limitations, its current utility as a stand-alone methodology, and possible directions in its further development.

The Annexes illustrate the revised checklists of the HFAT and examples on how to apply the tool in practice.

2. DEVELOPMENT OF HF METHODOLOGICAL FRAMEWORK

This chapter presents an overview of the sets of criteria and indicators selected for the SAFER-LC HF methodological framework to evaluate LC safety measures from a HF perspective. This framework was built based on a combined methodology covering:

- Review and summary of important HF and psychological models which provide theoretical foundations for the HF methodological framework in the LC context (Subchapter 2.1);
- Literature review conducted as part of Task 2.1 with the focus on identifying key safety indicators concerning human errors and violations at LCs (Subchapter 2.2);
- Review of relevant past evaluation studies conducted in road and rail contexts to collect ideas on classification and evaluation criteria, behavioural safety indicators, and structure of the framework (Subchapter 2.3);
- Panel discussions between WP2 partners to formulate the HF assessment framework based on the information collected and analysed in the above steps (Subchapter 2.4).

The developed framework represents the theoretical backbone of the HFAT. The first version of the assessment tool was reported in the intermediate deliverable D2.2 (Havârneanu et al., 2018). The revised version will be presented and discussed later in this deliverable.

2.1. Theoretical background

The HF methodological framework developed in the SAFER-LC project is based on two principles: to make LCs *self-explaining* and *forgiving*. These two concepts are adapted from the field of road safety, where these principles have been previously presented and studied.

A *self-explaining* road is designed and built to create correct expectations and hence to guide road users in behaving safely in traffic (Bekiaris & Gaitanidou, 2011; SafetyCube Glossary, 2008). Correct expectations reduce the probability of driving errors and increase comfort. The main tools to make a road self-explaining are the layout, design and signage. Expectations are always based on road users' previous experiences, and therefore, consistency within the road infrastructure is also important. Motorways are a good example of self-explaining roads, because drivers usually know very well what to expect and how to behave there. On the other hand, motorway ramps are not always self-explaining, as demonstrated by drivers who end up driving on an off-ramp. A *forgiving* road is designed and built to minimize the consequences of drivers' errors – either by allowing recovery from an error or by mitigating the consequences of a crash (SafetyCube Glossary, 2008). A roundabout is a good example of forgiving road design. The design reduces the possibilities to make an error in visual search by limiting the directions where other vehicles are coming. In case of an error, crashes occur with lower driving speeds than in an intersection.

To make LCs self-explaining and forgiving, we need to consider LCs as socio-technical systems, where individual road users and the technical infrastructure interact (Read et al., 2013). Therefore, safety improvements require considering both psychological and behavioural aspects in road users as well as the design and functioning of the LC and, in addition, the interaction between the two. In the following, we discuss how to make LCs more self-explaining by evoking the right expectations for the road users, and how to design LCs to be more forgiving if an error occurs.

A classical information processing approach to analyse the human behaviour in a LC context, is to conceptualise human behaviour in a linear way: the road user needs to *perceive* the LC, allocate *attention* to the relevant features, *interpret* the perceived information, then *decide* how to *act* and finally *execute* the chosen behaviour (Wogalter, 2006; Grippenkoven, 2017). Such a linear approach provides useful ways to categorize errors by their mechanisms as well as to suggest ways to address those (Rasmussen, 1986; Reason, 1990; Grippenkoven et al., 2012; Laapotti, 2016). While easy to understand and use, conceptualisation of human behaviour as a linear process from perception to action does not showcase the importance of expectations for human behaviour. Instead of a linear process, the human cognition and action should be rather thought of as a cycle (Neisser, 1976): Expectations guide perception and the interpretation of the perceived information, the perceptions are interpreted relative to the expectations, and new expectations are formed which further guide perception. A more recent predictive processing viewpoint of human cognition has provided evidence on how actions are also performed in a way to match expectations and that the key information which the human brain processes is not necessarily the perception itself, but the mismatch between the expected and the perceived information (Clark, 2013; Engström et al., 2017).

Expectations come from experience. Thus, LCs should be consistently equipped with signs and designed to confirm that the road users can interpret that they are approaching a LC. The provided warnings should also be related to actual hazards. If a warning is too frequent without a materialized hazard, road users learn that there is no need to expect a hazard. This is a fundamental problem in passive LCs with low train traffic, where road users may stop looking for trains due to the low expectation of train arrival. A way to counteract this issue would be either to provide a clear and reliable *warning when a train is coming*, or to modify the technical system to guarantee that the hazard related information will be picked up by the bottom-up processing of information in perception. For example, if the driver *slows down* before a LC, he/she may have enough time to notice an approaching train even when not expected. Also, *making a train visually more salient* e.g. by using a flashing lights in the front of the train, should make it easier to detect a train even when not expecting it (Wickens & McCarley, 2008b).

Preventing distraction also helps in creating the right expectations. If a road user is distracted, he/she is less likely to engage in visual search of relevant cues of the approaching LC (or train), thus relying only on the bottom-up perception of them (Wickens & McCarley, 2008a). Being distracted with a visual task such as manual operation of a mobile phone, may also impair the use of peripheral visual cues by increasing the eccentricity between the direction of the gaze and the relevant visual cues at the LC, for example when the driver is looking down (Wickens & McCarley, 2008b). One way to counteract distraction would be thus to provide multimodal warnings for an approaching LC, e.g. via rumble strips.

Driving errors at LCs can occur due to wrong expectations, and consequently not looking for the relevant cues, or incorrectly interpreting the meaning of detected cues (Grippenkoven et al., 2012; Laapotti, 2016). Expectations also play an important role in violations, for example when driving around barriers in an active LC. In such a situation, drivers most likely have the expectation – and are very certain about with it – that the train is not yet arriving or that there is no other train coming after the first one has passed.

Driver behaviour is also influenced by other motives than safety (Näätänen & Summala 1976, Summala 1988). In some cases violations can be based on deliberate risk taking due to a personality trait of sensation-seeking (Ulleberg & Rundmo, 2003), or being in a hurry (Summala, 1988; Block & Zakay, 1996). A safe behaviour can incur some subjective cost, such as time loss, which the road user aims to balance by choosing a riskier option (Maddux & Rogers, 1983).

Making sure that the drivers have time to spot the cues related to an approaching train even when not expecting it, is also an example of 'forgiving' design. The possibilities to reduce the consequences of errors in a LC context are more limited than in road traffic. If a crash with a train occurs, in most cases the massive kinetic energy of the train causes severe consequences to the involved road users. Another example of forgiving design can be demonstrated considering active LCs with barriers. If a driver makes an error and gets stuck between the barriers for example due to traffic congestion, a design which leaves space between the barriers or *allows escaping between barriers* enables a recovery from an error without severe consequences. Another example of forgiving design in LCs is installing *remote monitoring systems* at LC, which can stop the train traffic or alert the train operator in case there are any obstacles like a stopped vehicle on the rails.

Table 1 provides a short description of theory-derived indicators for use in the SAFER-LC HF methodological framework.

Table 1. *Theory-derived indicators for the evaluation of level crossing safety measures, taking into account the road and rail users' perspective (adapted from D2.2, Havârneanu et al., 2018).*

Indicator	Definition	Examples and possible quantification
Impact on safe behaviours	Positive behavioural adaptation when approaching a LC	Speed reduction (-km/h) Looking left and right (yes/no, how often) Timing of these reactions (seconds before crossing) Speed choices in relation to the time that would theoretically be needed to stop in front of the rails if necessary
Impact on unsafe behaviours (involuntary)	Positive or negative effect on the errors committed by road users or rail users	Type of error (e.g. perception, memory etc.) Number of errors
Impact on unsafe behaviours (voluntary)	Positive or negative effect on the risky behaviours and violations committed by road users at LC (mostly at active LCs)	Type of violation (e.g. zig-zagging) Number of violations
Impact on the user's needs / motivations	How the measure integrates the needs of different road user categories	Short waiting time at LCs Time pressure
Impact on user's habits	How the measure is able to break the unsafe routines of frequent LCs users	Assuming they know the trains timetable at a specific LC (level of confidence)
Impact on VRUs	How the measure is adjusted to the vulnerability of road users such as pedestrians and cyclists	Type of VRUs (e.g. people with hearing disability)
Level of self-explaining nature	Level of implicit understanding of the measure by the end-user (i.e. easy to perceive and understand)	Possibility of language barriers in understanding signage Easily understood by children, elderly, people not familiar with technological measures

2.2. Information drawn from earlier phases of SAFER-LC project

Deliverable D2.1 (SAFER-LC consortium, 2018a) generated a knowledge base drawing on existing data sources and analytical tools with a view to enhancing the safety performance of LC infrastructures from a HF perspective. The main outcome of the task was the identification of key safety indicators concerning human errors and violations at LCs based on a review of relevant HF literature and a process of individual expert evaluation. From this process, a total of thirty indicators were identified and grouped under seven categories (see Table 2).

Table 2. *Key safety indicators concerning human errors and violations at LCs identified in D2.1 (SAFER-LC consortium, 2018).*

Indicator category	Indicators
Indicators related to personal conditions	Gender; age; disability; substance use.
Indicators related to distraction and inattention	Tiredness; stimuli overload; external distraction; internal distraction; distraction in general.
Indicators related to conspicuity of crossings and trains	Conspicuity; visual contrast; crossing angle; sight distances; signs.
Indicators related to lack of knowledge	Traffic rules; signalling; correct action; general knowledge of LCs.
Indicators related to inaccurate risk perception	Perception of risk in general; familiarity with place; frequent user; perception of train speed and distance.
Indicators related to deliberate risk-taking behaviour	Frustration and impatience; risk-seeking personality; low costs of fines; signal unreliability; suicide.
Indicators related to information about the context	Time of day; weather conditions; infrastructure layout; LC setting.

Each of these indicator categories and their validity for inclusion in the framework were further discussed in-depth in face-to-face group consultation with WP2 partners. During the discussion it was agreed that focus should be placed on those factors that are feasible to be measured, detected and controlled through external intervention (e.g. LC design). For example, whilst internal distraction was identified to affect safety, it is a difficult safety factor to detect and influence through a countermeasure. In this way, whilst the distinction between the different causes of distraction is important from a theoretical perspective, it was decided not to include the specific causes in the final checklist of criteria, which will focus only on external and visual distraction.

Based on this group work, several indicator categories listed in Table 2 above were regrouped, renamed and included as part of two broader sets of criteria relevant for the HF methodological framework: 'Classification criteria' and 'Criteria to assess the behavioural safety effects'. The renaming concerned, for example, the category of personal conditions (e.g. gender, age, disability etc.) which was renamed as 'Socio-demographical factors' and the category of 'Distraction and inattention' which was changed to 'Cognitive factors'. In addition, strong links were identified between knowledge and attention, whereby knowledge of traffic rules and the correct action can lead to better attention at LCs, which is now reflected within the 'Rule Knowledge' factor.

2.3. **Information drawn from related research projects**

The SAFER-LC Human Factors methodological framework also builds on indicators adapted from relevant past evaluation studies as transferrable lessons, such as classification and evaluation criteria, behavioural safety indicators, or ideas on how to structure and organize the framework based on other assessment approaches used in road or railway safety contexts.

Specifically, the main studies considered during the development of HF assessment tool were:

i) EU project RESTRAIL (*REduction of Suicide and Trespass on RAILway property*), which used 14 criteria to assess the most cost-effective measures to prevent railway suicides and trespassing (Ryan et al., 2018). These criteria were derived from those used in previous EU research (e.g. SUPREME project; Elvik, 2006), but have been adapted for use in a railway context.

ii) The assessment methodology developed based on the RESTRAIL project to assess measures aiming to improve the safety of LCs in Finland (Silla et al., 2015). Some criteria from the RESTRAIL project were slightly modified to fit better into the LC context and some additional (new) variables were created. Examples of these 'new' variables are 'Effect mechanism', 'Feasibility to different LCs', and 'Circumstances in which the measure is the most effective'.

iii) The ex-ante assessment method of Kulmala (2010) which is focused on road transportation and is targeted to assess the traffic safety impacts of ITS for cars, based on literature review and expert assessment. This method has been applied in several EU projects (see e.g. Wilmink et al., 2008; Scholliers et al., 2007; Kulmala et al., 2008; Wimmershoff et al., 2011; Fuerstenberg & Boehning, 2012; Innamaa et al., 2014; Silla et al., 2017). The most relevant criteria from Kulmala (2010) for the SAFER-LC HF methodological framework are the direct (short-term) and indirect (long-term) modification of road user behaviour.

iv) The method to assess the technical functioning, reliability, dependability, socio-economic benefits and costs of in-vehicle warning system for railway LCs as well as user experience, potential other applications areas and business models related to the in-vehicle warning system (Öörni et al., 2011).

Many of the criteria used in the above studies do not concern Human Factors – some criteria are more focussed e.g. on cost-benefit analysis (CBA) or socio-economic assessment. However, several of the criteria could be further developed and adapted in the SAFER-LC methodological framework and they provided a foundation for the classification of the relevant assessment criteria in the SAFER-LC HF methodological framework. Specifically, several of the 'Classification criteria' which are used in our framework are adopted from Ryan et al. (2018) and Silla et al. (2015), while the 'Criteria to assess behavioural safety effects' were motivated by Kulmala (2010). In addition, 'Criteria to assess user experience and social perception' were inspired by Ryan et al. (2018), Silla et al. (2015) and Öörni et al. (2011).

Table 3 provides a short description of these criteria adapted from previous studies for the purposes of the SAFER-LC HF methodological framework and their relevance in the LC context. The following list includes a criterion named as 'Reliability of the system' derived from the study of Öörni et al. (2011), which among other things assessed the technical functioning, reliability and dependability of an in-vehicle warning system for railway LCs.

Table 3. *Overview of the assessment criteria adapted from previous studies (as per Havârneanu et al., 2018).*

Category	Criterion	Definition	Source
Classification criteria	Effect mechanism	Specifies the type of effect mechanism (impact) expected with the intervention	Ryan et al., 2018; Silla et al., 2015
	Feasibility for different LCs	Specifies the types of LCs that the measure applies to	Silla et al., 2015
	Target of safety effects	Specifies the categories of users who are targeted by the measure	Ryan et al., 2018; Silla et al., 2015
	Circumstances where the measure is most effective	Specifies the circumstances where the measure is most effective or when it becomes ineffective	Silla et al., 2015
Criteria to assess behavioural safety effects	Short-term effect on road user behaviour	Describes the direct effects of the implemented safety measure on road user behaviour based on the strategic, tactical and operational level of behaviour	Kulmala, 2010
	Long-term effect on road user behaviour	Describes the indirect effects of the implemented safety measure on road user behaviour in the longer term. Long-term behavioural adaptation will often not appear immediately after a change but may show up later and is very hard to predict	Kulmala, 2010
Criteria to assess user experience and social perception	Acceptance (LC users, railway staff, people living nearby etc.)	Provides an estimate of how well the measure is accepted by the public and relevant stakeholders	Ryan et al., 2018; Silla et al., 2015
	Reliability of the system	Estimates if the users trust the system and how they know that it is fail-safe (i.e. the users are aware of the possible malfunction of the system)	Öörni et al., 2011
	Integration with road/railway environment, other safety measures	Describes how the measure is integrated with the road/rail environment and other possible preventative measures or interventions	Ryan et al., 2018; Silla et al., 2015

2.4. Criteria selected for the Human Factor methodological framework

The framework consists of three sets of criteria which are illustrated with different colours: *Classification criteria* (orange) as well as two sets of assessment criteria (*Criteria to assess the behavioural safety effects* – green –, and *Criteria to assess user experience and social perception* – blue) (Figure 1). Each of these categories is based on the sets of factors and indicators which represent the backbone of the HF Assessment Tool. Each criterion can be further broken down into a set of more specific and measurable indicators.

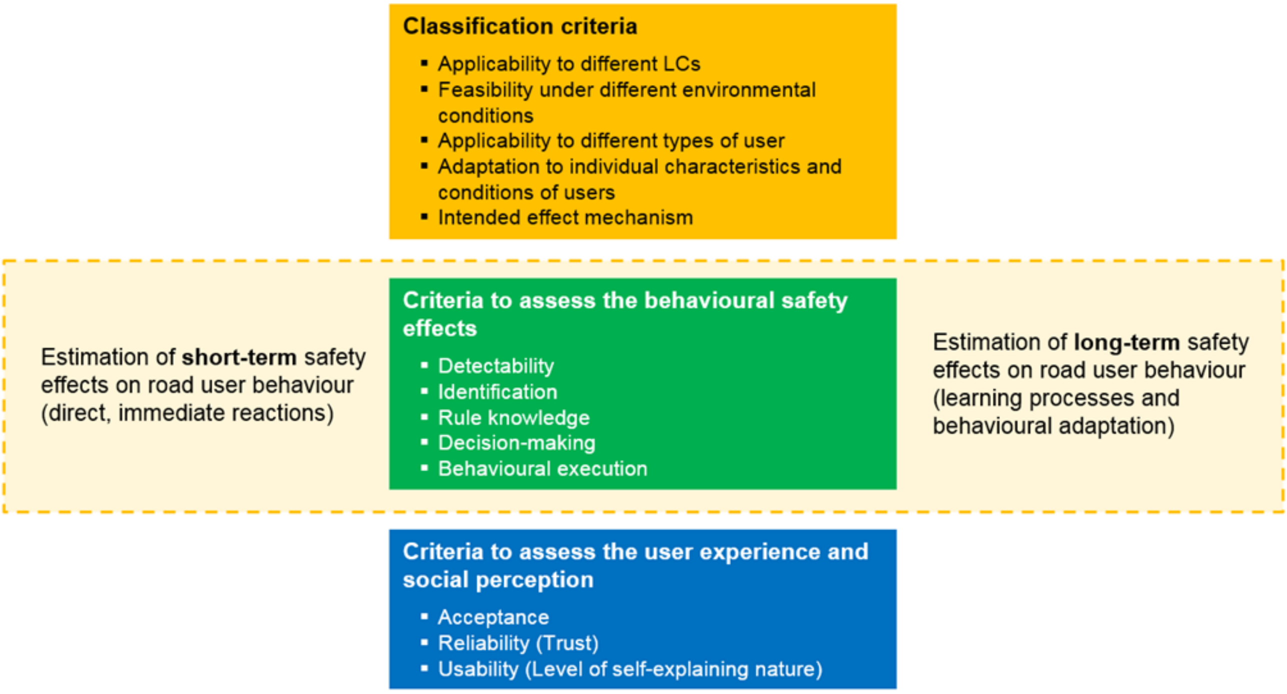

Figure 1. *The SAFER-LC HF methodological framework: Overview of the sets of classification and assessment criteria selected for the HF assessment tool.*

The upmost (orange) box of the assessment tool, *Classification criteria*, provides a description of the measure under assessment. It specifies the integration of the measure with different LCs and environmental conditions as well as its applicability to different LC user types and characteristics. This set of criteria also classifies the intended effect mechanism via which the measure is expected to affect road and railway safety (Table 4). These criteria are qualitative in nature and are used to define the context and environment in which the safety measure is expected to be effective. For example, if the safety measure is only installed at passive LCs and is targeted to improve the safety of children, the group of targeted LC accidents is rather limited and thus no high effects on Europe-wide LC safety performance can be expected, even though the effectiveness of that specific measure could be estimated as high.

Table 4. *Classification criteria.*

Factors	Descriptions	Indicators
Applicability to different LCs	Specifies the types and characteristics of LCs where the measure can be implemented	Types of LC ▪ Passive LCs without any warning devices ▪ Active (manual) ▪ Active LCs with barriers (half barriers, full barriers, skirts for pedestrians) ▪ Active LCs with light and sound warning ▪ Active LCs with other warning devices ▪ Active LCs with traffic lights, LCs with low vehicle traffic Characteristics of LC ▪ LCs with high / low vehicle traffic, paved road, gravel road, availability of electricity, sharp / wide crossing angle ▪ Other
Feasibility under different environmental conditions	Specifies the environmental circumstances in which the measure aims to be most effective and which may affect the perception or the behavioural adaptation of road users	▪ Time of the day (Daylight / Darkness / Dusk / Dawn) ▪ Peak traffic hours ▪ Weather conditions (Rain / Snowfall / Slipperiness / Fog / Bright sunshine, glare) ▪ Setting of the LC (urban/rural)
Applicability to different types of user	Specifies the categories of LC users who are targeted by the measure	▪ All road users ▪ MRU (car, motorbike etc.) ▪ VRU (cyclist, pedestrian etc.)
Adaptation to individual characteristics and conditions of users	Specifies if the measure can be targeted at individual characteristics and conditions of the user (e.g. socio-demographic characteristics, personal conditions, relevant individual traits)	▪ Gender ▪ Age (all ages, children, elderly etc.) ▪ Disability ▪ Under influence (e.g. alcohol, drugs, medication) ▪ Under skill impairing states (e.g. fatigue, stress) ▪ Risk-seeking personality ▪ Cultural/linguistic background (including e.g. different language needs)
Intended effect mechanism	Specifies the mechanism via which the measure is expected to have an effect on safety	▪ Improves the conspicuity of train ▪ Improves the conspicuity of LC ▪ Controls the access to LC ▪ Reduces the approach speeds of vehicles ▪ Increases the awareness of correct behaviour and dangerousness of LC ▪ Improves the physical environment of LC ▪ Improves the possibilities of vulnerable road users to cross LC safely ▪ Provides up-to-date information about the status of LC ▪ Supports the LC safety actions

In addition, the information gathered on the classification criteria can support road and railway stakeholders on deciding the locations where the specific safety measure could be implemented. For example, these criteria describe the types of LCs where the specific measure is implementable and in which circumstances it is most effective. Furthermore, if some LC has problems with specific road user groups, this framework allows the identification of safety measures which are targeted to that problem behaviour (e.g. safety measures targeted to pedestrians).

Table 5 presents the criteria to assess the short and long-term effects of safety measures on road user behaviour. These criteria are categorized according to the area of psychological function involved (Grippenkoven, 2017; Wickens et al., 2012). Once the estimated changes in road user behaviour have been identified (both short and long-term), the safety effects can be quantified, for example, based on Key Performance Indicators (KPIs) collected in earlier stages of the SAFER-LC project (SAFER-LC consortium, 2018b), literature, expert assessment, LC statistics etc.

Table 5. Criteria to assess the short and long-term effects of safety measures on road user behaviour.

Factors	Descriptions	Indicators
Detectability	Ease of detecting relevant visual and auditory stimuli taking into account: ▪ Conspicuity factors (sight distances, signs, crossings angle) ▪ Personal characteristics (individual visual/auditory capabilities)	▪ Detectability of LC and / or train ▪ Speed and timing of detection ▪ Prevalence of errors ▪ Number of errors (i.e. perception) / correct detections
Identification	Ease of identifying relevant information in the environment and not being distracted by irrelevant information taking into account: ▪ Cognitive factors (tiredness/fatigue, overload with stimuli / high workload, external and visual distraction) ▪ Personal characteristics (gender, age, disability) ▪ Use of addictive substances	▪ Ease of identifying relevant information ▪ Road users' focus of attention (focus on other road users and/or road) ▪ Looking left and right (yes/no, how often) ▪ Timing of reactions ▪ Type and number of errors (e.g. attention, memory etc.)
Rule knowledge	Ease of eliciting and retrieving relevant information or knowledge about required/safe behaviour taking into account: ▪ Prior acquired knowledge ▪ Understanding of the correct action	▪ Knowing the cue from the traffic rule / traffic sign etc. ▪ Knowing required behaviour (i.e. what to do when you detect the cue) ▪ Prevalence of errors ▪ Number of errors / correct replies ▪ Prevalence of violations ▪ Type and number of violations
Decision-making	Ease of taking more accurate decisions and arriving at safe behavioural intentions taking into account: ▪ Subjective risk estimates and cognitive biases (perception of probability, dangerousness, legal consequences and cost-benefits) ▪ Individual motivations (time pressure, suicide or vandalism intentions) ▪ Personal characteristics (personality of the road user, frustration and impatience)	▪ Prevalence of errors ▪ Type and number of errors (e.g. biased decision)
Behavioural execution	Focus on the motor execution of the action; ease of executing safe actions (required behaviours), and/or the difficulty of executing risky (non-adapted) behaviours taking into account: ▪ Behavioural intention and its antecedents (e.g. decision-making) ▪ Personal characteristics (e.g. movement ability, motor fitness)	▪ Risky behaviours and prevalence of violations; type and number of violations (at active LC), trajectories ▪ Speed choice / approach speed ▪ Verification behaviours for frequent users ▪ Time to collision (TTC) when a train is coming ▪ Interaction with other road users

Table 6 presents the three criteria to assess user experience and social perception regarding the safety measure. The indicators refer to the subjective opinions of road users and thus this information will most likely be collected through a questionnaire among relevant stakeholders and road users or through interviews with selected representatives of these categories. Social acceptance on the part of the end user and wider community is important, as it may affect their interaction and correct usage of the measure, potentially affecting safety. Information related to these indicators is proposed to be collected via a Likert scale, which means that the respondents specify their level of agreement or disagreement on a symmetric *agree-disagree* scale for a series of statements.

Table 6. *Criteria to assess user experience and social perception.*

Factors	Descriptions	Indicators
Acceptance	Provides an estimate of how well the measure is accepted by the public (e.g. social acceptance among road users) and by the relevant stakeholders (e.g. railway operator, rail infrastructure manager, train drivers, people living nearby, authorities, government). The estimates of acceptance by road and rail stakeholders should consider the perceived ease of implementation, namely the ease of integration within the road and rail environment and the ease to implement and use the safety measure with other safety measures.	Subjective self-report measure from the available categories of respondents (Likert scale)
Reliability (Trust)	Estimates if the users trust the system and whether they know that it is fail-safe	Subjective self-report measure from the road users (Likert scale)
Usability (Level of self-explaining nature)	Estimates to what extent the 'configuration' / 'design' of the safety measures is easy to perceive, understand and use by the road user (e.g. no language barriers to understand the signage)	Subjective self-report measure from the road users (Likert scale) Easily perceived, understood and used by all road users Easily perceived, understood and used by children, the elderly or the disabled

3. PILOTING OF THE HUMAN FACTOR ASSESSMENT TOOL

3.1. Application of the tool in the SAFER-LC pilot tests

The present section describes the process of testing the Human Factors Assessment Tool (HFAT) through its application to the SAFER-LC pilot tests. The aim of the testing, the methodology followed, and the process of collection of the feedback are described here below. The feedback obtained will be analysed through section 3.2 of this document.

The HFAT supports railway (or other relevant) stakeholders to estimate the effectiveness of a trialled LC safety measures from a HF perspective, by evaluating its impacts on human behaviour. The objective of testing the HFAT was to validate the tool and improve it considering the feedback received, so that by the end of the project it results in a valid assessment tool of the 'human' component in the safety of LCs which will be included at the final version of the SAFER-LC toolbox.

In order to do so, two reporting phases of testing at the pilot sites were planned: 1) *baseline assessment* (before measure implementation) and 2) *final assessment* (following a period of implementation)The objective was to obtain two types of information: 1) capturing HF impact data, and 2) providing feedback on the usability of the tool.

The HFAT tool was developed during the first part of Task 2.2 and presented in the deliverable D2.2, *Human Factor methodological framework and application guide for testing (interim report)*, (Havârneanu et al., 2018). That deliverable included the tool itself plus an application guide for its usage within the project.

The main Task 2.2. inputs came from other SAFER-LC activities such as WP1 and WP2. Specifically, deliverable D1.3 (SAFER-LC consortium, 2018a) provided preliminary requirements and recommendations to be considered in the evaluation activities of Task 2.2. Task 2.1 constituted the main source of input since it helped to define a set of criteria for self-explaining and forgiving LC design (D2.1; SAFER-LC consortium, 2018b). Furthermore, the risk evaluation activities of WP3 enabled the identification of behavioural models of *user-user* and *user-infrastructure* interaction at LCs.

The SAFER-LC project involved a strong connection between WP2 *Human Factors at Level Crossings* and WP4 *Lab test, field implementation and evaluation*. Specifically, the first version of the HF assessment tool was tested within WP4 pilot activities through its application in evaluating those LC safety measures which were found relevant for HF assessment. This HF assessment had important ties with WP4, which consisted of two parts: 1) detailed test site descriptions (i.e. locations where the measures can be tested) defined in deliverable D4.1 (SAFER-LC consortium, 2018c), used as implementation guidelines in the process of testing the safety measures , and 2) an evaluation framework proposed in deliverable D4.2 (SAFER-LC consortium, 2018d) including a collection of KPIs for use in collecting the data to assess the piloted LC safety measures. These KPIs were organised in five groups: *Safety* (accident and incident risk at LC), *traffic* (road and rail traffic flow), *human behaviour* (perception, understanding and compliance of LC users), *technical* (LC operation and maintenance), *business* (financial effort). The HF indicator subset (*human behaviour*) was developed in cooperation with WP2. While the Human Factors KPIs developed in WP4 represented a generic collection of indicators to assess the impact of safety measures on human behaviour, the aim of WP2 was to specify a set of indicators adapted to the purpose of the Human Factors Assessment Tool.

Further, WP4 linked all relevant KPIs to the test site descriptions, specifying which indicators were realistic to be tested and where.

The first version of the SAFER-LC HFAT and its application guide (Deliverable D2.2) were used to evaluate innovative solutions to enhance the safety of LCs during the SAFER-LC pilot tests. During the testing phase, the information from the pilot tests and piloted LC safety measures was shared between both work packages in order to fine-tune the final outcome of the HFAT, and to verify and improve it for the final end-users.

The HFAT is divided into three sections which correspond to the theory-derived criteria selected for the HF methodological framework:

- *Classification criteria* included in a classification checklist (marked in orange colour)
- *Criteria to assess the behavioural safety effects* included in five separate assessment sheets, one for each criterion (marked in green colour).
- *Criteria to assess the user experience and social perception* included in one assessment sheet (marked in blue colour).

For the testing, an additional section to gather information on the usability of the tool has been added under the heading *Human Factors Assessment Tool Feedback Form*. It includes forms and open questions to capture users' in-depth understanding of the tool and reflections on its clarity and usability, giving space to indicate any potential issue experienced while completing the tool. The HFAT template used for evaluation was presented in deliverable D2.2 (Havârneanu et al., 2018) and is included in Annex A of this document.

The measures developed within WP3 and/or selected for piloting within the SAFER-LC project were tested in different environments at several test-sites (e.g. laboratory, driving simulator, real-world conditions). The various test sites available in the SAFER-LC project were a perfect fit for LC safety measures at different stages of maturity. Early-stage developments could be tested in simulation environments or on controlled test tracks, while more readily developed measures were evaluated in field pilots. In addition, the various test environments allowed a mixed (complementary) method approach, combining advantages of different research methods (e.g. simulator study vs field study) and filing in the HFAT with data generated through different research setups. For example, the simulator studies allowed to confront drivers with different measures at a LC under the exact same conditions (traffic, weather, etc.) and to assess behavioural data that are very much harder to assess in the real world (e.g. gaze data).

The first action in the application of the HFAT was to select the pilot test sites and measures in which the HFAT could be tested. The decision was made based on discussions with the pilot test leaders to select the most appropriate test sites and piloted LC safety measures for the HF assessment. The list of selected test sites and LC safety measures for this validation exercise are presented in Table 7.

Table 7. *Selected human factors related LC safety measures and pilot test sites.*

Test site	LC safety measures
SNCF driving simulator (France)	1. Colour marks on the ground 2. Implementation of posts with a funnel effect. 3. Rings located upstream of the LC 4. Replacement of the LC warning light with a traffic light 5. Bumps in front of the LC and flashing sticks on the right edge of the roadway 6. Message in connected vehicles displaying the status of LC
DLR driving simulator (Germany)	1. Blinking peripheral lights drawing road user attention 2. Blinking Lights on the train 3. Sign “Look for train” 4. Noise-producing pavement (rumble strips) on advance to LC
TRAINOSE & CERTH Thessaloniki living lab (Greece)	1. LC and train proximity in-vehicle alert
INTADER real world pilot (Turkey)	1. Attractive sign for children 2. Coloured road markings 3. Flashing lights on barriers

Once the pilot test sites and the LC safety measures to participate in the validation exercise were selected, the HFAT was introduced to the pilot test leaders. In addition, a list of contact persons were collected for each pilot site. During the month of September 2018, a first email with the outline of the tool and a brief explanation on the required information and instructions on how to gather it was sent to the pilot test leaders. Afterwards, during the October SAFER-LC Progress Meeting there was a dedicated session between partners in both work packages to gather feedback on the use of the tool, to exchange views and experiences about the application of the tool, and to provide additional explanations to pilot test leaders if something was unclear. This session allowed addressing major questions and strengthening the project partners’ knowledge on the tool. As a final step, the HFAT and the application guide were sent to the pilot site leaders by the end of October. The dedicated session was seen important to confirm that the tool was completely understood and that the information that needed to be gathered by the pilot site leaders was clearly outlined. Individual support was available during the whole process to all pilot test leaders to guarantee an easy application of the tool and evaluation of the piloted LC safety measures from the HF perspective.

Figure 2 shows the main milestones in the application of the HFAT in the SAFER-LC pilot tests.

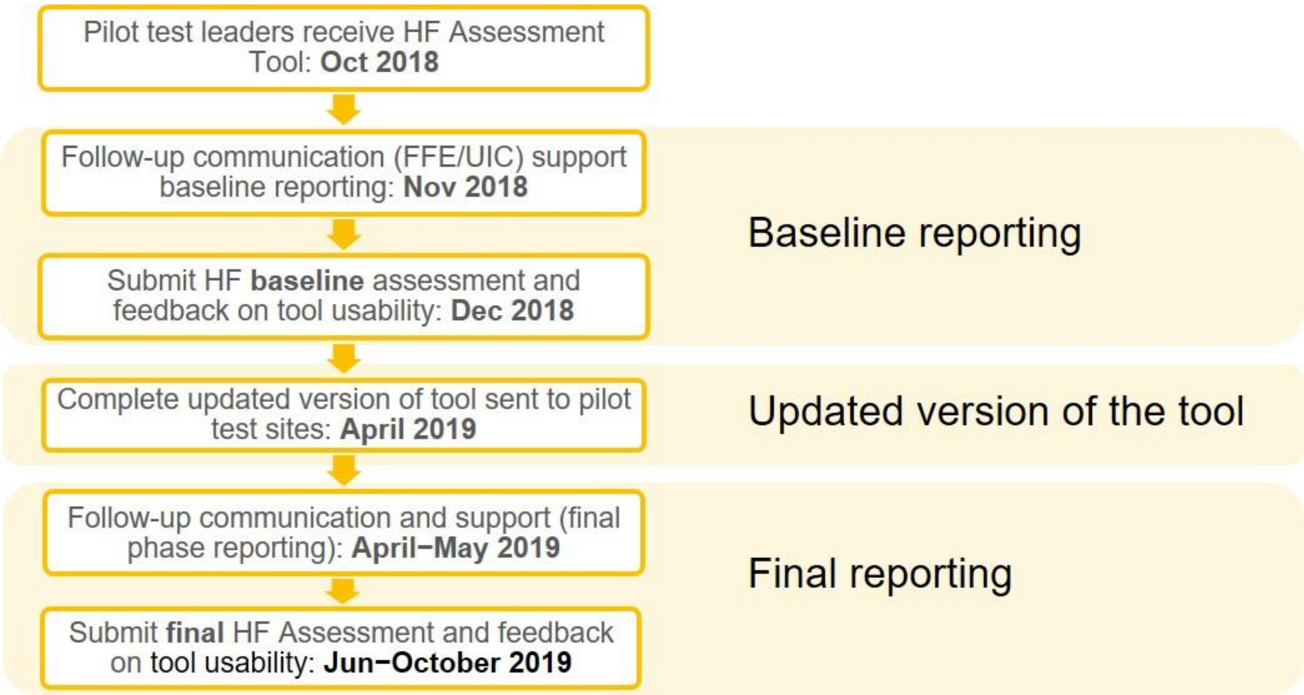

Figure 2. *Schedule of the collection of feedback from pilot sites.*

Baseline reporting

During the months of November and December 2018, the pilot site leaders were asked to submit their feedback on the usability of the tool, according to the characteristics of the pilot test site and the piloted LC safety measure(s). In addition, in order to have an understanding of the baseline situation (a "zero" / "control" case) with which to develop a comparative analysis, the pilot test leaders were asked to fill in the HFAT for the "before" situation where no measures were yet implemented at the pilot sites. This was done for the pilot tests where baseline data was already collected at that time. In some of the pilots, namely the simulation tests, the baseline was included as one of multiple conditions within one test session, while the studies were scheduled for spring and summer 2019. For these pilot sites, the "baseline" reporting involved the collection of evidence available from the literature and a HF assessment based on this information. The understanding on the baseline situation is important for the analysis on the impact of the implemented LC safety measures.

The baseline reporting was done in in close contact with the WP2 partners leading this work and different pilot test leaders (phone calls, emails and teleconferences) in order to solve any doubts or concerns. The aim was that the instructions on providing the feedback on the usability of the tool and on the baseline situations were clear and precise to be useful in the later stages of the process.

Information from the baseline application of the HFAT was gathered from the four pilot test sites (SNCF driving simulator – six measures, DLR driving simulator – four measures, Thessaloniki living lab – one measure, and real world pilot in Turkey – three measures). In addition, two other pilot test leaders volunteered to review the tool and provide initial feedback. This led to a total of six evaluations of the initial version of the HFAT.

Updated version of the tool

The information gathered from the baseline reporting was used to revise the HFAT, and an updated version of the tool (version 2) was issued in April 2019. From that moment on, the pilot test leaders could gather the information on the measures that were applied in each of the pilot test sites completing the HFAT in relation to the reporting of the results collected during the piloting of the LC safety measure(s).

Final reporting

Initially, the collection of information was planned from April 2019 until June 2019. However, the deadline was extended until October 2019, to be adapted to the implementation rhythm of the pilot tests. During this final reporting period, the pilot test leaders received support and advice regarding the application of the HFAT, when needed.

After the implementation of the LC safety measure(s), the final reporting for the HFAT was completed by three out of the four pilot test leaders (SNCF driving simulator – six measures, DLR driving simulator – four measures, Thessaloniki loving lab – one measure). The pilot test initially planned in Turkey (3 measures) could not be implemented due to political reasons.

3.2. Feedback obtained during piloting and implications

This section describes the feedback obtained during piloting the HFAT (in December 2018 and October 2019) and the implications for the revisions and further development of the HFAT.

Baseline assessment and feedback on tool clarity and usability

The pilot leaders were asked to indicate any issues they experienced while completing the HFAT, by responding to the open feedback questions and rating the clarity and usability of the tool.

All four pilot test leaders completed and returned their baseline assessment forms in December 2018. Two other pilot test leaders offered to check the tool and provided feedback on its clarity and usability along with suggestions for improvement. Therefore, a total of six HFAT evaluations were collected in the baseline assessment. The quantitative ratings of the clarity and usability are presented in Table 8. The average ratings suggest an overall satisfactory feedback (above three on a 5-point scale). However, there was an added value of the evaluation exercise, most of it coming from the qualitative feedback collected through open questions.

Table 8. *Overall rating of clarity and usability of the HFAT during the baseline evaluation (December 2018).*

	Pilot1	Pilot2	Pilot3	Pilot4	Bonus1	Bonus2	Mean
Clarity	4	3	4	4	3	4	3.66
Usability	3	2	4	4	3	4	3.33

The qualitative remarks on clarity and usability were encouraging:

- *The tool is clear and looks useful for me.*
- *We think it is clear and usable in general.*
- *The tool is as good as any other. The example is a good thing.*
- *Very huge work, complete assessment, a little bit complicated to understand and we have to read several times the instructions before filing.*

Some answers to the open questions were somewhat more critical and provided some insights on how clarity can be improved:

- *The assessment tool is very interesting but not always easy to understand. It needs an oral explanation and should be filled during a working session.*
- *The explanations provided in order to fill the assessment tool are clear but too long. My main suggestion is to have an oral explanation for every question and check if all the possibilities of responses are mentioned. Check also if this assessment tool is suitable for every measure in the project.*
- *It is very complete in terms of usability, safety, acceptability but I am not sure it is suitable for every measure planned in the project. For every question we have to suppose that the equipped LC will be integrated and working in the daily management of the railway system.*

A couple of comments provided some insights on how usability can be improved:

- *Once we have understood how to use the tool, it is easy to fill the questionnaire. The responses are not always immediately available for some questions and for some of them it is difficult to give the right choice.*
- *Simplify the questionnaire, explain shortly the procedure, provide possibility to have open responses if these latter are not in the list.*

Five suggestions were more detailed, targeting specific parts of the HFAT to improve clarity and usability. For example:

- To rephrase some items in the forms (e.g. “Conspicuity” of the train/LC -> “Detection” of the train/LC
- To provide additional instructions in the application guide
- To optimize the tick boxes in the classification criteria tables
    - Make them easier to tick/untick in the electronic version of the document
    - Provide an option for “all cases”
- To regroup the five behavioural assessment criteria categories into three categories:
    - Detectability + Identification
    - Rule Knowledge + Decision-making
    - Behavioural execution
- To revise the effect mechanisms list in the form containing classification criteria (add, group, rename?)

Final assessment and feedback on tool clarity and usability

In October 2019 (about 10 months after the initial evaluation of the HFAT), three out of four pilot test leaders submitted their final HFAT assessment, after having completed it with data collected during the pilot tests. The final assessment phase yielded HF Assessment Data for 13 LC safety measures: *blinking lights for the locomotive front, coloured road markings on approach to the LC, in-vehicle proximity warning, rings upstream of the LC, traffic light, blinking amber light with train symbol, funnel effect pylons, message ← Is a train coming? → written on road, peripheral blinking lights, rumble strips, sign ← Is a train coming? →*, and *speed bump and flashing posts*. The quantitative and qualitative data from the three HFAT sections were analysed in order to evaluate the measures from a HF perspective. The results have been described in SAFER-LC Deliverable 2.4 (Dreßler et al., 2019).

This second use of the HFAT allowed gathering new feedback on the improved tool as well as new ratings for its the clarity and usability. The quantitative ratings of clarity and usability are presented in Table 9. The number of evaluation data sets was lower compared to the baseline assessment, but the average ratings suggest again an overall satisfactory feedback (above three on a 5-point scale).

Table 9. *Overall rating of clarity and usability of the HFAT during the final evaluation (October 2019).*

	Pilot1	Pilot2	Pilot3	Mean
Clarity	4	3	5	4.00
Usability	3	2	5	3.33

The results suggest that the clarity of the HFAT slightly improved following the revisions implemented in the tool in January and February 2019. However, the usability score remained the same. The differences between the two evaluations are shown in Figure 3.

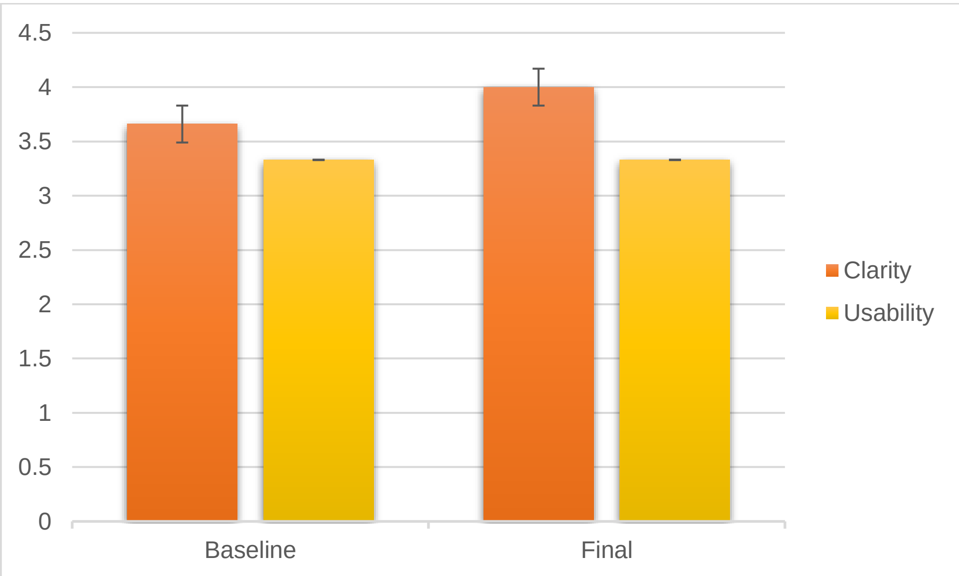

Figure 3. *Average ratings of clarity and usability of the HFAT during the baseline and final assessment.*

The second evaluation also brought some additional qualitative remarks. For example one pilot test leader said that *some questions are confusing* requesting *support from sociologists or psychologists* in order to answer them. Another pilot test leader stated that filling the tool was *generally easy with very few constraints (few unspecific formulations e.g. "… with other warning devices")*.

The HFAT was judged by one pilot test leader as *definitely feasible, but sometimes not very easy to provide the information requested*. The difficulties pointed out were related to the usability reasons (e.g. *un/ticking boxes*) or content-related reasons, *e.g. difficulty to separate stages of information processing from the available empirical data*.

Specific suggestions on how to further improve the HFAT included:

- *Adding navigation elements (e.g. a clickable Table of Content / Heading links) would be helpful to quickly jump from one section to the other while filling the forms. These would ease filling in the HFAT since information collection is not always a linear process.*
- *Implementation of an interactive form / online survey from which entries are directly saved as variables. This would facilitate analysis and comparison of multiple measures.*
- *Classification criteria form:*
    - *Do not use the "other warning devices" category because it is unclear. Proposal: Maybe include some examples. Include open entry field (please specify: _____).*
    - *Separate "LCs with low usage / not used at all" in two different options*
    - *For the "Type of road user" and "Age", use a more user-friendly way to place the tick boxes. Alternative options could be spatially separated, e.g. "all road users" on left side, all the other options on right side, so it will be clear that if you pick "all road users" you won't need to also check all the other options.*

Overall, the two-step evaluation of the HFAT by the pilot test leaders was a useful and productive exercise. It allowed collecting valuable inputs, suggestions and ideas on how to improve specific parts of the tool. Most of the recommendations were implemented during the HFAT revision process. The first round of feedback collected with the baseline evaluation was integrated in January and February 2019 leading to a revised version of the HFAT (Annex B). Its application on a real case study is shown in Annex C.

The second round of feedback collected with the final evaluation allowed gathering additional improvement ideas which were used to slightly revise and amend the HFAT. This led to the final version of the tool developed in the earlier stages of the SAFER-LC project (Havârneanu et al., 2018) (Annex D).

While most of the evaluation feedback was taken into account during the HFAT revision process, it is worth mentioning that not all received suggestions could be implemented. Some ideas were good, however it was not realistic to adopt them within the SAFER-LC timeframe and resources. Others were subject to group discussion during the project meetings and were implemented only partially, following the collective decision.

The main updates of the HFAT compared to the original version of the HFAT documented in Havârneanu et al. (2018) are described in the next chapter.

4. FURTHER DEVELOPMENT OF HF ASSESSMENT TOOL

This chapter describes the main changes which occurred during the HFAT review process leading to a tool with an improved clarity and usability. It points out the main differences between the initial version of the HFAT presented in deliverable D2.2 (Havârneanu et al. 2018) and the final version, which is described in Annex D of the current deliverable.

Based on the received feedback, changes concerned only the classification criteria (orange form) and the criteria to assess the behavioural safety effects (green forms).

4.1. Major changes

The following major changes were agreed to improve the tool.

Revision of effect mechanisms list in the classification criteria table

There were several proposals to add new effect mechanisms, to group some of the existing ones, to rename them and even to present them in a different reorder. A discussion on this topic was carried out between the consortium members to define the final list of effect mechanisms. The collective decision was made during a consultation exercise which led to an agreement on the final list of effect mechanisms.

Of the five newly proposed categories, two were omitted as they are seen to be covered within existing categories; two were added within existing mechanisms by slightly expanding the names of the current categories and one was added as a separate new mechanism, as follows:

- *Avoids / obviates use of LC* and *Helps keep track clear of obstacles* were seen to be covered by *Provides up-to-date information about the status of LC* and hence were not added.
- *Deters rule violation / incorrect behaviour* was added to an existing mechanism, yielding *Increases awareness of correct behaviour <u>and consequences of rule violation</u>*.
- *Facilitates egress from collision zone* was added to an existing mechanism, yielding *Controls access to <u>and supports egress from</u> LC*.
- *Improves waiting experience at LC* was adapted as *Makes waiting time more tolerable*.

Table 10 presents the original and revised list of the intended effect mechanisms. The revised list was validated during the consultation exercise with all SAFER-LC partners. The effect mechanisms were updated within the final version of the HFAT presented in this deliverable.

Table 10. Intended effect mechanisms in the HFAT: Original and revised list.

Before	After
1. Improves the conspicuity of train	1. Improves the detection of train
2. Improves the conspicuity of LC	2. Improves the detection of LC
3. Controls the access to LC	3. Controls access to and supports egress from LC
4. Reduces the approach speeds of vehicles	4. Reduces the approach speeds of vehicles
5. Increases the awareness of correct behaviour and dangerousness of LC	5. Increases the user's awareness of correct behaviour and consequences of rule violation
6. Improves the physical environment of LC	6. Improves the physical environment of LC
7. Improves the possibilities of vulnerable road users to cross LC safely	7. Improves the possibilities of vulnerable road users to cross LC safely
8. Provides up-to-date information about the status of LC	8. Provides up-to-date information about the status of LC
9. Supports the LC safety actions	9. Supports the LC safety actions
	10. Makes waiting time more tolerable

Regrouping of areas of psychological function in assessment of behavioural safety effects

There was also a proposal to regroup the five behavioural assessment criteria due to the closeness between some of the categories which made it complex to assign the empirical evidence collected to one category or another (most notably *detection* and *identification*). After some discussion among consortium partners, it was decided to reduce these to four categories, grouping *detection and identification* together. There was also some debate as to whether to group *rule knowledge* and *decision-making* together, or *decision-making* and *behavioural execution*, but with no clear consensus it was decided to keep them as three separate categories. In practice, this was reflected in a shorter HFAT which now has only four green forms to be filled in.

4.2. Minor changes to improve clarity

Based on the feedback, the numerous but rather minor changes were implemented to improve the clarity of the tool. Most of these concerned the *Classification criteria* form and consisted in renaming or rephrasing different words, separating one item into two, etc. The following list is not exhaustive but presents the most representative examples of changes:

- *Conspicuity of the train / LC* has been changed to *Detection of the train / LC*;
- *LCs with low usage / not used at all* was separated in two different options;
- The *LC with other warning devices* category was changed into a category with an open entry field *(please specify: _____)*

4.3. Minor changes to improve usability

Based on the feedback, the following changes were implemented to improve the usability of the tool. Most of these concerned the *Classification criteria* form and referred to the usability of the tick boxes:

- The original tick boxes were replaced with a new type of box and it became easier to tick/untick them in the electronic version of the document.
- An option for *all cases* was included, enabling the user to tick only once instead of having to tick every single option.
- The tick boxes were placed in a more user-friendly way for most indicators, using an additional column (i.e. a spatial separation) which makes it clear to choose the option and tick it.

5. DISCUSSION AND CONCLUSIONS

This chapter provides an overall discussion of the HFAT, its strengths and limitations, its current utility as a stand-alone methodology, and possible directions in its further applicability and development beyond the SAFER-LC project.

The HFAT presented in this deliverable was developed during the SAFER-LC project to evaluate the efficiency of LC safety measures with respect to road users' needs, cognitive processes, and behaviour. The first version of this tool (and the related HF methodological framework) was built in earlier stages of the SAFER-LC project and reported in Havârneanu et al. (2018). The tool was verified and further developed based on the feedback received during the piloting of LC safety measures. The developed tool includes several sets of criteria and associated indicators to classify and assess LC safety measures, taking into account the road and rail users' perspectives and requirements.

The HFAT was validated in four pilot tests and through two evaluation rounds. In general, the clarity and usability of the HFAT was assessed to be rather good (both were above three on a 5-point scale). After the revisions, the clarity of the HFAT was even further improved.

Two major changes were implemented to the revised version of HFAT: 1) rewriting and reordering of the intended effect mechanisms as part of classification criteria, and 2) grouping of *Detection* and *Identification* as part of the behavioural assessment criteria. In addition, several minor changes were made to the wording and categories used in the tool, and the usability of tick boxes was improved.

The tool was developed to support both road and rail stakeholders involved in LC safety to better understand road users' needs and related requirements to be able to consider them in the implementation of future designs for LCs. For example, the HFAT could be used as a checklist to support the consideration of the HF perspective in the evaluation of LC safety measures. Specifically, the application of the tool supports the LC stakeholders in tailoring unique solutions for different LC environments and to design the layout of LCs and the safety measures in such a way that they enhance the self-explaining and forgiving nature of the LC infrastructure for the road users.

Feedback collected through the HFAT in the demonstration phase allowed making recommendations and evaluate the developed measures regarding technical specifications and human and organisational processes. At the same time, the HF methodological framework and assessment tool were adjusted and improved based on feedback from the pilot test sites, as reported in this deliverable.

5.1. Additional feedback drawn from the evaluation of HFAT data

Besides the direct feedback that the test site leaders gave on their experience with using the HFAT, the experiences made in joining, comparing and analysing the obtained HFAT data are another source of information that provides insights on how the tool could be further developed (cf. Dreßler et al., 2019). The main points are reviewed here (cf. p. 67 and following).

In its current form, the tool contains detailed instructions on how it should be applied, including an exemplary completion of the forms for one measure. In the evaluation of HFAT data, a few more details were discovered in terms of how these instructions could be refined. In the first HFAT section, *applicability*, two apparently different interpretations were observed among the contributing test site leaders concerning the use of the subcategories *disability, under influence of*, and *under skill impairing states* within the indicator subset *adaptation to individual characteristics and conditions of users*. One apparent understanding was that a measure should be indicated as suitable for a certain user characteristic if it has the potential to facilitate safe behaviour in persons with this characteristic, too. The other one was probably to indicate a measure as suitable only if it specifically addresses persons with the respective characteristic. In a further development, the HFAT instructions could specify how this categorization should be handled.

Further specifications could also address the second HFAT section, behavioural safety effects. An explicit standard could be defined of how to deal with the assignment of a score to stages to which no finding can be directly allocated: Should a score be assigned based on theoretical reasoning, or should no score be assigned at all? As a related topic, an explicit standard could be defined of how to deal with stages that are not directly influenced by a measure. One possible way would be to reason that an influence on one stage can also affect the processing at following stages (e.g. if a user is more likely to detect a train approaching, this may also influence her decision not to cross the tracks in a critical phase). Another way would be to require that a 0 should be assigned to the following stages if they were not measured by independent indicators. Whatever method is chosen should fit the method of integrating all the results in the end.

Considering the results in terms of the assignment of findings to the stages of information processing, another idea for the further development of the HFAT could be to include more specific behavioural descriptions of the target effects on behaviour within the stages. For example, evidence concerning an observed speed reduction was sometimes cited for decision-making, sometimes for behavioural execution. If there was a behavioural indicator as, e.g., “induces speed reduction on approach”, it would be easier for the users of the tool to find the right place to insert a finding from a study. A more specific description could moreover be helpful in the specification of requirement profiles to allow a standardised integration of all information into an overall assessment. For example, knowing that reduced speed on approach of a passive LC is a good prerequisite to enable effective visual search and coming to a stop in time if necessary, but futile if it is not indeed combined with increased visual scanning, this could be reflected in the requirements.

In the evaluation mentioned, a qualitative approach was used to integrate information from all the three HFAT sections in an overall assessment of measures. This approach could be used as a starting point to further refine the integration of results in a future procedure. For example, it may be possible to devise a reasonable procedure for the computation of an overall score if the relevance of single stages (or behavioural indicators) can be defined a priori for a given measure, based on its scope (application context), and these relevance values can be used as weights in the computation of the overall score. In the reported analysis, the scope of the measure was considered by using the information on the main psychological functions and intended effect mechanism. This may be refined by including further information such as the target LC type (e.g. Is approach speed relevant to accidents there?) and the target road users (e.g.: How fast are they usually?) in the definition of relevant stages or indicators (e.g.: Is speed reduction a desirable target behaviour?). In order to avoid an abundance of possible combinations that would need to be parametrized in this way, the analysis could start with a few prototypical use cases in order to assess whether this is a promising approach. These use cases could be selected as combinations of road user type, LC type and potential other features that are specifically relevant in LC accidents (cf. Silla et al., 2017).

All these lessons learnt during the evaluation exercise can be considered in possible future revisions of the HFAT after the end of the SAFER-LC project.

5.2. Further inclusion of the HFAT in the SAFER-LC toolbox

After its application during the WP4 pilot tests, the HF methodological framework resulted in a valid assessment tool of the ‘human’ component in the safety of LCs which will be included in the SAFER-LC toolbox (http://toolbox.safer-lc.eu/) which will be accessible from the project final conference onwards. Its inclusion in the toolbox will ensure that the HFAT will be accessible through a user-friendly interface and that the HF perspective will be accounted for within the management of safety at LCs by the road and rail infrastructure managers.

Beyond its research purposes, the HFAT’s added value resides in the fact that it may inspire a particular policy vision in rail and road safety, by promoting a human- and organisational-factors approach when thinking about the implementation of LC safety measures.

In line with the systems approach, it reminds policy makers that LC safety measures should always be considered both in technical and HF terms. Technical solutions and in particular engineering solutions implemented at the infrastructure level must be considered with respect to their potential behavioural effects and analysed according to a set of relevant indicators which are strongly linked to the precursors of road user behaviour. In addition, safety measures that help the infrastructure become more self-explaining and forgiving should consider all aspects of information processing, such as perception, memory, action execution, etc. The HFAT checklists cover all these important aspects.

6. REFERENCES

- Bekiaris, E. & Gaitanidou, E. (2011). Infrastructure and Safety in a Collaborative World pp. 15-22 Chapter Towards Forgiving and Self-Explanatory Roads, Springer-Verlag Berlin Heidelberg.
- Block, R. A. & Zakay, D. (1996). Models of psychological time revisited. In H. Helfrich (Ed.), Time and mind (pp. 171-195). Kirkland, WA: Hogrefe & Huber.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36(03), 181–204. https://doi.org/10.1017/S0140525X12000477.
- Dreßler, A., Silla, A., Havârneanu, G., Boufidis, N., Taillandier, V., Herranz, A. & Grippenkoven, J. (2019). Evaluation of new human-centred low-cost measures. Deliverable D2.4 of the SAFER-LC project. http://safer-lc.eu/
- Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., & Victor, T. (2017). Great expectations: a predictive processing account of automobile driving. Theoretical Issues in Ergonomics Science, 1–39.
- Fuerstenberg, K. & Boehning, M. (2012). Final Report. D1.2, MINIFAROS project.
- Grippenkoven, J. (2017): Wahrnehmung und Verhalten am Bahnübergang. In: Deine Bahn (2), 10–15.
- Grippenkoven, J., Giesemann, S. & Dietsch, S. (2012). The Role of Human Error in Accidents on German Half-Barrier Level Crossings.
- Havârneanu, G. M., Silla, A., Whalley, S., Kortsari, A., Dressler, A. & Grippenkoven, J. (2018). Human Factor methodological framework and application guide for testing (interim report). Deliverable D2.2 of the SAFER-LC project. https://safer-lc.eu/IMG/pdf/saferlc_20180724_d22_v04web_uic_hf_methodological_framework.pdf
- Innamaa, S., Rämä, P., Hogema, J., Kroon, L., Heinig, I., Schindhelm, R., Ojeda, L., Servel, A., Visintainer, F. & Giosan, I. (2014). Safety impacts of cooperative systems. Deliverable IR4.y of DriveC2X project.
- Kulmala, R. (2010). Ex-ante assessment of the safety effects of intelligent transport systems. Accident Analysis and Prevention 42(4),1359–1369.
- Kulmala, R., Leviäkangas, P., Sihvola, N., Rämä, P., Francsics, J., Hardman, E., Ball, S., Smith, B., McCrae, I., Barlow, T. & Stevens, A. (2008). Final study report. CODIA Deliverable 5.
- Maddux, J. E., Rogers, R. W. (1983): Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. Journal of Experimental Social Psychology 19(5), 469–479.
- Neisser, U. (1976). Cognition and reality: Principles and implications of cognitive psychology. In Cognition and reality: Principles and implications of cognitive psychology. New York, NY, US: W H Freeman/Times Books/ Henry Holt & Co.
- Öörni, R., Hietikko, M., Kauvo, K., Lattunen, A. & Virtanen, A. (2011). In-vehicle Warning System for Railway Level Crossings. Pilot Tests on Hanko–Hyvinkää Railway Line [Autossa toimiva junavaroitusjärjestelmä. Kokeilu Hanko–Hyvinkää-rataosalla]. Espoo 2011. VTT Research Notes 2603. 106 p. + app. 30 p.

- Rasmussen, J. (1986). Information processing and human-machine interaction: An approach to cognitive engineering. New York: Elsevier.
- Read, G.J.M., Salmon, P. M. & Lenné, M. G. (2013). Sounding the warning bells: the need for a systems approach to rail level crossing safety. Applied Ergonomics, 44, 764–774.
- Reason, J., (1990). Human Error. Cambridge University Press.
- Ryan, B., Kallberg, V-P., Rådbo, H., Havarneanu, G., Silla, A., Lukaschek, K., Burkhardt, J-M., Bruyelle, J-L., El-Koursi, E-M., Beurskens, E. & Hedqvist, M. (2018). Collecting evidence from distributed sources to evaluate railway suicide and trespass prevention measures. Ergonomics, 1-41. DOI: 10.1080/00140139.2018.1485970.
- SAFER-LC Consortium (2018a). Needs State of the art of LC safety analysis: identification of key safety indicators concerning human errors and violations. Deliverable 2.1 of the SAFER-LC project. http://safer-lc.eu/IMG/pdf/saferlc_20180307_d21_v04_ffe_state_of_the_art_of_lc_safety_analysis.pdf
- SAFER-LC Consortium (2018b). Evaluation Framework. Deliverable D4.2 of the SAFER-LC project. https://safer-lc.eu/IMG/pdf/saferlc_20180430_d42_v13_dlr_evaluation_framework.pdf.
- Scholliers, J., Heinig, K., Blosseville, J., Netto, M., Anttila, V., Leanderson, S., Engström, J., Ljung, M., Hendriks, F., Ploeg, J. & Chen, J. (2007). D16.3 Proposal of procedures for assessment of preventive and active safety functions. PreVENT SP Deliverable.
- Silla, A., Peltola, H., Aittoniemi, E., Sintonen, H., Kortsari, A., Taillandier, V., Khoudour, L., Françoise, C., Petrelli, M., Renna, A., Kassa, E., Whalley, S., Lorenzo, L., Hatice Öztürk, E., Karahan, G. (2017). Level crossing accidents and factors behind them. Deliverable D1.2 of the SAFER-LC project. https://safer-lc.eu/IMG/pdf/saferlc_20171003_d12_v04_vtt_lc_accidents_and_factors_behind_them.pdf
- Silla, A., Rämä, P., Leden, L., van Noort, M., de Kruijff, J., Bell, D., Morris, A., Hancox, G. & Scholliers, J. (2017). Quantifying the effectiveness of ITS in improving safety of VRUs. IET Intelligent Transport Systems. lVol. 11 (2017) No: 3, 164–172.
- Silla, A., Seise, A., & Kallberg, V-P. (2015). Survey and assessment of measures aiming to improve the safety of level crossings. Finnish Transport Agency, Infrastructure and Environment. Helsinki 2015. Research Reports of the Finnish Transport Agency 7/2015. [Finnish with English summary]. http://www2.liikennevirasto.fi/julkaisut/pdf8/lts_2015-07_tasoristeysten_turvallisuustoimenpiteiden_web.pdf
- Ulleberg, P. & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. Safety Science, 41(5), 427–443. https://doi.org/10.1016/S0925-7535(01)00077-7.
- Wickens, C. D., Hollands, J. G., Banbury, S. & Parasuraman, R. (2012). Engineering psychology and human performance (Fourth edition). New York: Pearson Education.
- Wickens, C. D. & McCarley, J. S. (2008a): Applied Attention Theory, Chap. 4: Visual Attention Control, Scanning, and Information Sampling. Boca Raton: CRC Press.
- Wickens, C. D. & McCarley, J. S. (2008b): Applied Attention Theory, Chap. 8: Time Sharing and Multiple Resource Theory. Boca Raton: CRC Press.
- Wilmink, I., Janssen, W., Jonkers, E., Malone, K., van Noort, M., Klunder, G., Rämä, P., Sihvola, N., Kulmala, R., Schirokoff, A., Lind, G., Benz, T., Peters, H. & Schönebeck, S. (2008). Impact assessment of Intelligent Vehicle Safety Systems. eIMPACT Deliverable D4, 11 August 2008.

- Wimmershoff, M., Will, D., Pütz, A., Lach, A., Schirokoff, A., Pilli-Sihvola, E., Le, L., Zlocki, A., Sihvola, N., Weingart, J., Stimming, C. & Kulmala, R. (2011). Test and evaluation results. Deliverable D8.2, INTERSAFE2. 162 p.
- Wogalter, M. S. (2006): Communication-Human Information Processing (C-HIP) Model. In: Michael S. Wogalter (Hg.): Handbook of Warnings. Mahwah, N.J.: Lawrence Erlbaum Associates.

ANNEX A: HUMAN FACTORS ASSESSMENT TOOL FEEDBACK FORM

Together with the completed forms from the previous chapter, we are also requesting your feedback concerning the tool. The forms presented below help us capture your in-depth understanding of the tool and reflections on its clarity and usability. This can be returned following the first and final phase of HF Assessment Tool reporting.

Following your experience with the tool, please indicate any potential issues you have experienced with completing it, by providing responding to the feedback questions presented below.

For each of the following sentences please indicate your level of agreement / disagreement, by writing an "X" in the box which best suits your answer. 1 means that you strongly disagree with the proposed affirmation. 5 means that you strongly agree with the proposed affirmation.

1. The HF Assessment Tool is **clear**, and it is **easy to understand the information** that is being requested.

Strongly disagree				Strongly agree
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Describe how easy it is to **understand** the information being requested (optional):

………

Provide any **suggestions to improve clarity:**

………

2. The HF Assessment Tool is **easy to use, and it is easy to provide the information** that is being requested.

Strongly disagree				Strongly agree
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Describe how easy it is to **use the tool and provide** the information being requested (optional):

………

Provide any **suggestions to improve usability:**

………

ANNEX B: HUMAN FACTORS ASSESSMENT TOOL (REVISED VERSION USED FOR THE FINAL EVALUATION)

Fill in the following forms for a given safety measure under evaluation. Each form is colour coded to reflect the three different sets of criteria under assessment: the 'Classification criteria' are included in a classification checklist (orange form). The 'Criteria to assess the behavioural safety effects' are included in five separate assessment sheets, one for each criterion (green forms). The 'Criteria to assess the user experience and social perception' are included in one assessment sheet (blue form). Detailed instructions are provided in the forms' headers.

Name of the measure being assessed	Name of the pilot test and brief description of the tested measure
…………………………………	………………………………………………………………………………………

CLASSIFICATION CRITERIA			
Factor	**Brief description**	**Indicator** *(Tick only the cases that the measure applies to, or click the option 'All' if the measure covers all the cases)*	
Applicability to different LCs	*Specify the types and characteristics of LCs where the measure can be implemented (multiple answers are possible)*	Type of LCs	☐ All ☐ Passive LCs without any warning device ☐ Active (manual) ☐ Active LCs with half barriers ☐ Active LCs with full barriers ☐ Active LCs with skirts for pedestrians ☐ Active LCs with light and sound warning ☐ Active LCs with other warning device ☐ Active LCs with traffic lights
		Characteristics of LCs	☐ All ☐ LCs with low vehicle traffic ☐ LCs with high vehicle traffic ☐ LCs with paved road ☐ LCs with gravel road ☐ LCs with availability of electricity ☐ LCs with low usage / not used at all ☐ LCs with sharp / wide crossing angle ☐ Other (specify)……………………………
Feasibility under different environmental conditions	*Specify the environmental circumstances in which the measure aims to be most effective and which may affect the perception or the behavioural adaptation of road*	Time of the day	☐ All ☐ Daylight ☐ Darkness ☐ Dusk ☐ Dawn ☐ Peak traffic hours
		Weather conditions	☐ All ☐ Rain ☐ Snowfall ☐ Slipperiness

	users (multiple answers are possible)	 Setting of the LC	☐ Fog ☐ Bright sunshine/ glare ☐ All ☐ urban ☐ rural
Applicability to different types of user	*Specify the categories of LC users who are targeted by the measure (multiple answers are possible)*	MRU VRU	☐ All ☐ cars ☐ motorbikes / mopeds ☐ trucks / heavy vehicles ☐ buses / coaches ☐ farm / agricultural vehicles ☐ other (specify)………………………… ☐ All ☐ pedestrians ☐ cyclists ☐ other (specify)…………………………
Adaptation to individual characteristics and conditions of users	*Specify if the measure is applicable for people with the following characteristics or conditions (multiple answers are possible)*	Gender Age Disability Under influence of Under skill impairing states	☐ Male ☐ Female ☐ All ages ☐ children ☐ elderly ☐ vision loss and blindness ☐ hearing loss and deafness ☐ intellectual disability ☐ reduced mobility ☐ other (specify)………………………… ☐ alcohol ☐ drugs ☐ medication ☐ fatigue ☐ stress ☐ Risk-seeking personality
Intended effect mechanism	*Specify the mechanism via which the measure is expected to have an effect on safety (maximum 3 options can be ticked; undeline the main effect mechanism)*	☐ Improves the detection of train ☐ Improves the detection of LC ☐ Controls access to and supports egress from LC ☐ Reduces the approach speeds of vehicles ☐ Increases the user's awareness of correct behaviour and consequences of rule violation ☐ Improves the physical environment of LC ☐ Improves the possibilities of vulnerable road users to cross LC safely ☐ Provides up-to-date information about the status of LC ☐ Supports the LC safety actions ☐ Makes waiting time more tolerable ☐ Other (specify) ……………………………………	

CRITERIA TO ASSESS THE BEHAVIOURAL SAFETY EFFECTS OF MEASURES ON ROAD USERS (SHORT- AND LONG-TERM)

Criterion	Brief description
Detection and Identification	The measure can help the LC user detect relevant visual and auditory stimuli and identify relevant information in the environment which can increase their detection of the LC, an approaching train or other potential danger

Write down brief descriptions of the road user's detection and identification of relevant LC safety information (e.g. detection of LC or train) before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)

Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above.

Question: To what extent does the measure facilitate the detection and identification of the LC, train or possible danger while the user is approaching the LC?

Answer modalities	N	The LC user's visual or auditory perception can be impeded/distracted by this measure
	0	This measure has no intended influence on the visual or auditory perception of the LC user
	1	
	2	
	3	
	4	
	5	LC users can easily detect the LC or the approaching train with sufficient time to stop or to cross safely (and continue to do so in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* …………………………………………………………………………………………

Criterion	Brief description
Rule knowledge	The measure can help the LC user elicit and retrieve relevant information about the required safe behaviour to cross the LC

Write down brief descriptions of the road user's ability to elicit and retrieve relevant safety information before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure evoke the required behaviour while the user is approaching the LC?**		
Answer modalities	N	The LC user is confused about how to behave safely at LC, because the measure transmits unclear or misleading information
	0	This measure has no intention to remind the LC user the required/safe behaviour
	1	
	2	
	3	
	4	
	5	LC users understand how to cross the LC safely without prior knowledge or experience of the LC type and environment in question (in all situations, also in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* ………………………………………………………………………………………………

Criterion	**Brief description**
Decision-making	The measure can help the LC user take more accurate decisions that arrive at safe behavioural intentions

Write down brief descriptions of the road user's decisions before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure facilitate the user's decision-making towards a safe course of action while approaching the LC?**		
Answer modalities	N	The LC user decides to cross unsafely, because this measure encourages their inaccurate subjective judgment of risk
	0	This measure has no intended influence on the subjective decision-making factors of the LC user
	1	
	2	
	3	
	4	
	5	LC users decide to cross the LC safety, because they understand the risks and the associated consequences of their behaviour (in all situations, also in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* ………………………………………………………………………………………………

Criterion	Brief description
Behavioural execution	The measure can 'force' the LC user execute safe actions (required behaviours) or can impede the LC user from executing risky actions (non-adapted behaviours)

Write down brief descriptions of the road user's behavioural execution before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure directly influence the safe execution of the approach and crossing behaviour?**		
Answer modalities	N	The ability of the LC user to cross safely is made difficult by this measure
	0	This measure has no intended direct influence on the LC user's execution of actions
	1	
	2	
	3	
	4	
	5	LC users are physically impeded from illegally crossing the LC or are forced to cross the LC safety when this measure is in place (also in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* ……………………………………………………………………………………………

CRITERIA TO ASSESS THE USER EXPERIENCE AND SOCIAL PERCEPTION

Choose the most appropriate answer by ticking one box for each case

Factor	Definition	(0) Un-acceptable	(1)	(2)	(3)	(4)	(5) Excellent
Accep-tance	The estimated level of acceptance by the public (e.g. road users, people living near the LC)	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
	The estimated level of acceptance by relevant stakeholders (e.g. the railway operator, rail infrastructure manager, train drivers, authorities or Government)	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
	The estimated extent to which the measure can be integrated with the road and rail environment and with other safety measures	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
Reliability	The estimated extent to which the users of the LC trust the system and know that it is fail-safe	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
Usability	The estimated level of self-explaining nature of the design of safety measure (e.g. easy to understand or use) by all road users, all age categories and persons with various disabilities	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					

ANNEX C: EXAMPLE OF A FILLED-IN HFAT (DATA FROM THE THESSALONIKI PILOT TEST ON THE REVISED VERSION)

This annex presents the revised HFAT filled-in by the pilot leader during the Thessaloniki field test. It contains baseline and pilot evaluation data. The blue text represents the answers filled in by the pilot leader.

Fill in the following forms for a given safety measure under evaluation. Each form is colour coded to reflect the three different sets of criteria under assessment: the 'Classification criteria' are included in a classification checklist (orange form). The 'Criteria to assess the behavioural safety effects' are included in five separate assessment sheets, one for each criterion (green forms). The 'Criteria to assess the user experience and social perception' are included in one assessment sheet (blue form). Detailed instructions are provided in the forms' headers.

Name of the measure being assessed	Name of the pilot test and brief description of the tested measure
LC and train proximity in-car alert …………………………………	Thessaloniki living lab - Testing in real life conditions at 30 LCs. The piloted measure introduces a mobile application developed to enhance road user safety around level crossings. The app can be installed on any common mobile device like a smartphone or tablet, and it alerts users about the presence of a LC through a pop-up window and a short audio alert, whenever they approach a LC. The warning also includes an estimated time of arrival for the case of an incoming train. ………………………………………………………………………………………

CLASSIFICATION CRITERIA			
Factor	**Brief description**	**Indicator** *(Tick only the cases that the measure applies to, or click the option 'All' if the measure covers all the cases)*	
Applicability to different LCs	*Specify the types and characteristics of LCs where the measure can be implemented (multiple answers are possible)*	Type of LCs	☒ All ☐ Passive LCs without any warning device ☐ Active (manual) ☐ Active LCs with half barriers ☐ Active LCs with full barriers ☐ Active LCs with skirts for pedestrians ☐ Active LCs with light and sound warning ☐ Active LCs with other warning device ☐ Active LCs with traffic lights
		Characteristics of LCs	☒ All ☐ LCs with low vehicle traffic ☐ LCs with high vehicle traffic ☐ LCs with paved road ☐ LCs with gravel road ☐ LCs with availability of electricity ☐ LCs with low usage / not used at all ☐ LCs with sharp / wide crossing angle ☐ Other (specify)……………………………

			The tested measure is LC agnostic; it can be implemented to any LC as long as its location is recorded in the application and the in-car device has its location services enabled. Internet connectivity further enables the transmission of estimated time of train arrivals to LCs.
Feasibility under different environmental conditions	*Specify the environmental circumstances in which the measure aims to be most effective and which may affect the perception or the behavioural adaptation of road users (multiple answers are possible)*	Time of the day Weather conditions Setting of the LC	☒ All ☐ Daylight ☐ Darkness ☐ Dusk ☐ Dawn ☐ Peak traffic hours ☒ All ☐ Rain ☐ Snowfall ☐ Slipperiness ☐ Fog ☐ Bright sunshine/ glare ☒ All ☐ urban ☐ rural
Applicability to different types of user	*Specify the categories of LC users who are targeted by the measure (multiple answers are possible)*	MRU VRU	☐ All ☒ cars ☐ motorbikes / mopeds ☒ trucks / heavy vehicles ☒ buses / coaches ☐ farm / agricultural vehicles ☒ other (specify)…………………………. *Any vehicle driver could be targeted as long as there is a smart device (with the application installed) onboard. Drivers of motorbikes and agricultural vehicles may have such a device onboard, but it is expected that they will not be able to receive the visual and/or auditory alert in a desired way* ☐ All ☐ pedestrians ☐ cyclists ☒ other (specify)…………………………. *Pedestrians and cyclists could be targeted as long as they use a smart device with the application installed.*
Adaptation to individual characteristics and conditions of users	*Specify if the measure is applicable for people with the following characteristics or conditions (multiple answers are possible)*	Gender Age Disability	☒ Male ☒ Female ☒ All ages ☐ children ☐ elderly

		Under influence of Under skill impairing states	☒ vision loss and blindness ☒ hearing loss and deafness ☐ intellectual disability ☒ reduced mobility ☐ other (specify)………………………….. ☒ alcohol ☒ drugs ☒ medication ☒ fatigue ☒ stress ☐ Risk-seeking personality
Intended effect mechanism	*Specify the mechanism via which the measure is expected to have an effect on safety (maximum 3 options can be ticked; <u>undeline the main</u> effect mechanism)*	☒ Improves the detection of train ☒ <u>Improves the detection of LC</u> ☐ Controls access to and supports egress from LC ☐ Reduces the approach speeds of vehicles ☐ Increases the user's awareness of correct behaviour and consequences of rule violation ☐ Improves the physical environment of LC ☐ Improves the possibilities of vulnerable road users to cross LC safely ☒ Provides up-to-date information about the status of LC ☐ Supports the LC safety actions ☐ Makes waiting time more tolerable ☐ Other (specify) ……………………………………	

CRITERIA TO ASSESS THE BEHAVIOURAL SAFETY EFFECTS OF MEASURES ON ROAD USERS (SHORT- AND LONG-TERM)

Criterion	Brief description
Detection and Identification	The measure can help the LC user detect relevant visual and auditory stimuli and identify relevant information in the environment which can increase their detection of the LC, an approaching train or other potential danger

Write down brief descriptions of the road user's detection and identification of relevant LC safety information (e.g. detection of LC or train) before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				Phase 1 Questionnaire relevant questions: Q7. How easy is it to detect the presence of a LC or an approaching train based on the existing LC safety measures (e.g. signs)? Q8. How easy is it to identify LCs that you were not previously aware of or a possible danger at a LC based on the existing LC safety measures (e.g. signs) Q7: Approximately 10% and 26% answered 'Not at all' and 'Slightly' respectively. Less than 10% of drivers stated that they find it 'completely' easy. Similar results for Q8; 9.3% and 27.3% answered 'Not at all' and 'Slightly' respectively, while less than 10% of drivers stated that they find it 'completely' easy.

After / With the measure			Phase 2 Questionnaire relevant questions: Q7. How easy is it to detect a LC and approaching trains using the in-car alert system? Q8. How easy is it to identify LCs that you were not previously aware of or a possible danger at a LC using the in-car alert system? Q7: 14.8% and 3.7% answered 'Not at all' and 'Slightly' respectively. Almost 29.63% of drivers stated that they find it 'completely' easy. Q8: 11.11% and 7.41% answered 'Not at all' and 'Slightly' respectively, while 41.7% of drivers stated that they find it 'completely' easy.	

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure facilitate the detection and identification of the LC, train or possible danger while the user is approaching the LC?**		
Answer modalities	N	The LC user's visual or auditory perception can be impeded/distracted by this measure
	0	This measure has no intended influence on the visual or auditory perception of the LC user
	1	
	2	
	3	
	4	
	5	LC users can easily detect the LC or the approaching train with sufficient time to stop or to cross safely (and continue to do so in the long term)
Score	5	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* The measure influences both the visual and auditory perception of the LC user, generating a warning when the driver approaches a LC. The warning is generated at appropriate distance before the LC, so that the driver has sufficient space to adjust vehicle speed and cross safely. A behavioural change on the short and long term is not expected. ………………………………………………………………………………………………

Criterion	Brief description
Rule knowledge	The measure can help the LC user elicit and retrieve relevant information about the required safe behaviour to cross the LC

Write down brief descriptions of the road user's ability to elicit and retrieve relevant safety information before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				Phase 1 Questionnaire relevant question: Q9. To what extent do the current safety measures at LCs in Thessaloniki help you to know how to cross safely? Q9: more than 41% of respondents answered 'slightly' or 'not at all'. Another 40.5% 'moderately' and only 18% 'considerably' or 'completely'.
After / With the measure			Phase 2 Questionnaire relevant question: Q9. To what extent does the in-car alert system help you to know how to cross LCs safely in Thessaloniki? Q9: 28.5% of respondents answered 'moderately', while 50% 'considerably' or 'completely'.	

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure evoke the required behaviour while the user is approaching the LC?**		
Answer modalities	N	The LC user is confused about how to behave safely at LC, because the measure transmits unclear or misleading information
	0	This measure has no intention to remind the LC user the required/safe behaviour
	1	
	2	
	3	
	4	
	5	LC users understand how to cross the LC safely without prior knowledge or experience of the LC type and environment in question (in all situations, also in the long term)
Score	1	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* The tested safety system does not primarily focus on improving LC crossing rule knowledge. However, the generation of the auditory and visual warnings might affect the drivers by reminding them of the potentially dangerous situation. In this way, a driver might respect the existing rules to a greater extent. A change on the short and long term is not expected. ………………………………………………………………………………………

Criterion	Brief description
Decision-making	The measure can help the LC user take more accurate decisions that arrive at safe behavioural intentions

Write down brief descriptions of the road user's decisions before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				Phase 1 Questionnaire relevant questions: Q10. How important is it for you to know how far away the train is from the LC? Q11. How important is it for you to know when the train will arrive at the LC? Those questions received very similar answers. More than 66% answered 'completely' and another 15% 'considerably' important. Less than 10% stated 'slightly' or 'not at all'.
After / With the measure			Phase 2 Questionnaire relevant questions: Q10. How likely it is that you would ignore the information provided by the in-car alert system (e.g. crossing after being alerted to an approaching train)? Q10: Almost half (46.43%) answered 'not at all'. Only 7.14% answered considerably and no driver chose the option 'completely'.	

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure facilitate the user's decision-making towards a safe course of action while approaching the LC?**		
Answer modalities	N	The LC user decides to cross unsafely, because this measure encourages their inaccurate subjective judgment of risk
	0	This measure has no intended influence on the subjective decision-making factors of the LC user
	1	
	2	
	3	
	4	
	5	LC users decide to cross the LC safety, because they understand the risks and the associated consequences of their behaviour (in all situations, also in the long term)
Score	4	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* The results of the 2nd phase questionnaire indicate that most drivers are expected to adjust their decisions towards the LC according to the warning message they receive from the safety system. A change on the short and long term is possible, as some drivers might initially not trust the provided information and therefore not make decisions accordingly. On the long run they are expected to trust the information, after witnessing that warnings are meaningful. …………………………………………………………………………………………

Criterion	Brief description
Behavioural execution	The measure can 'force' the LC user execute safe actions (required behaviours) or can impede the LC user from executing risky actions (non-adapted behaviours)

Write down brief descriptions of the road user's behavioural execution before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				Phase 1 Questionnaire relevant question: Q12. To what extent do you take risks at LCs (e.g. crossing after being alerted of an approaching train)? For Q12, 70.5% answered '*Not at all*'. Almost 21% answered '*slightly'* and '*moderately*'. Only 2.6% stated '*considerably'*. However, 7% answered '*completely'* Furthermore, spatiotemporal data about the vehicle kinematics (Floating Car Data) were recorded and analysed to study changes in the behavioural execution of drivers when they approach LCs before and after the measure. The data were curated and processed to form groups of datapoints representing vehicle trajectories through LCs. The data utilized for this analysis were recorded until 15th April 2019, around 2 active and protected level crossings were the active warning pop-up was available. 1846 test vehicles trajectories were identified in the data, 88 of which occurred during the baseline period. 1379 trajectories identified at LC with id=3 (latitude-longitude coordinates: 40.668589, 22.885852) and 76 at LC with id=1 (latitude-longitude

				coordinates: 40.6614383, 22.8961283). The datapoints are map-matched to the street network to calculate the distance to the rail with the minimum possible error. The vehicle trajectories around each LC are aggregated with respect to vehicle's moving direction, since a LC may be approached from two directions. The mean speed and mean acceleration curves as a function of distance to the LC were generated for each LC, direction and period (before and after the measure). The results do not indicate a behavioural change before and after the application of the safety measure. It is found that the typical speed profiles, although differentiated between different LCs and/or approaching direction, are quite similar when comparing a certain LC-direction combination before and after the safety measure. The last 50 meters of vehicle trajectories around LC with id 3 (with the most vehicle trajectories) were further analysed. The GPS pulses were grouped with respect to LC proximity, in 11 levels (0,5, 10,…50 meters distance to the LC respectively). At each proximity level, statistical tests were performed to check correlation between the dichotomous variable period (with values "baseline" and "after") and the continuous variables vehicle speed and acceleration. The appropriate method to examine the association between such types of variables is point biserial correlation, which is valid under several

				assumptions, two off which were violated in the dataset. Those assumptions are a) no outliers for the continuous variable for each category of the dichotomous variable; and b: the continuous variable should have equal variances for each category of the dichotomous variable. The first assumption, regarding the outliers, was checked by interpreting of boxplots, where outliers in data are outlined. The second assumption was checked by performing the Levene's test, according to which there is a difference between the variances in the two categories. Consequently, there are not sufficient evidence to support the hypothesis that the safety system had a significant change in the behavioural execution of drivers when approaching a LC. This result was, to some extent, expected, considering that the test vehicle drivers are highly experienced, professional taxi drivers who are aware of the locations of LCs and approach LCs with safety.
After / With the measure			*Phase 2 Questionnaire relevant question:* *Q11.To what extent do you take risks at LCs (e.g. crossing after being alerted to an approaching train)* Q11: Almost all drivers (92.6%) answered 'not at all' and therefore would not take risks around LCs, after being warned by the safety system.	

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure directly influence the safe execution of the approach and crossing behaviour?**		
Answer modalities	N	The ability of the LC user to cross safely is made difficult by this measure
	0	This measure has no intended direct influence on the LC user's execution of actions
	1	
	2	
	3	
	4	
	5	LC users are physically impeded from illegally crossing the LC or are forced to cross the LC safety when this measure is in place (also in the long term)
Score	1	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* The tested safety system does not primarily focus on improving on the behavioural execution. However, the generation of the auditory and visual warnings might affect user behaviour by reminding the driver of the potentially dangerous situation. A change on the short and long term is not expected. …………………………………………………………………………………………………

CRITERIA TO ASSESS THE USER EXPERIENCE AND SOCIAL PERCEPTION

<table>
<tr><th colspan="8">Choose the most appropriate answer by ticking one box for each case</th></tr>
<tr><th>Factor</th><th>Definition</th><th>(0)
Un-acceptable</th><th>(1)</th><th>(2)</th><th>(3)</th><th>(4)</th><th>(5)
Excellent</th></tr>
<tr><td rowspan="6">Accep-tance</td><td rowspan="2">The estimated level of acceptance by the public (e.g. road users, people living near the LC)</td><td>0
☐</td><td>1
☐</td><td>2
☐</td><td>3
☐</td><td>4
☒</td><td>5
☐</td></tr>
<tr><td colspan="6">Reasoning behind the score (indicate the findings or assumptions the score has been based on):

Acceptance of road users is high, according to their feedback from the 2nd phase questionnaire (Q 14) answers, where over 40% answered that they would be very interested in using the system after the end of the test period and only 10% would not be interested at all. Furthermore (Q 5), 90% of drivers generally feel at least slightly safer using the measure

These numbers and results are very promising, taking into consideration that they concern professional (taxi) drivers who are extremely experienced and know the area and LC locations very well. Less experienced drivers might accept the measure to a greater extent.</td></tr>
<tr><td rowspan="2">The estimated level of acceptance by relevant stakeholders (e.g. the railway operator, rail infrastructure manager, train drivers, authorities or Government)</td><td>0
☐</td><td>1
☐</td><td>2
☐</td><td>3
☐</td><td>4
☐</td><td>5
☒</td></tr>
<tr><td colspan="6">Reasoning behind the score (indicate the findings or assumptions the score has been based on):

The measure does not affect the operation of relevant stakeholders. Therefore, they are expected to accept the measure because safety at LC will increase without negative aftermath.</td></tr>
<tr><td rowspan="2">The estimated extent to which the measure can be integrated with the road and rail environment and with other safety measures</td><td>0
☐</td><td>1
☐</td><td>2
☐</td><td>3
☐</td><td>4
☒</td><td>5
☐</td></tr>
<tr><td colspan="6">Reasoning behind the score (indicate the findings or assumptions the score has been based on):

The measure is by nature integrated with the road and rail environment, provided that trains are tracked with geolocation devices and road vehicles use navigation software. Those requirements are commonly met, considering the current technology standards.</td></tr>
<tr><td>Reliability</td><td>The estimated extent to which the users of the</td><td>0
☐</td><td>1
☐</td><td>2
☐</td><td>3
☐</td><td>4
☒</td><td>5
☐</td></tr>
</table>

	LC trust the system and know that it is fail-safe	*Reasoning behind the score (indicate the findings or assumptions the score has been based on):* Drivers trust the information provided by the in-car alert system, according to feedback from the 2nd phase questionnaire (Q 12), as only 18% of drivers think that the provided information is 'slightly' or 'not at all' reliable.					
Usability	The estimated level of self-explaining nature of the design of safety measure (e.g. easy to understand or use) by all road users, all age categories and persons with various disabilities	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☒
		Reasoning behind the score (indicate the findings or assumptions the score has been based on): The safety system operates automatically without requiring user input (after it has been installed to the mobile device). The generated LC warnings have self-explanatory nature.					

ANNEX D: FINAL HUMAN FACTORS ASSESSMENT TOOL

This annex presents the final revision of the HFAT at the end of Task 2.2 and the SAFER-LC project.

Fill in the following forms for a given safety measure under evaluation. Each form is colour coded to reflect the three different sets of criteria under assessment: the 'Classification criteria' are included in a classification checklist (orange form). The 'Criteria to assess the behavioural safety effects' are included in five separate assessment sheets, one for each criterion (green forms). The 'Criteria to assess the user experience and social perception' are included in one assessment sheet (blue form). Detailed instructions are provided in the forms' headers.

Name of the measure being assessed	**Name of the pilot test and brief description of the tested measure**
....	....

CLASSIFICATION CRITERIA			
Factor	**Brief description**	**Indicator** *(Tick only the cases that the measure applies to, or click the option 'All' if the measure covers all the cases)*	
Applicability to different LCs	*Specify the types and characteristics of LCs where the measure can be implemented (multiple answers are possible)*	Type of LCs	☐ All ☐ Passive LCs without any warning device ☐ Active (manual) ☐ Active LCs with half barriers ☐ Active LCs with full barriers ☐ Active LCs with skirts for pedestrians ☐ Active LCs with light and sound warning ☐ Active LCs with traffic lights ☐ Active LCs with other warning device (specify)...
		Characteristics of LCs	☐ All ☐ LCs with low vehicle traffic ☐ LCs with high vehicle traffic ☐ LCs with paved road ☐ LCs with gravel road ☐ LCs with availability of electricity ☐ LCs with low usage ☐ LCs not used at all ☐ LCs with sharp / wide crossing angle ☐ Other (specify)................................
Feasibility under different environmental conditions	*Specify the environmental circumstances in which the measure aims to be most effective and which may affect the*	Time of the day	☐ All ☐ Daylight ☐ Darkness ☐ Dusk ☐ Dawn ☐ Peak traffic hours ☐ All

	perception or the behavioural adaptation of road users (multiple answers are possible)	Weather conditions Setting of the LC	☐ Rain ☐ Snowfall ☐ Slipperiness ☐ Fog ☐ Bright sunshine/ glare ☐ All ☐ urban ☐ rural
Applicability to different types of user	*Specify the categories of LC users who are targeted by the measure (multiple answers are possible)*	MRU VRU	☐ All ☐ cars ☐ motorbikes / mopeds ☐ trucks / heavy vehicles ☐ buses / coaches ☐ farm / agricultural vehicles ☐ other (specify)………………………….. ☐ All ☐ pedestrians ☐ cyclists ☐ other (specify)…………………………..
Adaptation to individual characteristics and conditions of users	*Specify if the measure is applicable for people with the following characteristics or conditions (multiple answers are possible)*	Gender Age Disability Under influence of Under skill impairing states	☐ Male ☐ Female ☐ All ages ☐ children ☐ elderly ☐ vision loss and blindness ☐ hearing loss and deafness ☐ intellectual disability ☐ reduced mobility ☐ other (specify)………………………….. ☐ alcohol ☐ drugs ☐ medication ☐ fatigue ☐ stress ☐ Risk-seeking personality
Intended effect mechanism	*Specify the mechanism via which the measure is expected to have an effect on safety (maximum 3 options can be ticked; <u>undeline the main</u> effect mechanism)*	☐ Supports LC safety actions ☐ Improves the detection of LC ☐ Reduces the approach speed of vehicles ☐ Controls access and supports egress from LC ☐ Increases awareness of correct behaviour and consequences of rule violation ☐ Improves the possibilities of vulnerable road users to cross the LC safely ☐ Improves the physical environment of LC ☐ Improves the detection of train ☐ Makes waiting time more tolerable ☐ Provides up-to-date information about the status of LC ☐ Other (specify) ……………………………………	

CRITERIA TO ASSESS THE BEHAVIOURAL SAFETY EFFECTS OF MEASURES ON ROAD USERS (SHORT- AND LONG-TERM)

Criterion	Brief description
Detection and Identification	The measure can help the LC user detect relevant visual and auditory stimuli and identify relevant information in the environment which can increase their detection of the LC, an approaching train or other potential danger

Write down brief descriptions of the road user's detection and identification of relevant LC safety information (e.g. detection of LC or train) before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure facilitate the detection and identification of the LC, train or possible danger while the user is approaching the LC?**		
Answer modalities	N	The LC user's visual or auditory perception can be impeded/distracted by this measure
	0	This measure has no intended influence on the visual or auditory perception of the LC user
	1	
	2	
	3	
	4	
	5	LC users can easily detect the LC or the approaching train with sufficient time to stop or to cross safely (and continue to do so in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* …………………………………………………………………………………………

Criterion	Brief description
Rule knowledge	The measure can help the LC user elicit and retrieve relevant information about the required safe behaviour to cross the LC

Write down brief descriptions of the road user's ability to elicit and retrieve relevant safety information before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure evoke the required behaviour while the user is approaching the LC?**		
Answer modalities	N	The LC user is confused about how to behave safely at LC, because the measure transmits unclear or misleading information
	0	This measure has no intention to remind the LC user the required/safe behaviour
	1	
	2	
	3	
	4	
	5	LC users understand how to cross the LC safely without prior knowledge or experience of the LC type and environment in question (in all situations, also in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* …………………………………………………………………………………………

Criterion	Brief description
Decision-making	The measure can help the LC user take more accurate decisions that arrive at safe behavioural intentions

Write down brief descriptions of the road user's decisions before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure facilitate the user's decision-making towards a safe course of action while approaching the LC?**		
Answer modalities	N	The LC user decides to cross unsafely, because this measure encourages their inaccurate subjective judgment of risk
	0	This measure has no intended influence on the subjective decision-making factors of the LC user
	1	
	2	
	3	
	4	
	5	LC users decide to cross the LC safety, because they understand the risks and the associated consequences of their behaviour (in all situations, also in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* ……………………………………………………………………………………………

Criterion	Brief description
Behavioural execution	The measure can 'force' the LC user execute safe actions (required behaviours) or can impede the LC user from executing risky actions (non-adapted behaviours)

Write down brief descriptions of the road user's behavioural execution before and after the measure (including any numerical findings from pilot tests or literature to support the estimated behavioural changes)				
Period	Evidence from literature		Evidence from pilot test	
	Short-term*	Long-term*	Short-term*	Long-term*
Before / Without the measure				
After / With the measure				

* Refer to the Application Guide for examples of what can be considered a short- and long-term change

Answer the following question by choosing one score between 0 and 5 or the answer 'N'. Make the choice based on the descriptions you gathered above. **Question: To what extent does the measure directly influence the safe execution of the approach and crossing behaviour?**		
Answer modalities	N	The ability of the LC user to cross safely is made difficult by this measure
	0	This measure has no intended direct influence on the LC user's execution of actions
	1	
	2	
	3	
	4	
	5	LC users are physically impeded from illegally crossing the LC or are forced to cross the LC safety when this measure is in place (also in the long term)
Score	…	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* …………………………………………………………………………………………

CRITERIA TO ASSESS THE USER EXPERIENCE AND SOCIAL PERCEPTION

Choose the most appropriate answer by ticking one box for each case							
Factor	**Definition**	**(0)** Un-acceptable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)** Excellent
Accep-tance	The estimated level of acceptance by the public (e.g. road users, people living near the LC)	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
	The estimated level of acceptance by relevant stakeholders (e.g. the railway operator, rail infrastructure manager, train drivers, authorities or Government)	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
	The estimated extent to which the measure can be integrated with the road and rail environment and with other safety measures	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
Reliability	The estimated extent to which the users of the LC trust the system and know that it is fail-safe	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					
Usability	The estimated level of self-explaining nature of the design of safety measure (e.g. easy to understand or use) by all road users, all age categories and persons with various disabilities	0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
		Reasoning behind the score (indicate the findings or assumptions the score has been based on):					