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

Deliverable D2.2

Human Factor methodological framework and application guide for testing (interim report)

Due date of deliverable: 31/07/2018

Actual submission date: 27/07/2018

Authors:

Grigore M. Havârneanu, UIC
Anne Silla, VTT
Sarah Whalley & Laura Lorenzo, FFE
Annie Kortsari, CERTH-HIT
Annika Dressler & Jan Grippenkoven, DLR

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 and application guide for testing (interim report)
Workpackage	WP2
Date of the document	24/07/2018
Version of the document	04
Responsible partner	UIC
Reviewing partner	Task 2.2 partners
Status of the document	Final
Dissemination level	Public

Document history:

Revision	Date	Description
01	28/05/2018	First draft for review by Task 2.2 partners
02	15/06/2018	Second draft for review by WP2 partners after a new round of inputs
03	29/06/2018	Third draft for review by all consortium partners
04	24/07/2018	Fourth and final version for submission

Consortium - List of partners

Partner No	Short name	Name	Country
1	UIC	International Union of Railways	France
2	VTT	VTT Technical Research Centre of Finland Ltd	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 first version of the Human Factors methodological framework which has been developed in the SAFER-LC project as part of Work Package 2. The purpose of this deliverable is to describe the theoretical background of the Human Factors methodological framework, including the sets of criteria selected for the Human Factors assessment tool, and to describe the Human Factors assessment tool and its application guide for testing. The overall objectives of this deliverable in the context of the wider workplan of the SAFER-LC project is described in Chapter 1.

The first part of the document sets out the theoretical foundations of the framework before going on to define the sets of evaluation criteria for self-explaining and forgiving level crossing design. Chapter 2 reviews and summarises the most important human factors and psychological models which provide theoretical foundations for the Human Factors methodological framework in the level crossing context. These models and theories were selected from the wider Human Factor and Traffic Psychology literature and refer to psychological aspects that can be manipulated through safety measures implemented at level crossings (e.g. perception, information processing, motivation, decision-making etc.).

Further, Chapter 3 shows how the Human Factors methodological framework builds on indicators adapted from relevant past research as transferrable lessons, such as classification and evaluation criteria, safety behavioural indicators, or clues on how to structure and organize the framework based on other assessment approaches used in road or railway safety context. The framework consists of a set of ‘Classification criteria’ as well as two sets of assessment criteria: ‘Criteria to assess the behavioural safety effects’ and ‘Criteria to assess the user experience and social perception’. Each of these criteria is based on the sets of factors and indicators which represent the backbone of the Human Factors assessment tool.

Each criterion included in the framework can be further broken down into more specific and measurable indicators. Chapter 3 describes how the selected sets of criteria can be operationalized through measurable indicators. The proposed sets of assessment criteria exploit maturity-type evaluation scales and Likert-type scales, which are used by the evaluators to estimate the effectiveness of the safety measure from a Human Factors perspective. The estimated extent and permanence of behavioural safety effects are defined according to the maturity scale from 0 to 5, and the level of agreement or disagreement regarding the user experience and social perception criterion is defined on a symmetric inadequate-excellent scale for a series of questions which also score from 0 to 5. This allows an aggregated overall quantitative estimation of how a safety measure implemented in a given setting is likely to perform according to relevant Human Factors criteria.

The second part of the document presents the Human Factors assessment tool and its application guidelines for field testing. Chapter 4 presents the assessment scales and Chapter 5 provides guidance on how to use the assessment tool during the pilot trials and illustrates this with a specific example.

The first version of the Human Factors assessment tool and its application guide will be used during the project trials to evaluate innovative measures aiming to improve the safety of level crossings from safety and human factors point of view. Most of the measures selected or developed within the SAFER-LC project will be tested and further developed under different environments in several test-sites (e.g. laboratory, driving simulator, living lab).

The information collected in the demonstration phase through the Human Factors assessment tool will allow the evaluation of the developed measures and the drawing of recommendations on human-centred improvements and organizational processes related to the evaluated measures. At the same time, based on the experiences gathered at the test-sites, the proposed HF assessment tool will be validated and improved at the later stages of the SAFER-LC project. For example, the evaluation scales proposed in the Human Factors assessment tool will be adjusted according to the feedback collected from the pilot sites and some criteria may be further refined or excluded. These issues will be addressed in the second part of Task 2.2.

Table of content

Abbreviations

Short name	Meaning
ADAS	Advanced Driver Assistance Systems
AOIs	Areas Of Interest
C-HIP	Communication-Human Information Processing (model)
CMM	Capability Maturity Model
COBIT	Control Objectives for Information and related Technology
CWA	Cognitive Work Analysis
CWA-DT	Cognitive Work Analysis Design Toolkit
D	Deliverable
DoW	Description of Work
DSS	Decision Support System
GEMS	Generic Error Modelling System
GIDAS	German In-Depth Accident Study
HF	Human Factor(s)
ITS	Intelligent Transport System
KPI	Key Performance Indicator
LC	Level Crossing
PMT	Protection Motivation Theory
PRM	Persons with Reduced Mobility
PTW	Powered Two Wheeler
RTH	Risk Homeostasis Theory
SEEV	Salience, Expectancy, Effort, Value (model)
SRK	Skills, Rules, Knowledge (model)
TTC	Time to collision
V2I	Vehicle to Infrastructure (exchange of information between vehicles and road infrastructure)
V2V	Vehicle to Vehicle (exchange of information between vehicles)
VRU	Vulnerable Road User
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 has a ‘Give Way’ sign 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 will 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 level-crossing infrastructure.

The solutions and tools that are being 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 level crossings, (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 services and advanced infrastructure-to-vehicle communication systems and on human processes to adapt infrastructure designs to road user needs and 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.

Within the project, the objective of Work Package 2 (WP2) is to enhance the safety performance of level crossing infrastructures from a human factor (HF) perspective, making them more self-explaining and forgiving.

1.2. General description of Task 2.2

The objective of Task 2.2 of WP2 is to develop a Human Factors methodological framework that evaluates the effectiveness of selected safety measures in terms of making level crossings more self-explaining and forgiving and therefore safer. The methodological framework includes a practical Human Factors assessment tool accompanied by an implementation guide to be applied in the evaluation of safety measures in the field.

The framework will allow the analysis and evaluation of the following types of measures (effectiveness for road and rail users):

- Non-technological: layout and design of LCs and different physical measures (e.g., angle of approach for road users, visibility of LC and trains, light or sound warnings at LCs, signage, ground markings, type of barriers, stopping distances).

- Technological safety measures (e.g., exchange of information between vehicles (V2V) and between vehicles and road infrastructure (V2I), vehicle-activated signage, in-vehicle Advanced Driver Assistance Systems; ADAS).

The framework consists of a set of evaluation criteria for self- explaining and forgiving design against which the safety measures will be objectively evaluated with the assignation of a score rating. The user perspective will encompass several of the following:

- Human behaviour analysis (focus on individual factors that influence safe/unsafe behaviour):
    - Factors concerning the type of user (motorized road user, vulnerable road user, age of user).
    - The usage frequency of level crossings (frequent vs. occasional user).
    - Knowledge related factors (actual meaning and subjective understanding of signage, knowledge of the level crossing, traffic rules, necessity to look to left and right).
    - Motivational factors (time pressure, other traffic participants, readiness to take risks).
- Involuntary unsafe behaviour (focus on errors and failures): distraction, fatigue, gaze and scanning habits of road users, ways to compensate for such failures (e.g. through forgiving infrastructure design or other measures).
- Voluntary unsafe behaviour (focus on violations): risk perception, understanding but ignoring the safety measures, determining when users make the final decision to cross, speed adaptation, etc.
- User acceptance of measures (focus on how users perceive the safety measures): usability of the different measures, understanding of the functioning of the measures (difficult to comprehend vs. self-explaining), user acceptance of novel warning devices, etc.

Train drivers are not in the main focus of the project. However, they are included in the methodological framework indirectly for example by analysing whether a measure implemented at a LC has an impact on the train drivers' actions, performance, etc. Therefore, only the indirect impact of the designed safety measures on train drivers or on rail operations is considered through the 'acceptance' indicator.

Similarly, suicides at or near the LCs are not the focus of the project. Railway suicide is acknowledged as a subcategory of 'intentional risky behaviour' but it does not receive special attention within the theoretical framework or the specific design of measures. Information available on this issue is reviewed and mentioned where relevant.

1.3. Interactions with other tasks and evolution within the project

The main inputs to Task 2.2 work from other SAFER-LC activities are coming from WP1 and WP2. Specifically, D1.3 (SAFER-LC consortium, 2018a) provides preliminary requirements and recommendations to be taken into account in the evaluation activities of Task 2.2. Task 2.1 constitutes 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 enable the identification of behavioural models of user-to-user and user-to-infrastructure interaction at level crossings.

The HF assessment tool presented in this deliverable will be tested within WP4 pilot activities through its application in the evaluation of tested safety measures which will be supported by an application guide. Task 2.2 has important ties with WP4 which includes two parts: (1) detailed test site descriptions defined in Task 4.1 (i.e. locations on where the measures can be tested) to be used as implementation guidelines in the process of testing the safety measures (D4.1; SAFER-LC consortium, 2018c), and (2) an evaluation framework proposed in Task 4.2 including a collection of Key Performance Indicators (KPIs) which are used to collect the data for the assessment (D4.2; SAFER-LC consortium, 2018d). These KPIs are 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 indicators subset (*human behaviour*) has been developed in close cooperation with Task 2.2. While the Human Factors KPIs developed in Task 4.2 represent a generic collection of indicators that can be used to assess the impact of safety measures on human behaviour, the aim of Task 2.2 is to specify a set of indicators adapted to the purpose of the Human Factors Assessment Tool. Further, WP4 will be relating all the KPI groups to the test site descriptions, specifying which indicators will be tested and where.

The first version of the SAFER-LC Human Factors assessment tool and its application guide (D2.2) will be used to evaluate the innovative solutions to enhance the safety of LCs during the project trials. Most of the piloted safety measures will be tested and further developed under different environments in several test sites. The various test sites available in SAFER-LC are a perfect fit for solutions and measures at different stages of maturity. Early stage developments can be tested in simulation environments or on controlled test tracks, while more readily developed measures will be evaluated in field experimentations.

The feedback collected in the demonstration phase through the HF assessment tool will allow the evaluation of the developed measures and recommendations to be made regarding technical specifications and human and organizational processes. At the same time, based on the feedback from the test sites, the HF methodological framework and assessment tool will be adjusted and improved (D2.5), for example by further refinement of evaluation criteria or by improving the issues related to the usability of the tool.

By the end of the project the HF methodological framework will result in a valid assessment tool of the ‘human’ component in the safety of LCs which will be included in the toolbox (WP6, Task 6.3, D6.7). Its inclusion in the toolbox will ensure that the human factor perspective will be accounted for within the management of safety at LCs by the road and rail infrastructure managers.

1.4. Purpose and structure of the document

This deliverable represents the first version of the framework that will be tested and further refined following its implementation in WP4 and consists of the following chapters:

1. Introduction
2. Theoretical background of the Human Factors Methodological Framework
3. Criteria selected for the Human Factors Assessment Tool
4. The Human Factors Assessment Tool for testing
5. Application Guide for Testing
6. Discussion and Conclusions

The document starts by setting out the theoretical foundations of the framework before going on to define a set of evaluation criteria for self-explaining and forgiving LC design. Chapter 4 presents the assessment tool that applies the criteria in the evaluation of the tested measures and the following Chapter 5 provides guidance on how to use the assessment tool in practice. Finally, we discuss the implications of this framework for the implementation and evaluation of user-centred safety measures at LCs and its planned validation scheduled for the second part of the project.

2. THEORETICAL BACKGROUND

The development of the HF methodological framework is based on existing publications, data sources, and analytical tools in the field of road and railway safety, traffic and transport psychology, and Human Factors research. The sets of relevant criteria are identified and selected based on rail human factors literature, published studies, and suitable approaches from related research projects. In other words, the framework is driven both theoretically and from applied studies and lessons learnt in practice.

The core of the HF framework relies on two innovative concepts for the LC context: the 'self-explaining' and 'forgiving' nature of the safety measures at LCs. The two concepts have been borrowed and adapted from road safety.

A self-explaining road is designed and built to evoke correct expectations from road users and proper driving behaviour (either due to its layout or through adequate signage), thereby reducing the probability of driver errors and enhancing driving comfort (Bekiaris & Gaitanidou, 2011). It includes the concepts of intuitive and understandable design, consistency, readability and psychological traffic calming (Van Geem et al., 2013). According to the European research project SafetyCube glossary (2018), roads are self-explaining when they are in line with the expectations of the road user, eliciting safe behaviour simply by design. The psychological processes of categorisation and expectancy are central to this concept. According to Theeuwes and Godthelp (1993), through experience road users will develop a prototypical representation with respect to different types of roads. This process is helped when the physical appearance of a specific road environment is homogenous and physically different from others. Having a clear categorization of the road traffic environment helps evoke correct expectations from road users and therefore it is crucial that the design of these roads is adjusted to these expectations (Van Geem et al., 2013). Motorways are a good example of a self-explaining road, as road users will clearly know when they are on a motorway and therefore will know what to expect (e.g. speed limits, lane positioning, where to expect signage, etc.).

A forgiving road is one that is designed and built in such a way as to interfere with or block the development of driving errors and to avoid or mitigate negative consequences of driving errors, allowing the driver to regain control and either stop or return to the travel lane without injury or damage (Bekiaris & Gaitanidou, 2011). In other words, a forgiving road is a road which is designed to help avoid driver errors from resulting in any serious injuries or a collision at all (SafetyCube glossary, 2018).

Similar definitions were already proposed in the LC context by the UNECE Working Party on Road Traffic Safety in their final strategic report (UNECE, 2017) which states that the next generation level crossings for pedestrians and vehicle usage should minimize the opportunities for human error and deliberate violations.

In the context of the SAFER-LC project, both concepts refer to the user-centric design of LC safety measures:

- 'Self-explaining' refers to the clear and good design of the safety measures implemented at the LC which supports adequate situation awareness (detection and perception of the situation; understanding the meaning of signs and measures; and projection of the current status into the future). The self-explaining nature of the LC design is therefore linked to the cognitive level of the LC user (easily perceived and understood by the user).

- 'Forgiving' means that the safety measures implemented at a LC include appropriate measures to counteract road user misbehaviour (e.g. error, violation, or deficient behavioural adaptation), and if a misbehaviour occurs, the system is able to mitigate the consequences. The forgiving nature of the LC design is therefore linked to the actual behaviour (action) of the LC user (easily compensates for misuse or misbehaviour).

2.1. Sociotechnical Systems Theory

The basis for developing and evaluating LC safety measures from the Human Factors viewpoint can be found in the Sociotechnical Systems Theory. Human Factors and Ergonomics use systems-based methods to support the design of complex and safe systems. An increasing number of researchers are supporting for the use of 'systems approach' when analysing and redesigning rail LC systems (e.g. Read et al., 2013; Salmon and Lenné, 2015; Stefanova et al., 2015). The advantage of the systems approach is that it considers all the relevant components within a LC context and the complex interactions between these elements: level crossing users (e.g. pedestrians, older drivers); vehicles (e.g. heavy vehicles, rolling stock); level crossing infrastructure (e.g. sight distances, signage); and the broader environment (e.g. weather conditions) (Searle et al., 2012). This is important as road users with different individual characteristics interact with the various technologies in different LC environments. Countermeasures adopted through the safe systems approach seek to make the characteristics of level crossings more forgiving of human error, and to minimise the level of unsafe road user behaviour (Searle et al., 2012).

2.1.1. Focus on level crossing environments

There are a number of external factors related to the environmental context of a level crossing which can influence the safe use of crossings (voluntary or involuntary). Task 2.1 studied the following aspects of the level crossing context and their importance in terms of safety:

- Level crossing type
- Level crossing setting
- Infrastructure layout
- Weather conditions
- Time of day
- Traffic volume

An overview of the factors identified as most important according to the results of Task 2.1 (i.e. in terms of prevalence in the literature and its evaluation by experts) are presented below, together with supporting evidence from the literature and a reflection on their relevance according to the type of level crossing (passive versus active).

Level crossing type

The relevance of environmental factors may depend on the type of level crossing; some human factors are relevant to all types of level crossings whilst others may apply exclusively to passive or active crossings. For example, having good visibility of the crossing and tracks is essential at passive level crossings as it determines one's ability to check for the presence of trains and therefore to know when it is safe to cross. On the other hand, at an active crossing the road user will be warned of the presence of a train through the closing of the barrier or activation of light and/or sound signals.

Another issue related specifically to the type of crossing is deliberate risk-taking behaviour. There is a clear relationship between deliberate risk-taking behaviour and active level crossings, due to the fact that passive crossings have no safety controls to violate. This situation seems to be particularly prevalent at crossings with barriers, with risk-taking often taking place due to the user's frustration and impatience at having to wait. The length of the warning time is found to be important, with some studies detecting that violations tend to increase significantly when the time between warning activation and train arrival exceeds 20–30 seconds (Coleman & Venkataraman 2001; Richards & Heathington 1990, cited in Searle et al., 2012).

Road users may also deliberately violate crossing controls if they consider them to be unreliable, possibly due to previous experience of right-side failures (signalling failure when the LC provides warning without a reason). Indeed, frequent or prolonged failures of this type may cause road users to lose confidence in the warning which can in turn influence their driving performance. This situation potentially facilitates the creation of mental models of when the train approach warning is credible based on other factors such as known train schedules, resulting in a mismatch between real risk and perceived risk (Rongfang Liu, 2010).

Another issue found in the literature in relation to the type of crossing is the user's understanding of the difference between passive and active level crossings. A number of studies referenced in Searle et al. (2012) have pointed to the fact that many drivers do not look for trains at passive crossings, suggesting they lack awareness of the nature of passive crossings and difference between protected and non-protected crossings.

Level crossing setting

The level crossing setting is considered to be a particularly relevant environmental factor affecting safety at level crossings. There are features of the level crossing setting that can impact on the conspicuousness of the crossing and trains, and most notably the sight distances. For example, sight distances can be obstructed by trees, buildings, and the roadway-crossing geometry (Caird et al., 2002). As indicated above, poor sight distance and impediments to level crossing visibility are of particular importance at unprotected crossings where the decision to cross safely depends on one's ability to detect an oncoming train and having a safe time margin for stopping.

Physical elements in the surroundings of the level crossing, such as traffic lights and give way signs belonging to neighbouring streets, pedestrian traffic, shops etc. can act as external distractors to level crossing users' attention. When overloaded with other stimuli, the situational awareness of the road user can be compromised and attention is taken away from the level crossing. In this situation, stimuli such as trains or flashing lights may be fully visible but unnoticed, a phenomenon referred to as 'attentional blindness', or 'looked but failed to see' (Searle et al., 2012). The "looked-but-failed-to-see" error describes a situation in which drivers fail to identify hazards despite looking at the hazard's source, due to limitations in human information processing. The factor "faulty activation of schemas" is characterized by learned misbehaviour and maladaptive expectations, resulting from the infrequent occurrence of a train at many level crossings (Grippenkoven & Dietsch, 2015). On the other hand, a potential inattention issue that can be experienced at passive crossings is low states of arousal and inattention to the broader environment, due to the rural isolation of these type of crossings and their low train and road traffic, meaning that the user may fail to notice either crossings or approaching trains (Searle et al., 2012).

Infrastructure layout

A related factor and one that was also found to be relevant in D2.1 (SAFER-LC consortium, 2018b), is the level crossing infrastructure layout. Its importance for safety particularly relates to its effect on visibility, train detection and perception of speed and distance of another vehicle. The driver's ability to detect the presence of a train is fundamental in influencing their decision to stop at a passive level crossing. Rudin-Brown et al. (2014) proposed four perceptual human factors that could contribute to a driver's ability to detect the presence of a train: sightlines; train conspicuity; unchanging retinal image, and train horn audibility. For pedestrians, awareness of a train depends on their ability to look and listen, or otherwise search, for trains and their associated warning devices. At 'passive' crossings there are no specific warnings when a train approaches, so adequate sight distance must be provided for crossing users to see approaching trains (Edquist et al., 2011).

Weather conditions

According to the results of D2.1 (SAFER-LC consortium, 2018b) the consideration of weather conditions (e.g. ice, snow, hard wind, etc.) can be important in terms of impact on the functioning and maintenance of level crossings. However, weather conditions can also influence the user's ability to detect a train. For example, weather-related problems such as fog, sun glare, and blowing snow may obscure a train or reduce the distance at which the train can be viewed in time to stop (Caird et al., 2002). In addition, weather conditions can influence the possibilities of road users to cross the LC safely (e.g. the driver might have difficulties to stop the car on time before the LC if the road is slippery because of ice or heavy snow).

Time of day

Time of day can also be a useful indicator of safety when referring to lighting conditions and impact on visibility. For example, at night, drivers' judgment is usually worse, because they can find it difficult to compare train movement against a dim background with indistinct landmarks. The time of day may also have a relation with different road volumes, traffic density, rush hour related hurry and fatigue, etc.

Traffic volume

The literature reviewed in D2.1 (SAFER-LC consortium, 2018b) identified traffic volume as a variable frequently included in the study of level crossing safety. One study found that annual daily road traffic and annual daily train traffic are significant predictors of injury accident frequency (Borsos et al., 2016). Findings from the survey reported in D1.1 (SAFER-LC consortium, 2017) also point to the importance of road and rail traffic volumes as criteria in deciding level crossing protective arrangements.

One of the issues detected in the literature is that user expectations regarding daily train volumes can influence the degree of attention paid at a crossing. For example, crossing familiarity and an expectation that a train will not be present have the potential to lull drivers into complacency or poor looking habits (Caird et al., 2002). This can be particularly true for frequent passive crossing users due to the typically low daily train volumes at this type of crossing. In their study of unintentional non-compliance at rail level crossings, Salmon et al. (2013) echo this view and propose that rail bodies undertaking LC risk assessments should consider crossings with low train volumes as a risk factor together with psychological failures of the users such as schema and Looked-but-failed-to-see errors.

Another issue related to road and rail traffic volume may be the increased level of impatience experienced by some users when delayed at level crossings during rush hour traffic (Caird et al., 2002) and as such a greater propensity for risk-taking.

2.1.2. Focus on vulnerable road users

The following section provides a reflection on the different level crossing user groups drawing on information studied within the D1.1 (SAFER-LC consortium, 2017) and D2.1 (SAFER-LC consortium, 2018b). In general terms, level crossing users can be divided into two key groups: motorised and non-motorised road users who can be further distinguished as vulnerable or non-vulnerable.

According to the European Commission (2011), vulnerable road users are defined as "*non-motorised road users, such as pedestrians and cyclists as well as motor-cyclists and persons with disabilities or reduced mobility and orientation*". The European Road Transport Research Advisory Council (ERTRAC, 2011) broadens this definition to: "*those participants in traffic that are not protected by any mechanical system: pedestrians, motorcyclists, bicyclists, and users of mopeds. This includes road users with impairment, e.g. using a mobility aid, or children playing on the road. Car occupants, even when this refers to impaired people, senior people or children, do not belong to the category of VRU according to this definition*".

The literature review conducted as part of Task 2.1 identified different level crossing user categories. These user groups are presented in Table 1, classified under either motorised road user (MRU) or vulnerable road user (VRU). Please note however that motorbike and moped riders, classified as MRU in the table, are often referred to as VRU, as reflected in the EU Intelligent Transport Systems (ITS) Directive (Directive 2010/40/EU).

An analysis of the prevalence of these user groups within the literature shows that *car users* are the most frequently studied group followed by *pedestrians*, *cyclists* and *heavy vehicle* users. According to the results of an in-depth LC accident analysis conducted in Task 1.2, the victims in LC accidents are most often car occupants (drivers or passengers) or pedestrians (Silla et al., 2017). These findings are also reflected in the results of D1.1 (SAFER-LC consortium, 2017) where survey responses from 24 countries most frequently reported the existence of safety arrangements related to motorised road users (transport professionals, heavy vehicles and farm vehicles). In terms of vulnerable users: pedestrians, cyclists and persons with reduced mobility (PRM) are the most widely recognised, both in terms of their representation in the literature and having specifically targeted safety measures. One response specifically highlighted the importance of age (older users) together with disability, due to the potential effects of these on the fitness or cognitive abilities of the user. Young users together with users with different cultural and language background (e.g. refugees), have also been highlighted as being at greater risk, due to language barriers or the lack of knowledge of traffic and level crossing safety rules. Such groups of VRUs are reflected in the choice of pilot sites in WP4 (e.g. Turkey). Altogether, four categories of users with a disability have been studied in D1.1 (SAFER-LC consortium, 2017) and D2.1 (SAFER-LC consortium, 2018b): persons with reduced mobility (PRM); vision loss and blindness; hearing loss and deafness; and persons with disability in general.

Table 1. Type of level crossing user variables identified in the Task 2.1 literature review (SAFER-LC consortium, 2018b).

Motorised road user (MRU)	**Vulnerable road user (VRU)**
Car	Pedestrians
Heavy vehicle	Cyclists
Motorbike / moped*	Vulnerable age groups (e.g. children, the elderly)
Transport professional	Persons with reduced mobility (PRM)
Farm vehicles	Horse riders
	Person with vision loss and blindness
** Please note that motorbike and moped riders are also often defined as vulnerable road users.*	Person with hearing loss and deafness
	Person with different cultural and language background
	Person with disability in general

2.1.3. Cognitive Work Analysis

The Cognitive Work Analysis (CWA) framework (Vicente, 1999) has emerged as a promising approach for supporting the design of safe LC systems. The literature review reported in D2.1 (SAFER-LC consortium, 2018b) conducted on 125 publications found that of those publications applying analytical models (n=14), Cognitive Work Analysis (CWA), was the most commonly cited approach. CWA refers to a 5-step analysis which can also be successfully applied in the LC context:

1. **Work Domain Analysis:** describes the environmental constraints on behaviour (e.g. LC type: passive/active, LC context: urban/rural, etc.)
2. **Control Task Analysis:** describes the needed decisions and tasks (e.g. decision ladder, information queues, etc.)
3. **Strategies Analysis:** identifies various strategies to fulfil the tasks (e.g. violations at active LC, personal motivations, habits due to exposure over time, etc.)
4. **Organisational Analysis:** refers to the social organisation, cooperation, division of work, allocation of functions between humans and technology (e.g. technology or design compensates for the human errors at LC, etc.)
5. **Competencies Analysis:** refers to the skills required by the actors operating within the domain (e.g. individual differences and capabilities of different users of LC, VRUs, cultural differences, differences between frequent and infrequent LC users, etc.)

The recently proposed Cognitive Work Analysis Design Toolkit (CWA-DT) (Read et al., 2016) provides guidance and tools to assist in applying the outputs of CWA to design processes to incorporate the values and principles of sociotechnical systems theory (see for review, Walker et al., 2008) in rail transport settings such as LCs. For example, the following indicators are included in the CWA-DT: the view of humans as assets, the view of technology being a tool to assist humans, the promotion of the quality of life for rail LC users, the respect for individual differences, and upholding the responsibility to all stakeholders.

Implications for the HF framework:

From these principles, one can derive more specific criteria of LC design and safety measures such as: the suitability for all kind of road users (MRUs as well as VRUs), acceptability by involved parties, or the degree to which LC infrastructure is self-explaining. In addition, the type of the LC and environmental context of the LC are very important (e.g. passive/active, rural/urban etc.).

However, the application of CWA uses tools such as the abstraction hierarchy, contextual activity template or decision ladders which are not always easy to use by those undertaking the analysis. The results of CWA are likely to vary according to the specific LC environments being analysed. Taken together with the fact that a full CWA requires a lot of resources, there are limits to its practicability in the face of the existing diversity of LC. Still, the approach offers a broad methodological toolkit from which approaches fitting the most important research needs can be chosen. In summary, the main weakness of CWA is its limited practical applicability to variety of contexts. For an easier applicability in LC settings, CWA can be used as a very general theoretical framework which structures the analysis on 5 levels which can be enriched with elements derived from other models and theories of human cognition and behaviour.

2.2. Human cognition and behaviour theory

This section reviews and summarises the most important human factors or psychological models which provide additional theoretical foundations for the HF methodological framework in the LC context. These models were selected from the wider Human Factors and Traffic Psychology literature to enrich the mainstream CWA approach with aspects that can be manipulated through safety measures implemented at LC. Some of these models were identified as valuable enhancements in the literature review which was initiated in Task 2.1 (D2.1; SAFER-LC consortium, 2018b).

The models described below bring additional insights on:

- Cognitive aspects (e.g. information processing; attention and distraction; perception; task performance and errors),
- Motivational aspects (e.g. risky decision-making; action choices; habit and routines; behavioural intentions), and
- Behavioural aspects (e.g. risk taking, violations, behavioural adaptation linked to risk compensation).

2.2.1. Errors and violations

From classical accident research, collisions at LCs can be linked to errors of perception, knowledge or decision-making (Grippenkoven & Dietsch, 2016; Lobb et al., 2001; Ward & Wilde, 1995). Errors have been defined as acts where the subject intends to follow the rules, yet the actions deviate from this intention. Most approaches from a cognitive psychology perspective define errors in terms of a specific step within a sequence of information processing steps. The benefit of adopting this approach is that by identifying the critical step in information processing the error can help to gather insights into the weaknesses of a system design (Grippenkoven et al., 2012). For example, in the LC environment a road user may fail to see the warning lights because of fatigue, inattention, poor lighting, limited sight distance, etc. (Freeman et al., 2013). Further well-known definitions of errors as well as in-depth classifications of errors can be found in the work of Hollnagel (1993a, 1993b), Reason (1990) or Rasmussen (1982).

An error categorisation framework can support understanding of the underlying mechanisms of human error in level crossing accidents. Reason's generic error modelling system (GEMS) is a classification scheme that focuses on cognitive factors in human error. Reason's model defines three basic error types: slips, lapses and mistakes. The first two, slips and lapses, are skill-based errors which result from some failure in the execution stage of an action sequence (Reason, 1990). Slips refer to errors of attention and emerge when a correct plan is applied incorrectly, whereas lapses involve memory failures or lack of attention leading to omissions of necessary actions (Grippenkoven et al., 2012; Reason, 1990). Whilst these errors are unintended, the third type of error, mistakes, are a result of an intended action. Mistakes can be defined as failures in planning and/or judgemental processes and occur at the planning stage of information processing. They can be divided into two kinds: rule-based mistakes (involving the misapplication of a good rule or application of a bad rule) or knowledge-based mistakes (due to incomplete/inaccurate understanding of system, confirmation bias, overconfidence) (Dian et al., 2011; Grippenkoven et al., 2012; Reason, 1990). According to Reason they are normally caused by either a failure of expertise or lack of expertise.

Similar to Reason's classification, Rasmussen's model of internal human malfunction (Rasmussen, 1982) is based on the assumption of distinct stages of human information processing, each of which can lead to the emergence of specific errors. In his categorisation, he differentiates seven types of human errors that enable the rater to identify the critical step of information processing which in turn leads to the occurrence of an error. These errors are ordered as structural, information, diagnostic, goal setting, strategy selection, procedure and action errors (Grippenkoven et al., 2012).

The GIDAS (German In-Depth Accident Study; Graab et al., 2008) framework for human error categorisation draws on both of these models. Given its development within the road accident domain, this framework is of particular interest as it can be applied to the analysis of human errors in level crossing accidents on the part of road traffic users. The framework maintains the sequential procedure of human information processing, with the following categories: information access, information admission, information evaluation, planning and operation. It should be noted though that the information processing is not only sequential, but also entails feedback loops from "later" stages to "earlier" ones (e.g. if I do not expect a danger I will not look out for it and have a lesser chance to detect it). Each of the categories are assigned general influences and specific indicators that comprise the cause of human error. This GIDAS error categorisation is presented in Table 2.

Table 2. GIDAS error categorisation (Graab et al., 2008).

Error category	**Description of influence**	**Indicator**
Information access	Relevant information cannot be perceived	E.g. glare, parking vehicles that cover signs, foggy weather
Information admission	Interfering information and influences in and outside the car	E.g. crying children, fatigue, focus on other car
Information evaluation	Information interpreted in the wrong way	E.g. lack of experience, underestimation of speed or distance
Planning	Violations of the rules, wrong decisions	E.g. bypassing active barriers
Operation	Wrong actions taken	E.g. confusing controls

Another class of unsafe acts described by Reason in his GEMS are violations (Figure 1). These are different from errors because they manifest through deliberately faulty actions. These unsafe actions are associated with an intention to deviate from regulations, rules and procedures, although the person has no intention of injury (Reason, 1990).

A recent Australian study into the origins of rule-breaking at pedestrian train crossings has shown that 24.5% of respondents intentionally violated the rules (Freeman & Rakotonirainy, 2015). Besides the fact that at least a fraction of the errors can go unnoticed by the LC user, these findings suggest that unsafe behaviour at LCs is likely to be driven not only errors, but also by violations. Several studies have also shown that some drivers are willing to ignore active warnings and some found it exciting to "beat the train" (Abraham et al., 1998; Meeker et al., 1998; Witte and Donohuie, 1998 cited in Yeh & Multer 2008). Unsafe behaviour resulting from violations may have different motivational roots and is likely to be associated with different personal, socio-cultural, and environmental variables (Lobb, 2006). It is therefore important to differentiate between these motives because the potential countermeasures or designs should also depend on the nature of the motivational context. Summala (1997) emphasized that external motives influence the level of risk individuals are willing to take (e.g., time pressure, mood). For example, when in a bad mood, drunk or simply in a hurry, drivers are prepared to compromise. This implies that external motives influence peoples' willingness to take risks.

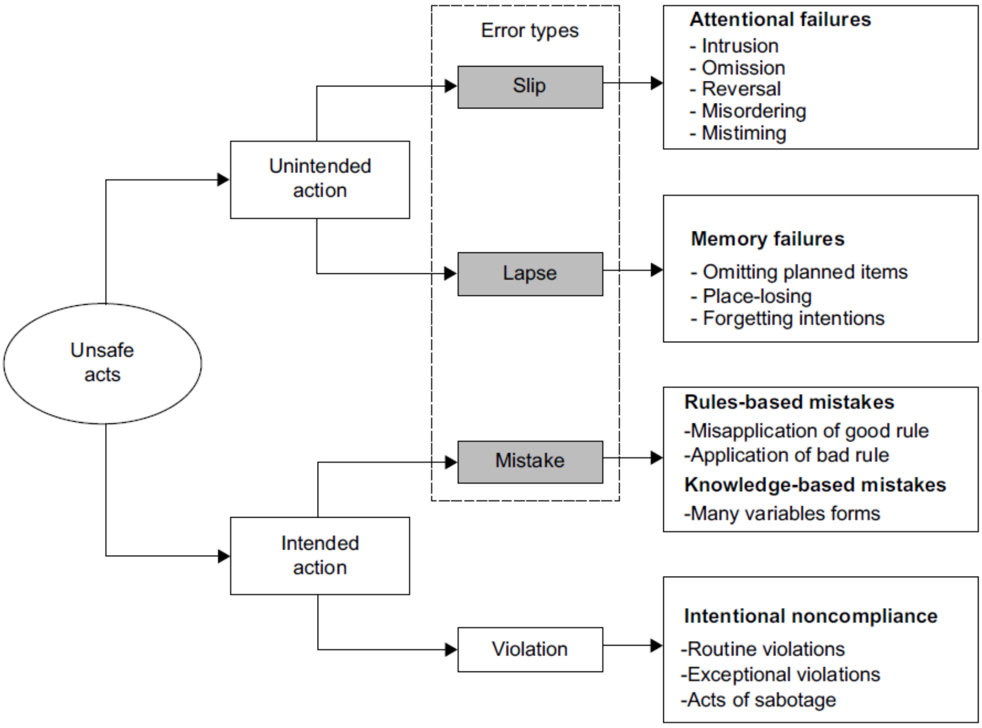

Figure 1. Classification of unsafe acts into errors and violations (adapted from Reason, 1990).

This is also in line with the more general literature of psychology and social sciences where there is widespread agreement that behaviour is influenced by its perceived benefits and, even more, by its perceived costs (e.g. Kahneman & Tversky, 1979). Indeed, several theories of behaviour at level crossings propose that the decision to violate level crossing rules is the result of a cost-benefit analysis, where the perceived benefits of committing a violation outweigh its potential costs (Fambro et al. 1995; Yeh & Multer 2008 cited in Searle et al., 2012). This means that the subjective discomfort caused by time loss can outweigh the perceived benefits of safe waiting behaviour, with factors such as expectancies regarding the likelihood of a train, familiarity with the crossing and perceived credibility of the warning devices being weighed up. According to a report by Lerner et al. (1990) of particular concern are drivers who are risk-takers and simply accept higher levels of risk. Social pressure, from either peers or other drivers, also plays a role in increasing risk-taking behaviour (Yeh & Multer, 2008).

Laapotti (2016) classified the immediate risk factors based on the in-depth analysis of LC accident data. These findings differentiate between different types of errors and other individual risks linked for example with particular individual motivations:

- Observation error: A driver failed to see or hear an oncoming train or did not pay attention to warning signals or sounds
- Anticipation or evaluation error (does not include deliberate risk-taking): A driver failed to anticipate danger or evaluated a situation wrongly when approaching or driving through a level crossing. For example, the driver approached the level crossings with too high speed and was unable to stop in time, or the driver misinterpreted the speed of the oncoming train or misunderstood the warning signals
- Vehicle handling error: A driver handled the vehicle incorrectly. For instance, the driver selected the wrong gear when driving through a level crossing, thereby stalling the car on the tracks
- Other human-related risk factors: A driver took a deliberate risk or drove on purpose in front of an oncoming train (suicide)
- Vehicle risk factors: A sudden technical malfunction or breakdown of the vehicle

Implications for the HF framework:

The inclusion of motivational aspects in the human factor analysis is important especially since the criteria on motivation, habits and systematic violations as voluntary unsafe behaviours are theoretically interrelated.

Whilst both errors and violations should be considered, one should bear in mind that violations are mostly if not only relevant at active LCs.

2.2.2. Models of human information processing

The C-HIP model (Communication-Human Information Processing; Wogalter, 2006) was originally developed to describe human processing of warnings. It combines a classic communication model (Source-Channel-Receiver) with specifications of crucial stages of human information processing where the message must successfully pass within the receiver. The newest version also specifies Delivery of information as a process between Channel and Receiver, in which information can get lost or changed because of other environmental factors (e.g. persons, objects, lighting, noise etc.).

It is essentially a framework of human information processing and behaviour which includes the following aspects:

- Source
- Channel
- Delivery (and possible influences of environmental factors)

- Receiver (and its individual factors)
    - Attention switch
    - Maintenance of attention
    - Comprehension and Retention
    - Attitudes and Beliefs
    - Motivation
- Behaviour

Relevance for LC context: The C-HIP model provides a checklist of conditions that must be fulfilled for successfully changing a person's behaviour by an explicit warning or other kinds of signal. However, the factors and processes influencing information processing at each stage must be more closely defined, using other theories. In addition, these elements could theoretically contribute to phases 1 (Work Domain Analysis), 2 (Control Task Analysis) and 3 (Strategies Analysis) of the CWA.

Similar to C-HIP, a second **Model of Human Information Processing** at LC was proposed (Grippenkoven, 2017). It is essentially an information processing model because it includes the necessary stages of human information processing in the encounter with LC infrastructure. However, compared to C-HIP, this conceptualization is closer to the behavioural aspects, as it brings forward the idea of behavioural intention and behavioural adaptation of users at LCs. It sums up some elements that are separated in C-HIP (e.g. Detection, including attentional processes that are co-determined by properties of the signal source and channel, or 'formation of behavioural intention' that includes the influence of beliefs and motivational issues). The importance of having useful knowledge of correct behaviour is stressed somewhat more explicitly than in C-HIP as well as the possibility that behaviour could still go wrong even if correct intentions are formed.

It is essentially a framework of human information processing and behaviour which includes the following aspects:

1. Detection of LC infrastructure
2. Information processing and understanding
3. Retrieval of knowledge from long-term memory
4. Behavioural intention formation
5. Execution of action

Relevance for LC context: As with the C-HIP model, it is useful in providing a checklist of processes that can succeed or fail in generating adaptive behaviour at LC and that can be influenced by the design of LC and/or its surroundings. The focus on behavioural intention points to the important link with behavioural theories such as the Theory of Planned Behaviour. The factors and processes influencing information processing at each stage can be defined in more detail using other theories. In addition, these elements could theoretically contribute to phases 2 (Control Task Analysis) and 3 (Strategies Analysis) of the CWA.

2.2.3. Models of attention

The SEEV model (Salience, Expectancy, Effort, Value; Wickens & McCarley, 2008a) explains and predicts the allocation of visual attention (gaze) to areas of interest (AOIs) in a scene. The gaze depends on certain objectives, expectations, environmental factors, and the effort necessary to access a given piece of information.

It is essentially a model of visual attention allocation which includes the following aspects:

- Habit (individually learned ways of scanning for information, e.g. "look left-right-left" before crossing a road) *[not directly included in the model, but mentioned as a factor]*
- Salience (power of AOI to capture attention, depending on onsets, perceptual salience and attentional set) *[S in "SEEV", for "Salience"]*
- Information content: Bandwidth (expected event rate in a given AOI - the lower, the fewer fixations) *[first E in "SEEV", for "Expectancy"]*
- Information Content: Context (expected momentary availability of information in the given AOI [e.g. driven by cues] or momentary task demands) *[first E in "SEEV", for "Expectancy"]*
- Information Access Effort (Effort necessary to sample information from the given AOI, largely defined by the distance of AOI from currently fixated AOI with three zones:
    1. no scan [no eye movement necessary]
    2. eye field [eye movement necessary but sufficient]
    3. larger head field [head or even body movement necessary] - other factors aggravate subjective effort as in LC the need to reduce speed) *[2nd E in "SEEV", for "Effort"]*
- Information Value (subjective utility of having the information from the given AOI, including utility potentially lost by omitting a scan – can be multiplicated with bandwidth in style of an expectancy*value model) *[V in "SEEV", for "Value"]*

Relevance for LC context: The SEEV model can inform the evaluation and development of measures to sustain desirable attention allocation (and help identify and avoid psychological obstacles to that) in the approach to LC environments. The model represents one possible specification of the processes at the "attention switch" and "attention maintenance" stages of the C-HIP model and the "detection" and "information processing" stages of the LC behaviour model. In addition, these elements could theoretically contribute to phases 3 (Strategies Analysis) and 5 (Competencies Analysis) of the CWA.

The Multiple Resources model (Wickens & McCarley, 2008b) explains the possibility and quality of timesharing (i.e. how well can two or more activities requiring attention be carried out at the same time), depending on a number of available attentional resources and the way they are required by simultaneous tasks.

It is essentially a model of attention division and allocation which refers to independent resources for:

- Different stages of information processing: "early" (Perception and Cognition) and "late" (Responding)
- Different sensory modalities: visual and auditory (other modalities, e.g. haptic, could potentially be included)
- Different representational codes used in perception/cognition: Spatial (images, positions etc.) and Verbal (language)
- Different areas of the visual field: Focal (fovea, fixated area) and Ambient (retinal periphery) - only within visual modality

Relevance for LC context: The model can be used to inform the evaluation and design of countermeasures with respect to avoiding overload and maximizing the chances of the relevant messages being processed by the user. In addition, these elements could theoretically contribute to phases 2 (Control Task Analysis) and 5 (Competencies Analysis) of the CWA.

The Attentional-Gate Model (Block & Zakay, 1996) explains how individuals estimate the waiting time. It states that duration judgments largely depend both on arousal level and on the amount of attention allocated to time. An increased arousal level (e.g. when you are in a hurry and you have to wait at a closed barrier of a LC) may lead the pacemaker to produce more pulses in a given time unit; thus, the waiting duration subjectively appears even longer than it is. Furthermore, the model states that the amount of attention allocated to temporal cues is inversely related to the attentional demands of a concurrent task. In other words, if people are doing something else or attending something else at a LC, fewer pulses pass the gate and the experienced time interval should be smaller. If workload during a task is high, less attention is focused on time and therefore more clock time is needed in order to reach the target time. Based on this model, in order to effectively reduce subjective time intervals, more cognitive engagement is needed to subjectively "take time from the clock".

It is essentially a model of human time perception which includes the following aspects:

- Pacemaker that autonomously produces pulses at a rate that is influenced by arousal
- When a person attends to time, as opposed to external stimulus events, the attentional gate opens, and the pulse stream is sent to subsequent components.
- A cognitive counter accumulates a pulse count, which is transferred to a working memory store.
- When duration has ended the switch closes and the accumulated pulse total is sent to the reference memory store.
- The accumulated pulse total from the working memory is compared with various durations, which have been previously stored in the reference memory.
- If fewer than the criterion number of pulses have accumulated, this cognitive comparison results in the person judging that the duration is shorter than what is required (e.g. for a task).

- When the pulse count in working memory is approximately equal to that in reference memory, the person judges that the duration is appropriate for some response.
- Retrospective time estimation widely depends on the amount of contextual changes that happened during encounter in a time interval. More changes lead to longer time estimations.

Relevance for LC context: Along with other subjective perceptions, time estimation plays an essential role during the waiting period in front of an active level crossing protected by half barriers. If road users experience the waiting time as too long, they are prone to violations that subsequently may lead to dangerous situations. In addition, these elements could theoretically contribute to phases 2 (Control Task Analysis) and 3 (Strategies Analysis) of the CWA.

2.2.4. Skill-related Human Factors models and hierarchical behavioural models

Human Factor skill-based models can basically be categorized as either person models (e.g., the generic error modelling system by Reason (1990), focusing on the errors made at an individual level; see section 2.2.1 above) or system models, focusing on the interaction between wider systematic failures and errors made by an individual operator.

According to **the Swiss Cheese model (Reason, 1990)**, safety focuses on the interaction between latent and active conditions/failures within and between layers (i.e., culture and climate) and unsafe acts and their contribution to accidents. Each layer can include its own defences (i.e. countermeasures) against unsafe acts. An accident occurs when there is a synchronised set of failures between various layers of the system. Safety is therefore the responsibility of actors at all layers and levels of the system, which should include effective defence barriers (see Figure 2).

Relevance for LC context: The Swiss Model is important because it highlights the systemic view in accident occurrence and the need of achieving redundancy between several safety layers. If one countermeasure fails at one level of the system, another one should come into action at a different level. This is in line with our approach to the ‘forgiving infrastructure’ which concerns measures able to compensate for the individual unsafe acts. In addition, this model supports the CWA by considering several layers of possible failures which partially correspond to the levels of analysis proposed by the CWA. For example, both theoretical perspectives highlight the importance of individual preconditions and capabilities (phase 5 – competencies analysis) but also the importance of organisational issues (phase 4 – organisational analysis).

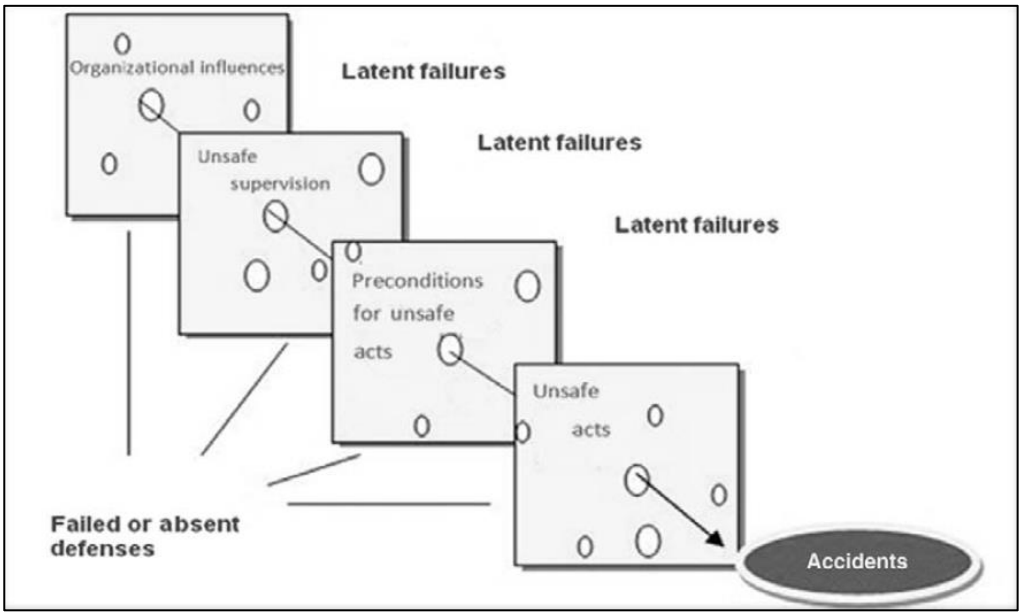

Figure 2. "Swiss cheese" model (adapted from Reason, 1990).

The three-level model of behavioural control also known as **the Skills, Rules, Knowledge (SRK) framework (Rasmussen, 1983, 1986)** describes three different modes of generating behaviour, depending on task characteristics and training level. The model also relates to human error taxonomies, which themselves are related to research on human reliability as the engineering approach to expanding concepts of technical reliability and barriers to human factors (see Sharit, 2006, for summary).

It is essentially a hierarchical descriptive model of operational levels of behaviour and defines three different modes of generating behaviour (Figure 3):

- Knowledge-based: operation mode in situations for which no learned rules or automatized reactions are available, involves most cognitive effort.
- Rule-based: involves the application of learned rules for a given situation. Requires having (learnt) an adequate set of rules and correctly recognizing the "signs", i.e. the conditions that define the "if"-part of a rule. Requires some attentional resources.
- Skill-based: involves highly trained, automatic behavioural programmes that are executed nearly without requiring attentional resources in response to certain stimuli.

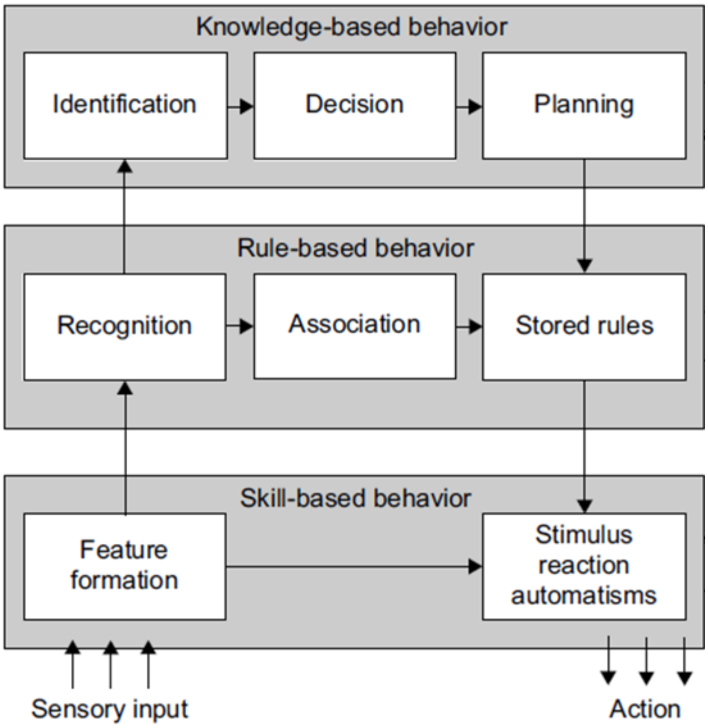

Figure 3. Combination of performance levels according to Rasmussen (1986).

Relevance for LC context: One useful feature in the LC context is that it reminds the importance of implicit, highly automatized processes in human behaviour (skill-based) that are hardly covered by models that involve conscious processing (attention, comprehension, beliefs etc.). It also implies that the best LC design solutions will be those that do not require a lot of knowledge or explicit thinking to elicit the desired response. In case this cannot be achieved, and processing has to go up to the rule-based level. This would stress the importance of conveying clear rules on how to behave at LC in each situation and puts emphasis on the designing for easy recognition of the respective situations. It also points to the problems occurring when rules for different situations are not equally practiced, and to the fact that the existence of a wide variety of rules might be confusing and thus it would be desirable to reduce the number of needed rules as much as possible. The SRK model lays the theoretical ground of phase 2 (Control Task Analysis) of the CWA, by defining three types of information processing which range from fully conscious to fully automatic.

The three-level model of driver behaviour (Michon 1985) divides the information processing in the driving task into three levels of skills and control: strategic, tactical, and operational control of car driving. The information needed for decision-making is different at each of the three levels; examples are described in the following:

- The strategical level concerns the general planning of a trip, including route finding and selection to achieve goals such as to minimise the time to reach the destination, to avoid a traffic congestion or to avoid bad weather conditions etc.

- The decision-making on the tactical level includes interaction with other road users and environment. At the tactical level drivers exercise manoeuvre control according to the prevailing circumstances and the decision-making occurs in a couple of seconds and is based on perceptions, estimation of distances and velocities, and anticipation of traffic situations. The decisions include e.g. speed choice, gap acceptance, and interaction with other road users.
- The operational level decisions are automatic and they occur in milliseconds. These decisions cover vehicle control, i.e. the use of steering wheel, pedals and any assistance system in the car.

Relevance for LC context: In the LC context, the tactical level and its components are probably the most interesting for the HF framework. In addition, these elements could theoretically contribute to phases 2 (Control Task Analysis) and 5 (Competencies Analysis) of the CWA.

The Model of Intuitive Use (Naumann et al., 2007) is essentially a characterization of "intuitive use" and proposes several levels:

- The first and lowest level of the continuum consists of innate knowledge. Generally, this is what reflexes or instinctive behaviour draw upon. Purists will see this as the only valid level of knowledge when talking about intuitive interaction, because it assures universal applicability and non-conscious processing.
- The next level is sensorimotor. It consists of general knowledge, which is acquired very early in childhood, and is from then on used continuously through interaction with the world. Scientific notions like affordances reside at this level of knowledge.
- The next level is about knowledge specific to the culture an individual lives in.
- The most specific level of knowledge is expertise that is specialist knowledge acquired in one's profession. Across the sensorimotor, culture and expertise levels of knowledge, we also distinguish knowledge about tools. Tool knowledge seems to be an important reference when designing user interfaces.

Relevance for LC context: This model could serve as a heuristic for the intuitive quality of a solution, tool or countermeasure designed for the safety of LCs. Measures intuitive at an innate or sensorimotor level are more promising in terms of their effectiveness than others, since no cultural background or professional knowledge is needed. On the sensorimotor level, it could furthermore be worth reflecting on the psychological concept of "affordances" (as defined by Norman et al., 2007). This model is also a theoretical support for the 'self-explaining' concept which is central to the HF methodological framework. In addition, this model can theoretically enrich phase 3 (Strategies Analysis) of the CWA, by making the link between good usability and the motivation to use the system.

2.2.5. Risk theories and models of risky decision-making

Zero-risk theory (Näätänen & Summala 1976, Summala 1988) deals with the risk adaptation and argues that due to human perceptual, cognitive, and motivational processes, drivers adapt to risks on the road and are hence motivated towards faster speeds and objectively more risky behaviour. According to the zero-risk theory, the drivers have a target risk of crashing of zero, which is achieved by attempting to maintain the level of task difficulty within the target boundaries. This means that when drivers get into a situation which they assess as risky they tend to eliminate this risk immediately by certain behavioural changes (e.g. by reducing the driving speed), and in the future they tend to avoid such situations which elicit fear or may lead to fearful situation. However, with repeated confrontations, drivers adapt to situations which at first elicited a 'risk response' and thus they drive the majority of the time with overlearned habitual patterns based on safety margins, with no concern for risk.

Novice drivers, for example, initially feel a sense of uncertainty or fear in many traffic situations, but these feelings disappear with increased experience. With experience, the novice drivers acquire greater control and fluency of driving and their driving becomes a habitual, increasingly automatized activity in which risk control is based on maintaining safety margins. Because of risk adaptation, drivers cannot rationally take traffic risks into account and speed limits, for example, become an important safety measure.

Relevance for LC context: The theory highlights the importance of behavioural adaptation in time and the fact that some safety measures become obsolete in the long run because drivers get used to them and to the level of safety they bring. Based on this theory, frequent level crossing users are a population at higher risk since they can manifest a more prominent behavioural adaptation based on their experience and familiarity with specific LCs. Therefore, safety measures implemented at LCs should be able to achieve a long-term effect and should also be able to target frequent users of LCs. This theory is also essentially a motivational theory which underlines the importance of individual motivations in taking decisions. For example, time pressure to reach the destination very quickly can act as a risk factor at active LCs with long waiting times. In addition, this model can theoretically enrich phase 3 (Strategies Analysis) of the CWA, by bringing forward individual factors such as motivation and experience in risky decision-making.

Risk homeostasis theory (Wilde 1982, 1986, 1994, 1998) postulates a concept of target risk, which is determined by the expected costs and benefits of people's behaviour. Target risk is the level of accident risk at which the individual believes that the net benefit of his or her action is maximized. Therefore, the persons are trying to optimize the risk instead or minimizing it. Specifically, this theory proposes that people accept a particular level of subjectively assessed risk to their health and safety to gain a range of benefits associated with that activity. Whenever a person perceives a discrepancy between target risk and experienced risk in one or other direction, he or she will make some behavioural adjustment to restore the balance (e.g. drive faster or slower). Consequently, this theory proposes that the basic strategy of injury prevention should be to reduce the level of risk that people are willing to accept.

In the field of Traffic Psychology, the theory also addresses the impact of engineering and technology on driving behaviour. It postulates that with the implementation of more and more safety measures, drivers will perceive a higher level of safety (therefore a lower level of risk) which will encourage them to accept more risk and take more risks.

Relevance for LC context: The risk homeostasis theory brings forward the concept of “perceived risk” which is influenced by subjective factors including the perceived costs and benefits of the actions but also by the technological innovations implemented at the vehicle or infrastructure level. The risk homeostasis theory can theoretically enrich phase 2 (Control Task Analysis) of the CWA, by pointing to additional factors which can influence the user’s decision. Thus, it helps in identifying unsafe decisions at LCs and approaches to modify the decision.

Protection Motivation Theory (PMT, Maddux & Rogers, 1983) helps in understanding behavioural decisions under uncertainty. It is especially suitable for situations in which there are two behavioural options: 1) the behaviour is safe but associated with some immediate subjective costs or effort (e.g. having to put on a helmet, having to wait, having to slow down, etc.) and 2) the behaviour that is risky, but can be executed right away, without much (immediate) cost or effort. For understanding the (irrational) subjective weighting of perceived risk of the maladaptive behaviour and response cost of the adaptive behaviour see also findings in Neth, Sims and Gray, 2006 (which apply not only to gains, but also to the avoidance of losses).

PMT is essentially a model of decision in a risk context and includes the following aspects (*the brackets with +/- indicate the direction in which the respective variable is assumed to influence protection motivation and thus adaptive behaviour*):

- Appraisal of maladaptive behaviour ("threat appraisal"):
    - Perceived Severity of potential adverse consequences (+)
    - Subjective Vulnerability to adverse consequence (probability of consequence occurring) (+)
    - Intrinsic + Extrinsic Rewards associated with maladaptive behaviour (-)
- Appraisal of adaptive behaviour ("coping appraisal"):
    - Perceived Response Costs (*effort* [mental, physical, time, financial, emotional, social etc.] to be invested to execute adaptive behaviour) (-)
    - Perceived Response Efficacy (power of adaptive behaviour to eliminate threat) (+)
    - Perceived Self-Efficacy (level of perceived control to personally execute the adaptive behaviour - generally and in the respective situation) (+)

Relevance for LC context: The PMT identifies factors that influence the decision ("cognitive mediating processes"). Thus, it helps in identifying unsafe decisions at LCs and approaches to modify the decision. It can be considered as one possible specification of important processes at the "attitudes and beliefs" and "motivation" stages of the C-HIP model and the "formation of behavioural intention" stage of the LC behaviour model. In addition, the PMT can theoretically enrich phase 2 (Control Task Analysis) of the CWA, by pointing to additional factors which can influence the user’s decision.

2.2.6. Synthesis of the reviewed models and theories

In the previous sections, a series of models and theories have been presented. As pointed out, their focus is variable, and therefore they contribute in different ways to the development of the HF methodological framework by pointing out various individual factors that are likely to shape the road user's behaviour at LCs.

Our theoretical review revealed a set of models focusing on human skills with a particular interest in the performance linked to specific individual capabilities. For example, the **Model of Human Information Processing** provides a checklist of processes that can succeed or fail in generating adaptive behaviour at LC: (1) detection of LC infrastructure, (2) information processing and understanding, (3) retrieval of knowledge from long-term memory, (4) behavioural intention formation, and (5) execution of action. This process identifies five antecedents of the user's behaviour, and every phase can be influenced by the LC design.

Further, **the Multiple Resources model and the SEEV Model** give more insight on the human cognitive capabilities such as attention and perception. They point out that a countermeasure should avoid cognitive overload and maximize the user's attention allocation and maintenance by simulating different sensory modalities (e.g. visual, auditory, haptic), by using different representational codes (e.g. spatial and verbal) or by using different areas of the visual field (e.g. central and peripheral). **The C-HIP model** provides a checklist of conditions that must be fulfilled for successfully changing a person's behaviour by an explicit warning or other kinds of signal (e.g. environmental conditions to deliver the message, attention switch and maintenance, comprehension, individual attitudes and motivations, etc.). In addition, the **Model of Intuitive Use** suggests that countermeasures that target the innate or sensorimotor level of the user could be more effective than others since no cultural background or professional knowledge is needed to detect and understand them. It also provides theoretical support for the 'self-explaining' concept.

Other models are purely descriptive but help understand the different levels where human failure can occur leading to errors or violations. For example, **the Swiss Model** highlights the layers in the whole system and suggests that a measure or combination of measures should be able to 'safeguard' several layers, providing theoretical support to the concept of 'forgiving infrastructure'. Complementary to this, the **SRK model** zooms in at the individual level showing that errors can occur at the skill-, rule- or knowledge levels which range from fully automatic to fully conscious information processing. It suggests that the LC design should not require a lot of knowledge or explicit thinking to elicit the desired response and should have in-built easy recognition. The **hierarchical model of Michon** points to important actions that drivers can take at LCs: at the tactical level drivers behave based on perceptions, estimation of distances and velocities, and anticipation of traffic situations. The decisions include e.g. speed choice, gap acceptance, and interaction with other road users. At the strategical level they will consider their planning of the trip, including route finding and selection to achieve goals such as to minimise the time to reach the destination, to avoid a traffic congestion or to avoid bad weather conditions etc.

Risk theories focus even more on the individual level of analysis but shift the focus from skill and task performance to the style of behaviour and behavioural adaptation. Thus, they allow an in-depth understanding of the factors that shape behaviour through deliberate and fully conscious decisions. For example, **the zero-risk theory** highlights the importance of individual motivations in taking decisions (e.g. time pressure to arrive at the destination faster), or the behavioural adaptation of users based on their experience and familiarity with specific LCs. The issue of subjective time perception is further considered in **the Attentional-Gate Model.** Since time estimation plays an essential role during waiting period in front of an active level crossing protected by half barriers, if road users experience the waiting time as too long, they could be prone to violations. **The risk homeostasis theory** brings forward the concept of "perceived risk" which is influenced by subjective factors including the perceived costs and benefits of the actions but also by the technological innovations implemented at the vehicle or infrastructure level. It also argues the continuous behavioural adaptation of user to implemented safety measures and challenges the long-term effect of measures. The **Protection Motivation Theory** brings further insight with respect to the users' unsafe decisions at LCs through their attitudes, motivations and subjective appraisals of the costs and benefits of their actions.

Implications for the HF framework:

Overall, the models and theories reviewed above propose a set of factors that can be considered at specific levels of the CWA. They highlight the importance of: (1) the main individual capabilities which shape the road user's performance at a LC; (2) the hierarchical behaviour precursors where errors can occur (e.g. skills, rules, knowledge); (3) the factors influencing the subjective risk perception and risky decision-making; and (4) the individual and external motivational factors affecting the risky behaviour at LC and the behavioural adaptation in the long term.

The HF methodological framework will include these four issues as further enhancements of the levels of analysis proposed in the CWA (according to the Figure 4 below). In our framework, the traditional Work Domain Analysis from the CWA is replaced by the 'Analysis of the LC physical domain and environmental constraints'. These conditions (e.g. LC type: passive/active, LC context: urban/rural, etc.) create a defined context that may have specific constraints on behaviour.

The next level is the 'Organisational analysis of social and societal issues' which refers to the harmonisation between the road and rail cooperation, public acceptance of the implemented safety measure, the ability that technology or design compensates for the human errors at LC, etc.

The Control Task Analysis is replaced by the 'Activity analysis of LC approach and crossing'. This level incorporates the hierarchical behaviour precursors where errors can occur (e.g. skills, rules, knowledge) as well as the needed decisions and tasks (e.g. information queues, information processing, etc.). The decision-making process is influenced by the subjective risk perception among other factors.

The 'Mental strategies to perform the crossing' is the level of analysis that considers various strategies to fulfil the crossing tasks according to the individual and external motivational factors affecting the risky behaviour at LC and the behavioural adaptation in the long-term (e.g. systematic violations at active LC, habits due to exposure over time, etc.), lack of trust in the system, etc.

Finally, the traditional Competencies Analysis is assimilated to the 'Individual skill (abilities) of LC users'. These are the main individual capabilities which shape the road user's performance at a LC.

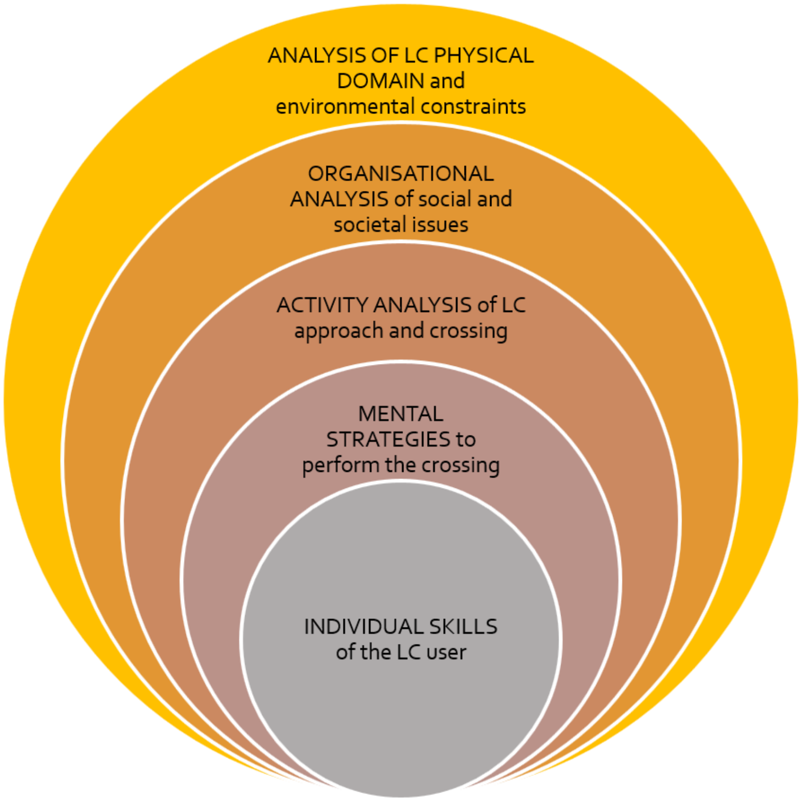

Figure 4. An adaptation of the five levels of analysis proposed in the CWA adapted to the LC user.

On this wide theoretical basis, the SAFER-LC Human Factors methodological framework partially builds on the sets of criteria described in Table 3.

Table 3. Theoretically driven indicators for the evaluation of level crossing safety measures, taking into account the road and rail users' perspective.

Indicator	Definition	Examples and possible quantification	Related models/theories
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	Model of Human Information Processing; Multiple Resources model; SEEV; C-HIP; Model of Intuitive Use; Swiss Model; SRK; Hierarchical model of Michon; Zero-risk theory; Attentional-Gate Model; RHT; PMT
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	Model of Human Information Processing; Multiple Resources model; SEEV; C-HIP; Model of Intuitive Use; Swiss Model; SRK; Hierarchical model of Michon
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	Zero-risk theory; Attentional-Gate Model; RTH; PMT
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	Zero-risk theory; Attentional-Gate Model; RHT; PMT
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)	Zero-risk theory; Attentional-Gate Model; RHT; PMT
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)	Model of Human Information Processing; Multiple Resources model; SEEV; C-HIP; Model of Intuitive Use; Swiss Model; SRK
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	Multiple Resources model; Model of Intuitive Use

2.3. Analytical tools and empirical approaches driven from related research projects

The SAFER-LC Human Factors methodological framework also builds on indicators adapted from relevant past research as transferrable lessons, such as classification and evaluation criteria, safety behavioural indicators, or clues on how to structure and organize the framework based on other assessment approaches used in road or railway safety context. The studies considered and their main contributions to the SAFER-LC HF methodological framework are described in the following subchapters.

2.3.1. EU project RESTRAIL

The EU project RESTRAIL (REduction of Suicide and Trespass on RAILway property) used 14 criteria to assess the most cost-effective measures to prevent rail suicides and trespassing (Ryan & Kallberg, 2013). These criteria were derived from those used in previous EU research (e.g. SUPREME project, Elvik 2006), but have been adapted for use in this rail related context. The criteria include RAMSHEC principles (Reliability, Availability, Maintenance, Safety, Health, Environment, Cost) (Jovanovic & Zoeteman 2010) which are core components of rail infrastructure business. Some of these RAMSHEC principles are obvious within the main evaluation criteria, whilst the remaining are subsumed within other evaluation criteria.

The list below includes an overview of these RESTRAIL criteria accompanied by a brief description:

1. **Definition of the measure** provides a description of relevant features of the measure.
2. **Definition of target incidents** answers the question of what kinds of incidents the measure is intended to reduce. It refers to specific types of incidents, but can focus on a specific group of people, e.g. school children.
3. **Size of the problem** provides a quantitative estimate of the frequency of target incidents (e.g. trespassing accidents in the target group per year).
4. **Effect on incidents** means the expected effect (in per cent) on target incidents (as defined above). The effect in absolute number of incidents can then be calculated by multiplying this estimate by "size of the problem".
5. **Durability of effects** concerns the durability of the effects on target incidents: are they likely to remain fairly stable or is there reason to believe that they will erode with time.
6. **Costs and benefits** should provide approximate estimates of the costs and benefits, if available. A more detailed cost-benefit analysis will be conducted for a limited number of most promising measures that will be identified in the first group evaluation.
7. **Integration with other policy measures** describes how the measure is integrated with other preventative measures or interventions.
8. **Impact on railway operations** means the positive or negative effect on the running of trains.

9. **Impact on people and jobs** means especially the effects on the health and jobs of people within railway industry (e.g. the number of staff in different job categories and changes in the roles of people) but also elsewhere if relevant.
10. **Technological issues** concern changes in the existing technology and infrastructure caused by the implementation of the measure, including the readiness of technology for new interventions.
11. **Environmental issues** concern impacts on the environment in general (e.g. different kinds of pollution, impacts on scenery and wildlife).
12. **Acceptance** provides an estimate of how well the measure is accepted by the public and relevant stakeholders (e.g. policy makers, industry).
13. **Transferability issues** concern the functionality of the measure in different environments and in different scales (e.g. is it likely that the effects are different in different countries or depend of the scale of the implementation).
14. **Additional information** can be any relevant information that is not addressed within the issues listed above, e.g. notes on the strengths, weaknesses, opportunities and threats concerning the conducted evaluations.

The RESTRAIL methodology which used these criteria was successful in achieving the project goals, namely to identify cost-effective measures to prevent rail suicides and trespassing (Ryan et al., 2018).

2.3.2. Evaluation framework of Silla et al. (2015)

The assessment methodology developed in the RESTRAIL project was adapted and applied 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. The 'new' variables included 'Family of measures', 'Effect mechanism', 'Type of measure', 'Feasibility to different LCs', and 'Circumstances in which the measure is the most effective'. The variables 'family of measures' and 'effect mechanism' were already discussed and applied in RESTRAIL project even though these criteria are not included in the above list of criteria.

The short descriptions of the additional (new) variables are presented in the following:

- **Family of measures** which refers to different groups of measures aiming to improve the safety of LCs. Family of measures included the following categories: 'Warning devices', 'Warning lights', 'Traffic signs and marking of LCs', 'Improvement of the detection of approaching train', 'Changes to the road leading to LC', 'Risk analysis', 'Campaigns and education', 'Legislation and specifications', and 'Other measures'
- **The effect mechanism** specifies the type of impact expected with the intervention and can typically include subcategories referring to various preventive layers with completely different goals:

- Improve the conspicuousness of train (colouring of the head of the train, LC mirrors, lighting systems etc.)
    - Improve the conspicuousness of LC (active warning signs etc.)
    - Restrict the access to LC (barriers)
    - Reduce the approach speeds of vehicles (rumble strips, speed bumps, road swivelling etc.)
    - Increase the awareness of correct behaviour and dangerousness of LCs (information campaigns, education etc.)
    - Improve the physical environment of LC (wait platforms, inclination, maintenance etc.)
    - Improve the possibilities of vulnerable road users to cross LC safely (rubber pads between the rails, barriers for VRUs, gates etc.)
    - Provides up-to-date information about the status of LC (navigator, smart phone etc.)
    - Supports the LC safety actions (accident modelling, risk analysis etc.)
- The different **types of measures** were grouped according to the following categories including some examples of measures related to some of the categories: 'Technical, low-tech (e.g. active warning devices)', 'Technical, high-tech (e.g. in-vehicle warning system)', 'Structural (Speed bumps, barriers, portals etc.)', 'Traffic signs and signals', 'Campaigns and education', 'Tool, method, practice', 'Other'.
- **The feasibility to different LCs** specifies the types of LCs where the measure can be implemented and includes a detailed classification of LC types (passive LCs without any warning devices, active LCs with barriers, active LCs with light and sound warning, active LCs with other warning devices, active LCs with traffic lights, 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, other).
- **The circumstances** where a specific measure is the most effective can include particular environmental conditions affecting perception or behavioural adaptation such as daylight, darkness, twilight, rain, snowfall, slipperiness or poor visibility due to weather (fog).

The study of Silla et al. (2015) described the issues affecting the safety of level crossings and surveyed and systematically assessed measures aiming to improve it. The results of the assessment process enabled the authors to list the most effective and promising measures for improving the safety of level crossings in Finland in the future.

2.3.3. Ex-ante assessment method of Kulmala (2010)

Furthermore, the SAFER-LC Human Factors methodological framework is inspired by an 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). During the application of the method, the assumptions on possible effects of ITS applications are made transparent to enable the validation of the results. The validation could be done e.g. in the field operational tests focussing on measuring driver behaviour.

The assessment framework of Kulmala (2010) follows the generally accepted theoretical background, according to which the traffic safety consists of three dimensions, which are (1) exposure, (2) risk of a collision taking place during a trip, and (3) consequences (= risk of a collision to result in injuries or death) (Nilsson, 2004). In order to ascertain that all possible impacts (both positive and negative impacts on road safety; direct, indirect and unintended effects of systems) are covered, and no effects are counted twice, the approach of Kulmala (2010) exploits a set of nine mechanisms via which ITS can affect road user behaviour and hence road safety. These nine mechanisms systematically cover the three aspects of road safety and are based on a ten-point list compiled by Draskóczy et al. (1998):

- Mechanism 1: Direct modification of the task of road users by giving information, advice, and assistance or taking over part of the task
- Mechanism 2: Direct influence by roadside systems mainly by giving information and advice
- Mechanism 3: Indirect modification of user behaviour in many, largely unknown ways
- Mechanism 4: Indirect modification of non-user behaviour
- Mechanism 5: Modification of interaction between users and non-users
- Mechanism 6: Modification of road user exposure by for example information, recommendation, restrictions, debiting or increased comfort in car driving, Powered Two Wheeler (PTW) riding, cycling or walking
- Mechanism 7: Modification of modal choice by for example demand restraints (area access restriction, road pricing, area parking strategies), supply control by modal interchange and other public transport management measures, and travel information systems
- Mechanism 8: Modification of route choice by route diversions, route guidance systems, dynamic route information systems, and hazard warning systems monitoring incidents
- Mechanism 9: Modification of accident consequences by intelligent injury severity reducing systems at crashes, by quick and accurate crash reporting and call for rescue, and by reduced rescue time

The most relevant criteria from Kulmala (2010) for the SAFER-LC HF methodological framework are the direct and indirect modification of road user behaviour (mechanisms 1, 2 and 3). The direct effects refer to short-term changes in road user behaviour which appear in few milliseconds or seconds, whereas the indirect effects refer to the long-term changes in road user behaviour which involve learning processes and experiences leading to this behavioural adaptation.

Direct modification of road user (including driver) behaviour aims to describe the direct effects of the implemented safety measure (in-vehicle information or roadside systems) on the driving task or road user behaviour by giving information, advice, and assistance or taking over part of the task. This may influence the road user’s attention, mental load, and decision about the performed action (for example, road user’s choice of speed). These effects are direct consequences of the safety measure which means that they are direct reactions to the system outputs and appear in few milliseconds or seconds. The direct (short-term) modification of road user behaviour covers both intended (e.g. decrease in speed to avoid a collision) and unintended (e.g. road user distraction) impacts.

Indirect modification of road user (including driver) behaviour is more long-term than the very direct, short-term reactions to the system. The road users will always adapt to the changing situation. This is often called behavioural adaptation and will often not appear immediately after a change (i.e. implementation of the safety measure) but may show up later and it is commonly very hard to predict. Long-term behavioural adaptation may appear in many different ways (for example, by reallocation of attention resources, by a change of expectation of the behaviour of other road users, or by risk compensation). This adaptation may often be due to the delegation of responsibility of the current task partly or totally to the system, which the road users have learnt to rely on. This can occur, for example, when road users receive warnings of approaching dangerous situations via their mobile phone or other applications. The road users learn to rely on this information and thus might observe their surroundings less carefully.

2.3.4. Main inputs from related previous assessment studies

One should note that many of the criteria used in the previous studies do not concern Human Factors. Moreover, it should be noted that some of the criteria used in the earlier studies are related to the cost-benefit analysis (CBA), socio-economic assessment, or other issues that fall out of the scope of Task 2.2, but which will be considered in other SAFER-LC tasks. However, several of these criteria (especially the ones related to the assessment of safety effects) take into account HF issues and can be further adapted in the SAFER-LC methodological framework. These features refer to categorical variables and are therefore more qualitative in nature.

Table 4 presents the overview of criteria adapted for SAFER-LC HF methodological framework and its assessment tool from the previous road and railway safety related assessment projects. The criteria from previous studies were divided into three main categories: 'Classification criteria', 'Criteria to assess the short- and long-term behavioural safety effects', and 'Other assessment criteria'.

Some of the criteria presented in Table 4 have not been discussed until now in the LC context in this deliverable. Therefore, the following list provides a short description of these criteria adapted from the RESTRAIL project (Ryan & Kallberg, 2013) and the evaluation framework of Silla et al. (2015) for the purposes of the SAFER-LC HF methodological framework and their relevance in the LC context. In addition, 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 in-vehicle warning system for railway level crossings.

- **Target of safety effects** specifies the categories of users who are targeted by the measure. The safety measures can be targeted to all road users, specifically either to motor vehicle drivers or vulnerable road users (e.g. pedestrians and cyclists) or to some specific user groups such as children or professional drivers.
- **Reliability of the system** estimates the degree to which 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). This is an important topic when assessing the effect of ITS applications on driver behaviour and can be relevant for LCs as well, especially when innovative or unusual measures are implemented bringing a high degree of novelty to the public.

- **Integration with road and/or railway environment or other safety measures** identifies if the implementation of a measure is expected to have a positive or negative effect on the existing context: e.g. it will not create any problems, only minor problems may occur that may be easily solved, or major problems may occur.
- **Acceptance** provides an estimate of how well the measure is accepted by the public and relevant stakeholders. For example, it can be assessed with the categories: 'No major resistance is expected', 'Resistance is possible (including an explanation on possible groups affected)', and 'Resistance is probable (including an explanation on possible groups affected)'. One could quantify acceptance (from total acceptance to total resistance) referring to all important categories of users and involved stakeholders: LC users, railway staff, people living nearby etc. Acceptance is also closely related to the level of trust, therefore it has close ties with the reliability criterion.

Table 4. Overview of the assessment criteria adapted from previous studies.

Category	Criterion	Definition	Source
Classification criteria	Effect mechanism	Specifies the type of effect mechanism (impact) expected with the intervention	Ryan & Kallberg, 2013; Silla et al., 2015
	Feasibility to 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 & Kallberg, 2013; 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 the short- and long-term behavioural safety effects	Short-term effect on road user behaviour	Describe the direct effects of the implemented safety measure on road user behaviour based on the strategic, tactical and operational levels 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
Other assessment criteria	Reliability of the system	Estimates if the users trust the system and how they know that it is fail-safe	Öörni et al., 2011
	Integration with road/railway environment, other safety measures	Describes how the measure is integrated with the road/rail environment other preventative measures or interventions	Ryan & Kallberg, 2013; Silla et al., 2015
	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 & Kallberg, 2013; Silla et al., 2015

Implications for the HF framework:

The findings and criteria adopted from relevant assessment studies 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 and Kallberg (2013) and Silla et al. (2015), while the 'Criteria to assess the short- and long-term behavioural safety effects' were motivated by Kulmala (2010) and 'Criteria to assess user experience and social perception' were inspired by Ryan and Kallberg (2013), Silla et al. (2015) and Öörni et al. (2011).

As these were adapted into the HF methodological framework, two out of three sets of criteria were renamed and two criteria were combined. Specifically, 'Criteria to assess the short- and long-term behavioural safety effects' was shortened to 'Criteria to assess behavioural safety effects' and 'Other assessment criteria' was named as 'Criteria to assess user experience and social perception'. The criterion 'Integration with road and/or railway environment or other safety measures' was included as an additional aspect of 'Acceptance' to cover this topic also from the road and rail stakeholder's viewpoint taking into consideration the integration of the measure to the environment.

2.4. Key Safety Indicators identified in SAFER-LC Task 2.1

Deliverable D2.1 (SAFER-LC consortium, 2018b) generated a knowledge base drawing on existing data sources and analytical tools with a view to enhancing the safety performance of level crossing infrastructures from a human factor perspective. The main outcome of the task was the identification of key safety indicators concerning human errors and violations at level crossings based on a review of relevant human factors literature and a process of individual expert evaluation. The methodology applied in the development of this task comprised five key phases.

- Phase I: To ensure the identification of a comprehensive set of literature, a bibliographical database regarding human factors at level crossings and safety systems was jointly constructed by Task 2.1 partners (comprising 125 validated documents).
- Phase II: The publications were then reviewed by partners using a Review Form template. This form, created in Excel, sought to capture relevant and comparable information from the publications and comprised six key information fields, including human factors variables.
- Phase III: In a third phase, the results of the Review Forms underwent a descriptive univariate analysis in order to identify a set of user requirement indicators, based on the most frequently cited variables in the literature.
- Phase IV: The user requirement indicators were then validated by task partners through an individual indicator rating exercise. On a five-point Likert scale partners provided a subjective rating on the relevance of the indicators in terms of measuring safety at level crossings from a human factors perspective, together with observations about the ratings assigned.
- Phase V: In a final phase, the validated user requirement indicators were mapped against the GIDAS human error categorization framework (Graab et al., 2008). This allowed the

classification of the indicators in terms of underlying mechanisms of human error and violation applied to level crossing accidents which led to the renaming of the indicators more specifically as key safety indicators concerning human error and violations.

From this process of data collection and analysis, a total of thirty indicators were identified, grouped under seven categories (see Table 5).

Table 5. Key safety indicators concerning human errors and violations at LCs identified in D2.1 (SAFER-LC consortium, 2018b).

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.

These groups of indicators constitute a first step towards building the Human Factors methodological framework through helping to define a set of self-explaining and forgiving evaluation criteria. The validation exercise carried out with experts in Task 2.1 did however reveal some differences in opinion regarding these indicators and the need to further refine and clarify their definition. The next stage in this process therefore, has been to reach a consensus on how to transfer the key safety indicators identified in Task 2.1 to the Human Factors methodological framework in Task 2.2. To this end, a face to face group consultation with WP2 partners was conducted in Paris, in March 2018, where each indicator category was discussed in depth and its validity for inclusion in the framework. The debate centred around the main results of D2.1 and most notably the arguments regarding the indicator ratings that emerged from the expert evaluation. A brief summary of this discussion and the consensus reached is presented in the following.

The debate first sought to determine which of the indicators identified in Task 2.1 should be considered within the HF framework. In terms of informing the development of criteria in the HF methodological framework, one of the most important results of this group discussion was the agreement that the indicators identified in Task 2.1 should be considered more as factors that influence the safe interaction of users with level crossings (i.e. independent variables). Therefore, from the categories proposed in Task 2.1, the discussion centred around agreeing on a list of factors (independent variables) and how they can be measured as indicators (dependent variables). In this way, it was decided that the "Criteria" to be used in Task 2.2 should include a checklist of factors and indicators from the categories proposed in Task 2.1, identified as relevant and able to help answer the question: "Would this measure work if this specific condition occurs?".

It should be highlighted that there was difficulty in achieving consensus on the relevance of one factor over another, this being relative to the context of where a safety measure is implemented and analysed (e.g. type of level crossing). For this reason, it was agreed that focus should be placed on those factors that are feasible to measure, detect and control through external intervention (e.g. LC design). For example, whilst internal distraction was identified to affect safety (in Task 2.1) 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.

The discussion offered the opportunity to further explore, clarify and agree on definitions of the factors and associated indicators which in some cases led to their renaming. By way of example: the category of personal conditions (e.g. gender, age, disability etc.) was renamed as *socio-demographical factors* and classified as *background classification criteria*. On the other hand, the category of distraction and inattention was changed to *cognitive factors* and classified as *criteria to assess the behavioural safety effects*. The listing of the factors has also been further refined or nuanced taking into account the relationship between some of the original indicator categories. For example, strong links were identified between knowledge and attention (dealt with originally in Task 2.1 as two separate categories) whereby knowledge of the traffic rules and the correct action can lead to better attention at level crossings which is now reflected within the 'Rule Knowledge' factor.

Implications for the HF framework:

Based on the results of D2.1 and the further group work between the SAFER-LC T2.2 partners, several relevant indicator categories were identified (as shown in the table above). Some of these were regrouped and renamed (e.g. 'Rule Knowledge') and were 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 results of the discussion regarding the transfer of key safety indicators identified in Task 2.1 to criteria for the HF methodological framework in Task 2.2 has been incorporated within the table of proposed criteria presented in Chapter 3. More generally and in addition to developing criteria in the form of factors and indicators, the discussion around the results of Task 2.1 has helped to focus ideas around how to structure the human factors framework criteria as presented in the aforementioned table.

3. CRITERIA SELECTED FOR THE HUMAN FACTORS ASSESSMENT TOOL

This chapter describes the sets of criteria (and the indicators behind them) selected for the SAFER-LC HF methodological framework. They represent the backbone of the HF assessment tool which is presented in the Chapter 4.

3.1. Selected sets of criteria and the indicators behind them

The overall framework for the HF Assessment Tool developed in SAFER-LC is presented in Figure 5. This offers an overview of the sets of criteria and indicators selected to estimate the safety measure from a human factors perspective. This framework was built based on a combined methodology covering the literature review conducted in D2.1, criteria-oriented review performed in Task 2.2 (based on relevant theories and models and summarized in Tables 2 and 3), and panel discussion between WP2 partners.

The framework consists of three sets of criteria summarized in Figure 5 and 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 the user experience and social perception' – blue). Each of these criteria categories are based on the sets of factors and indicators which represent the backbone of the HF Assessment Tool. As presented in detail in the following tables and explained in greater depth within this section, each criterion can be further broken down into a set of more specific and measurable indicators.

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 LC 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 the road and railway safety (Table 6). These criteria are more 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 installed only 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 European wide LC safety situation can be expected, even though the effectiveness of that specific measure can be estimated as high.

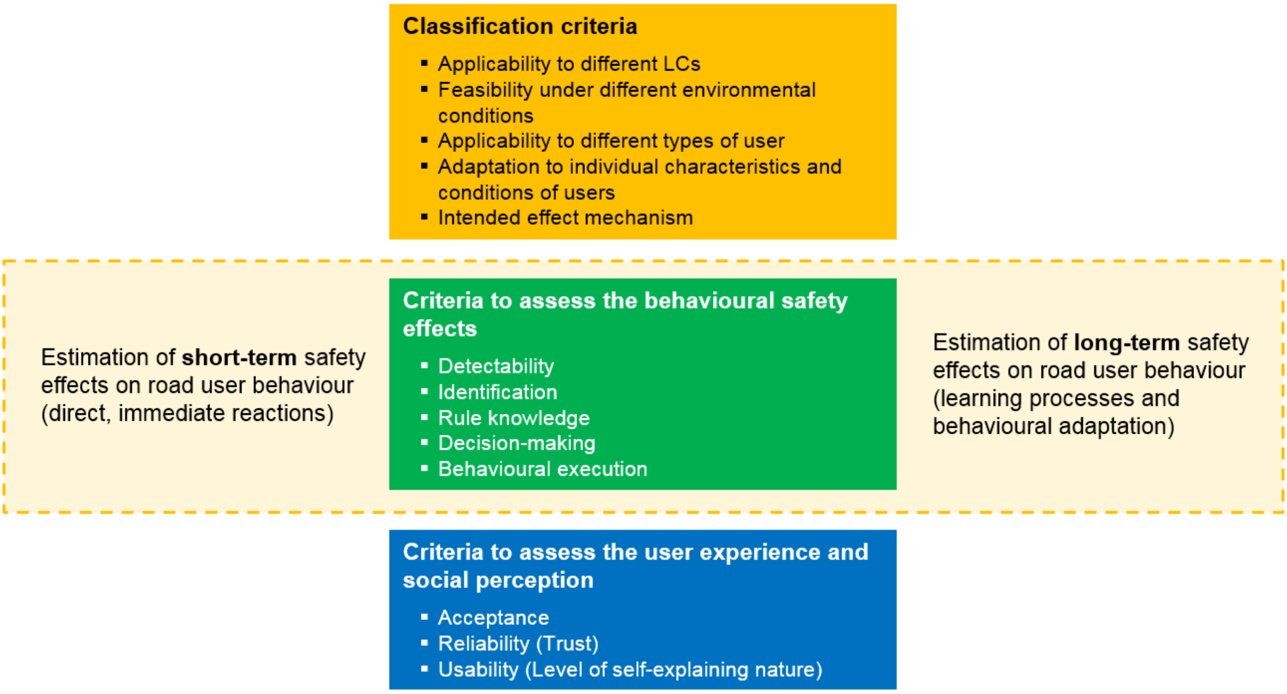

Figure 5. The SAFER-LC HF methodological framework: overview of the sets of classification and assessment criteria selected for the HF assessment tool.

Table 6. 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 ▪ LCs with paved road ▪ LCs with gravel road ▪ LCs with availability of electricity ▪ LCs with 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 can be implemented 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 7 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 (cf. Grippenkoven, 2017; Wickens et al., 2012). Once the estimated changes in road user behaviour have been identified (both short and long-term), the quantification of safety effects will be calculated, for example, based on KPIs collected in WP4, literature, expert assessment, LC statistics etc.

Table 7. 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
	  ○ Crossing angle
	▪ Personal characteristics
	  ○ Individual visual/auditory capabilities
		▪ Detectability of approaching 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 (e.g. multitasking)
	  ○ 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
	  o Perception of probability
	  o Perception of dangerousness
	  o Perception of legal consequences
	  o Perception of cost-benefits
	▪ Individual motivations
	  o Time pressure
	  o Suicide or vandalism intentions
	▪ Personal characteristics
	  o Personality of the road user (e.g. risk seeker)
	  o 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) ▪ Speed choice / Approach speed (at passive LC) (+/- km/h) ▪ Trajectories ▪ Verification behaviours for frequent users ▪ Time to collision (TTC) when a train is coming ▪ Interaction with other road users

Table 8 presents the three criteria to assess the user experience and social perception of the safety measure. The indicators refer to the subjective opinions of the 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 behalf of the end user and wider community is important as it may affect their interaction and correct usage of the measure which can impact safety. Information related to these indicators are proposed to be collected via 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 8. Criteria to assess the 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 how 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

3.2. Quantification of the sets of criteria

The HF assessment tool will be used in the SAFER-LC project to evaluate the effectiveness of safety measures from a HF point of view. More specifically, it provides a quantitative estimation of how a safety measure implemented in a given setting is likely to perform according to relevant HF criteria such as the road users’ cognitive processes, needs, behaviour and experience. Therefore, the specific objective of Task 2.2 is to identify, formulate and classify measurable sets of criteria that support road and rail decision-makers to take into account human factors when addressing safety at LCs.

The assessment tool which is included in the HF methodological framework consists of a set of evaluation criteria for self-explaining and forgiving design, against which the safety measures will be objectively evaluated with the assignation of a score rating. This means that the selected criteria should focus on *quantitative assessment* of a list of measures, based on very specific measurable subsets of criteria (e.g. absolute numbers or Likert-type evaluation scales with a preferred range from 0–5). However, in addition, some *qualitative criteria* are proposed to be used so that LC safety measures can be evaluated on an ordinal scale (e.g. level of self-explanation of LC signage: Good /

Average / Bad). It is therefore challenging to find a common scoring method for different sets of criteria which have a different quantitative-qualitative nature.

Various methods were reviewed to assign score ratings to the different groups of criteria. The chosen solution was to use a maturity scale which is able to transform and integrate quantitative, semi-quantitative, and/or qualitative data by creating a unique and comparable aggregated score. Finally, it was decided to use an adaptation of the Fleishman (1975) ability scales combined with the Capability Maturity Model (CMM; Humphrey, 1988) used in the Control Objectives for Information and related Technology (COBIT) framework as shown in Table 9. The term 'maturity' originally referred to the software development process and related to the degree of formality and optimization of a process, from ad-hoc practices to more and more elaborated steps. In more general terms, a maturity model can be viewed as a set of structured levels that describe how well the behaviours, practices and processes of an organization or system can reliably and sustainably produce required outcomes. It includes a 5-level maturity continuum where the uppermost (5th) level is a theoretically ideal state. The model provides a theoretical continuum along which a specific characteristic or function can be developed incrementally from one level to the next, therefore skipping levels is not possible or feasible.

Table 9. Example of a maturity scale ranging from 0 to 5 adapted from the CMM levels (Humphrey, 1988).

Quantitative Score	**Maturity level in the CMM**	**Level description in the COBIT framework**	**Possible semi-quantitative tag (for general purpose)**
0	Non-existent	Processes are not applied at all	Unacceptable
1	Initial	Processes are ad-hoc and disorganised	Very poor
2	Repeatable (Intuitive)	Processes follow a regular pattern	Poor
3	Defined	Processes are documented and communicated	Average
4	Managed (Capable)	Processes are monitored and measured	Good
5	Optimised (Efficient)	Good practices are followed and automated	Excellent

An adaptation of such a scale for other evaluation purposes might require a development of alternative descriptions for the 'levels' which are adapted to a given context. For example, Fleishman (1975) used a similar scale to illustrate the extent to which a human ability is developed, by defining the extremes of the scale and giving some examples for intermediate points of the scales. Essentially, the lowest end of a Fleishman scale illustrates the minimum required level of a human ability and the highest end of the scale the ideal level of the needed ability in a given context.

A maturity scale defined based on the CMM has a number of advantages. Firstly, it allows an aggregated score for each criterion (overall score between 0–5), which brings several types of criteria into the same reference system. Secondly, it gives the possibility to visualize the evaluation of the results in a type of chart and to conduct comparative assessment of different measures based on a common set of criteria.

On the other hand, the challenges of such a scale consist in defining the degrees to which a specific feature is completely present or absent and to formulate them in clear and illustrative way. Once the scale is defined, it can be used during an expert assessment but the score assignment remains a subjective process, a limit which is inherent to all self-report scales.

At this stage of Task 2.2 it was decided to develop the scales at the level of the general indicators, with definitions limited to the lower and upper end of the scale (0 and 5). The next chapter includes the scales for each main factor.

4. THE HUMAN FACTORS ASSESSMENT TOOL FOR TESTING

This chapter includes the list of HF items and assessment forms ready to be used in the field work (pilot sites). The sets of criteria described in the previous chapter are now operationalised in several forms which can be filled in for a given safety measure under evaluation. The colour of the forms (orange, green and blue) reflect the colour code represented earlier in Figure 5. The 'Classification criteria' are included in a classification checklist (Table 10; orange). The 'Criteria to assess the behavioural safety effects' are included in five separate assessment sheets, one for each criterion (Tables 11–15; green). The 'Criteria to assess the user experience and social perception' are included in one assessment sheet (Table 16; blue).

Table 10. Human Factors Assessment Tool: Classification Criteria Checklist

Name of the measure being assessed	**Brief description**
……………………………..	…………………………………………………………………………………

CLASSIFICATION CRITERIA		
Factor	**Brief description**	**Indicator** *(Tick all the cases that the measure applies to)*
Applicability to different LCs	*Specify the types and characteristics of LCs where the measure can be implemented*	Type of LCs ☐ Passive LCs without any warning devices ☐ 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 devices ☐ Active LCs with traffic lights Characteristics of LCs ☐ 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 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	*Specify the categories of LC users who are targeted by the measure*	☐ All road users MRU ☐ cars ☐ motorbikes / mopeds ☐ trucks / heavy vehicles ☐ buses / coaches ☐ farm / agricultural vehicles ☐ other (specify)………………………… VRU ☐ pedestrians ☐ cyclists ☐ other (specify)…………………………
Adaptation to individual characteristics and conditions of users	*Specify if the measure is specifically targeted at people with the following characteristics or conditions*	Gender ☐ Male ☐ Female Age ☐ children ☐ elderly ☐ all ages Disability ☐ vision loss and blindness ☐ hearing loss and deafness ☐ intellectual disability ☐ reduced mobility ☐ other (specify)………………………… Under influence of ☐ alcohol ☐ drugs ☐ medication Under skill impairing states ☐ fatigue ☐ stress ☐ Risk-seeking personality
Intended effect mechanism	*Specify 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 access to the LC ☐ Reduces the approach speeds of vehicles ☐ Increases the user's 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 ☐ Other (specify) ……………………………………

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

Table 11. Human Factors Assessment Tool: Assessment sheet for 'Detectability'

Criterion	Brief description
Detectability	The measure can help the LC user detect relevant visual and auditory stimuli, therefore increasing the detectability of the LC or the approaching train

Write down brief descriptions of the expected and/or observed changes in road user's detection of the LC or train as a result of 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	• … • …	• … • …	• … • …	• … • …

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 of the LC /or train 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* ………………………………………………………………………………………………

Table 12. Human Factors Assessment Tool: Assessment sheet for 'Identification'

Criterion	Brief description
Identification	The measure can increase safety by helping the LC user identify relevant information in the environment and not be distracted by irrelevant information

Write down brief descriptions of the expected and/or observed changes in road user's identification of relevant information as a result of 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	• … • …	• … • …	• … • …	• … • …

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 identification of a possible danger as the user is approaching the LC?**		
Answer modalities	N	The LC user is somewhat confused or distracted by this measure and therefore unable to identify potential dangers related to the crossing of the LC
	0	This measure has no intended influence on the attention or workload of the LC user
	1	
	2	
	3	
	4	
	5	LC users' attention is naturally directed towards identifying a potential danger despite being fatigued, distracted, or under high workload (also in the long term)
Score	**…**	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* ……………………………………………………………………………………………

Table 13. Human Factors Assessment Tool: Assessment sheet for 'Rule knowledge'

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 expected and/or observed changes in road user's ability to elicit and retrieve relevant safety information as a result of 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	• … • …	• … • …	• … • …	• … • …

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* ………………………………………………………………………………………………

Table 14. *Human Factors Assessment Tool: Assessment sheet for 'Decision-making'*

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 expected and/or observed changes in road user's decisions as a result of 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	• … • …	• … • …	• … • …	• … • …

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* ……………………………………………………………………………………………

Table 15. Human Factors Assessment Tool: Assessment sheet for 'Behavioural execution'

Criterion	Brief description
Behavioural execution	The measure can help 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 expected and/or observed changes in road user's behavioural execution as a result of 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	• … • …	• … • …	• … • …	• … • …

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 influence the safe execution of the approach and crossing behaviour?**		
Answer modalities	N	The LC user's crossing action is disturbed and becomes more difficult when this measure is in place
	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

Table 16. Human Factors Assessment Tool: Assessment sheet for '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):					

5. APPLICATION GUIDE FOR TESTING

This chapter provides guidance on how to use the HF assessment tool in practice. By way of example, the forms presented in the previous chapter have been filled out for a specific measure (i.e. speed bumps). The example aims only to be illustrative and uses imaginary data which are not based on any real measurement. Clarifications and instructions on how to fill out the assessment scales during the field work are provided in the introduction to each form.

From a practical application viewpoint, the HF assessment tool is conceived to evaluate specific measures of interest already selected for implementation. This assumes that before the application of the HF assessment tool the decision-maker has identified suitable measures in response to the local context and safety problematic to be addressed. Additionally, the decision-maker should be familiar with the name and the definition of the measure selected for assessment from a wider pool of possible measures.

Therefore, the first piece of information that needs to be filled out is the name and a brief description of the measure being assessed:

Name of the measure being assessed	Brief description
Speed bumps	A small raised area built across a road within the approach zone of the LC to force road users to drive more slowly

Then, the list of forms needs to be filled out one by one in the given order.

5.1. Measurement scales and scoring system

There are three sets of criteria:

1. Classification criteria on a checklist (orange table) where no quantitative score is assigned.
2. Criteria to assess behavioural safety effects (green table): evaluation on maturity scales with descriptions and examples (scoring 0–5).
3. Criteria to assess the user experience and social perception (blue table): evaluation on Likert-type scale (scoring 0–5).

By using this scoring system, all the assessment criteria (5 green and 4 blue) are measured in the same system of reference which allows the calculation of an aggregated score (total or average). The minimum theoretical score is 0 (supposing that the assessed measure scores 0 for all criteria) and the maximum theoretical score is 45 (if the measure scores 5 for all criteria). Further, an average score among all nine assessed criteria can be computed to give an overall quantitative estimation of how well the measure is likely to perform from a HF viewpoint.

5.2. Classification Criteria Checklist: Example applied to speed bumps

The classification criteria are presented in the orange checklist and need to be filled out by ticking all the relevant cases.

CLASSIFICATION CRITERIA		
Factor	**Brief description**	**Indicator** ***(Tick all the cases that the measure applies to)***
Applicability to different LCs	*Specify the types of LCs where the measure can be implemented*	Type of LCs ☒ Passive LCs without any warning devices ☐ 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 devices ☐ Active LCs with traffic lights Characteristics of LCs ☒ 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 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	*Specify the categories of LC users who are targeted by the measure*	☐ All road users MRU ☒ cars ☒ motorbikes / mopeds ☒ trucks / heavy vehicles ☒ buses / coaches ☒ farm / agricultural vehicles ☐ other (specify)………………………… VRU ☐ pedestrians ☐ cyclists

		☒ other (specify) ……*cyclists could be targeted depending on the width of the road and the width of the speed bump*……………………………….
Adaptation to individual characteristics and conditions of users	*Specify if the measure is specifically targeted at people with the following characteristics or conditions*	Gender ☒ Male ☒ Female Age ☐ children ☒ elderly ☐ all ages Disability ☐ vision loss and blindness ☒ hearing loss and deafness ☐ intellectual disability ☒ reduced mobility ☐ other (specify)…………………………. Under influence of ☒ alcohol ☒ drugs ☒ medication Under skill impairing states ☒ fatigue ☒ stress ☒ Risk-seeking personality
Intended effect mechanism	*Specify 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 access to the LC ☒ Reduces the approach speeds of vehicles ☐ Increases the user's 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 ☐ Other (specify) ……………………………………

5.3. Assessment Sheets for behavioural safety effects: Example applied to speed bumps

The following five green forms include assessment criteria for the estimation of behavioural safety effects. There is one form per criterion, and each criterion corresponds to a different area of psychological function (detectability, identification, rule knowledge, decision-making and behavioural execution). The purpose is to evaluate how a given implemented measure is able to influence these five elements.

This is a two-step process:

1. The respondent first needs to write a brief description of the expected and/or observed changes in road user behaviour resulting from the measure. This should be done by filling

out a summary of collected evaluation data taken from the existing literature and/or the pilot tests. The evaluation form makes a clear separation between evidence collected from previous studies (e.g. evaluation of same or similar measures in other settings), and findings from a dedicated SAFER-LC pilot study (i.e. behavioural effects of the measure).

If possible, the information given here should make reference to the indicators listed in Table 7 of Chapter 3 (e.g. type and number of errors, type and number of violations, etc.) and to the Human Behaviour KPIs listed in D4.2 (e.g. perception, queuing behaviour, traffic violations; check D4.2 pages 14–16 for more examples). In general, only the relevant cells for a specific case need to be filled in. If no information is available about a specific criterion, indicate by stating '*No information available*' or '*Not applicable – N/A*'.

The 'Before' and 'After' conditions are only relevant for studies with repeated measurements. 'Before' refers to the baseline condition (if available) or to a measurement conducted before the implementation of the measure or in absence of the measure.

Note that long-term effects are probably only valid for 'After' -measurements. The meaning of 'long-term' may vary depending on the study setting, research methodology or the type of measure under investigation. For example, in naturalistic study settings (e.g. observations conducted at a real-world LC) the recorded data could reflect habitual behaviours and may count as long-term behaviour. On the other hand, in a simulator study a long-term effect can refer to an additional measurement conducted a few weeks after the immediate measurement of the first effects. Therefore, the respondent filling in the short and long-term data should specify in their description what 'long-term' refers to (e.g. repeated measurement conducted after two months, observations conducted after 12 months, etc.).

2. Based on the pieces of evidence collected, the respondent can make an informed assumption about the short and long-term changes in road users' behaviour. Based on this assumption, the respondent will choose the most appropriate score between 0 and 5 and to write it down in the dedicated cell. For each question, only one answer must be selected. The reasoning behind the score must be written in the brief space provided next to the score.

   Only the upper and lower ends of the scale are defined. The definition for the minimum score (0) assumes that the measure does not tackle the respective aspect of behaviour and has no influence on it. On the other hand, the definition for the maximum score (5) refers to 'LC users' as a generic group. This assumes that the measure has a positive influence on almost all LC users in relation to the behavioural aspect under scrutiny, including MRUs and VRUs and accounting for various individual characteristics and personal conditions: age, disability, skill impairing states etc.

   In addition, there is an answer modality coded 'N' which stands for 'negative or adverse effect'. This is not part of the actual scoring scale, yet this option can be selected if the collected evidence indicates that the measure leads to an opposite effect than the one intended.

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

Criterion	Brief description
Detectability	The measure can help the LC user detect relevant visual and auditory stimuli, therefore increasing the detectability of the LC or the approaching train

Write down brief descriptions of the expected and/or observed changes in road user's detection of the LC or train as a result of 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	N/A	N/A	Some drivers did not direct gaze towards LC warning signs	N/A
After / With the measure	N/A	N/A	Most drivers directed gaze towards LC warning signs	N/A

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 of the LC /or train 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	**2**	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* Slowing down MRUs and cyclists will facilitate the detection of relevant visual and auditory stimuli such as LC signage and warnings (i.e. signs that might have been missed if travelling at speed) which alert the user to the LC and approaching train …………………………………………………………………………………………

Criterion	Brief description
Identification	The measure can increase safety by helping the LC user identify relevant information in the environment and not be distracted by irrelevant information

Write down brief descriptions of the expected and/or observed changes in road user's identification of relevant information as a result of 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	N/A	N/A	Drivers looked left and right only once	N/A
After / With the measure	N/A	N/A	Drivers looked left and right several times	N/A

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 identification of a possible danger as the user is approaching the LC?**		
Answer modalities	N	The LC user is somewhat confused or distracted by this measure and therefore unable to identify potential dangers related to the crossing of the LC
	0	This measure has no intended influence on the attention or workload of the LC user
	1	
	2	
	3	
	4	
	5	LC users' attention is naturally directed towards identifying a potential danger despite being fatigued, distracted, or under high workload (also in the long term)
Score	**2**	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* A lower speed gives MRUs more time to look left and right as they approach the LC …………………………………………………………………………………………

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 expected and/or observed changes in road user's ability to elicit and retrieve relevant safety information as a result of 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	N/A	N/A	20% of the drivers respected the speed limit while approaching the LC	N/A
After / With the measure	One study showed that speed bumps before passive LCs reduced the approaching speed by 20%	N/A	60% of the drivers respected the speed limit while approaching the LC	After 4 months, only 40% of the divers respected the speed limit while approaching the LC

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	**4**	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* Given that it is a traffic calming measure, most road users implicitly understand that they have to reduce the speed. However, the measure does not apply to all road users, for example it does not address pedestrians. ………………………………………………………………………………………………

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 expected and/or observed changes in road user's decisions as a result of 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	N/A	N/A	N/A	N/A
After / With the measure	When under time pressure the drivers declared that they are not willing to reduce the approaching speed	N/A	N/A	N/A

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	**3**	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* When they see a speed bump, most if not all MRUs understand that speed must be reduced, and they decide to slow down. In particular situations (e.g. time pressure) some drivers decide to maintain their speed. ………………………………………………………………………………………

Criterion	Brief description
Behavioural execution	The measure can help 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 expected and/or observed changes in road user's behavioural execution as a result of 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	N/A	N/A	The average LC approach speed was 47 km/h	N/A
After / With the measure	N/A	Reduction in average LC approach speed in the long term	The average LC approach speed was 35 km/h	Observed a zig-zagging behaviour, with drivers trying to avoid the bumps; No zig-zagging behaviour observed at control sites

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 influence the safe execution of the approach and crossing behaviour?**		
Answer modalities	N	The LC user's crossing action is disturbed and becomes more difficult when this measure is in place
	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	**4**	*Reasoning behind the score / Assumption on the short and long-term change in road user behaviour* Higher speed bumps can actually 'force' drivers to reduce speed while driving over the speed bump. If not for safety reasons, some drivers reduce speed not to damage their vehicles. …………………………………………………………………………………………………

5.1. Assessment Sheet for the user experience and social perception: Example applied to speed bumps

The last blue form includes the criteria to assess the user experience and social perception. The respondent should tick the most suitable answer for each affirmation (one answer between 0–5 where 0 stands for 'Unacceptable' and 5 for 'Excellent'). For each item, only one answer must be selected. The reasoning behind the score must be written in the brief space provided.

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): Frequent complaints from the drivers living nearby who claim that the speed bumps force them to slow down every time, even when not necessary and that they reduce their driving comfort					
	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): Endorsed by the railway stakeholders; strongly supported by the municipality which has an active local policy of traffic calming measures					
	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): In relation to the integration with other measures and the road/rail environment, the self-explaining nature of a LC may be compromised if too many measures (e.g. signs) are used in conjunction					
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): The measure aims to reduce speed and most road users are persuaded that lower speed correlates to safety. However, 10% of the observed drivers did not reduce their speed while driving on the speed bumps.					

<table>
<tr><td rowspan="2">Usability</td><td rowspan="2">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</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):

<u>Instinctively leads to speed reduction. In specific contexts it may also apply to cyclists. Usable for persons in wheelchair.</u></td></tr>
</table>

6. DISCUSSION AND CONCLUSIONS

This deliverable presented the first phase of the ongoing work to develop a Human Factors methodological framework conducted in the SAFER-LC project. This framework proposes several sets of criteria and associated indicators to classify and assess LC safety solutions, taking into account the road and rail users' perspectives and requirements.

The research work reported in this deliverable includes two parts. The first part of the report sets out the theoretical foundations of the HF methodological framework based on a review of relevant theories and models from the Human Factors and Traffic Psychology literature and identification of relevant sets of evaluation criteria. The second part of the report presents the HF assessment tool, the end result of this task which includes selected indicators and evaluation scales which are explained in the application guide for testing.

The HF assessment tool will be used to assess and categorise the safety measures piloted at the later stages of the SAFER-LC project. For example, the criteria to assess the behavioural safety effects of measures on road users are currently being used in Task 2.3 as part of the selection process of human centred low-cost measures. Further, the whole HF assessment tool will be verified and improved in the development and implementation phases of WP4. The assessment process planned in WP4 covers the user experience focusing among others on the acceptance, reliability and usability of the safety measures. Therefore, the pilots shall include collection of opinions of different stakeholders (e.g. road users, train drivers, road and rail transport authorities) taking into account age, gender and cultural background of the persons.

Through its testing in WP4 pilots, the application of the HF assessment tool in SAFER-LC will provide:

- Support to road and railway stakeholders to implement LC safety measures aiming to reduce human errors and violations related to infrastructure design;
- New approaches to raise awareness of HF related issues in collision prevention at LCs;
- Particular attention to vulnerable road users (pedestrians, cyclists, etc.);
- Better knowledge of the human requirements of LC users by unifying the existing research and analytical tools available; and
- An easy-to-use tool to evaluate the efficiency of different LC designs and safety measures with special focus on HF issues.

Therefore, the use of HF assessment tool will help both road and rail stakeholders involved in LC safety work to better understand the road users' needs and related requirements. This way, the road users' needs and requirements can be taken into account in the implementation of future designs for LCs. This will enable the optimization of the design of LCs and the associated safety measures in Europe and beyond by: (a) boosting the innovation potential for the industry in this area, (b) enhancing the safety levels of LC users, and (c) contributing to the development of more self-explaining LC infrastructure.

The HF assessment tool will be applied during the project lifetime to evaluate the efficiency of safety measures with respect to road users' needs, cognitive processes, and behaviour. As it is being designed not only as a theoretical tool but also as a practical one, it will allow the LC stakeholders to tailor unique solutions for different LC environments after the end of the project. Based on the evaluation carried out by applying the HF assessment tool, one will be able to make informed suggestions on the design of the layout of LCs to make them "user friendly". For example, if LCs are located in areas of high workload and visual clutter for road users, alternative signage directed to other areas of their cognitive performance can be suggested in order to enhance the self-explaining and forgiving nature of the LC infrastructure for the users.

6.1. Steps towards socio-economic assessment of safety measures

As indicated earlier in this deliverable, the proposed HF assessment tool will be used in the SAFER-LC project to evaluate the effectiveness of safety measures from a HF point of view. Hence, the aim is to provide a quantitative estimation of how a safety measure implemented in a given setting is likely to perform according to relevant HF criteria. To obtain this, the proposed sets of assessment criteria exploit maturity-type evaluation scales and Likert-type scales, which are used by the evaluators to estimate the effectiveness of the safety measure from a HF perspective. Specifically, this means that the estimated extent and permanence of behavioural safety effects are defined according to the maturity scale from 0 to 5, and the level of agreement or disagreement regarding the user experience and social perception criterion is defined on a symmetric inadequate-excellent scale for a series of questions which also score from 0 to 5.

However, for the purposes of socio-economic assessment of safety measures more quantitative assessment results are needed. Particularly, an estimate on the number of prevented LC accidents and related fatalities and injuries are typically requested as an output of safety impact assessment. In order to proceed a step further towards this objective Figure 6 proposes a method for more detailed quantification of safety effects based on the work done as part of the HF assessment tool. The more detailed quantification utilises the information collected via 'Classification criteria' and 'Criteria to assess the behavioural safety effects'.

The more detailed quantification is proposed to be done in two steps:

1. 'Classification criteria' are used to define the accidents which are targeted to be prevented with the implementation of the safety measure. This also includes the identification of locations and circumstances where the measure can be implemented and is estimated to be effective.

   By combining this information with available statistics on the share of different types of LCs in Europe and by the share of relevant accidents we can define the amount of LC accidents which can potentially be prevented by the implementation of the safety measures. Especially for the identification of the share of relevant LC accidents, the use of expert assessment might be needed in addition to information obtained from statistics. Information on the share of some specific type of LC accidents might be available only from one or few countries and thus there is a need to make a European-wide estimate on that share based on the available numbers. For example, we could have a safety measure which is designed for passive level crossings with gravel roads and is estimated to work only during darkness. In this case, the assessment should proceed in three steps: 1) estimation of the share of passive level crossings with gravel roads of all European LCs (to be used as background information), 2) estimation of the share of LC accidents which occur at passive LCs with gravel roads, and finally 3) estimation of the share of LC accidents which occur during darkness at passive LCs with gravel roads.

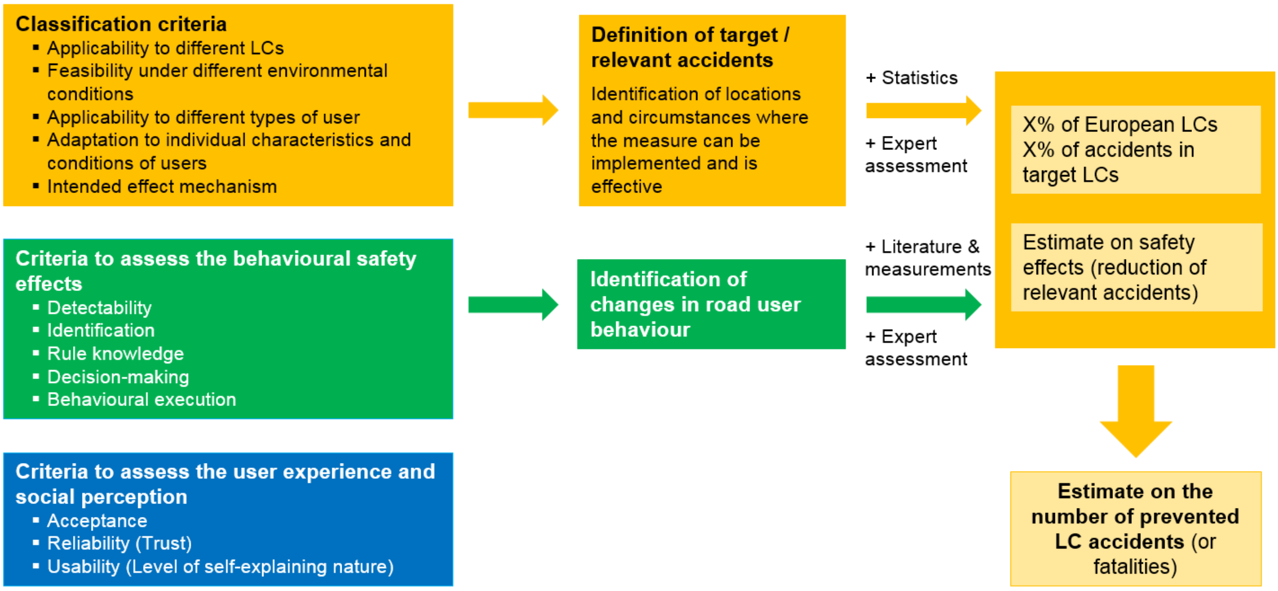

Figure 6. Method for more detailed quantification of safety effects.

2. ‘Criteria to assess the behavioural safety effects’ are used to identify the short- and long-term changes in road user behaviour which are estimated to occur due to the implementation of the safety measure.

   The quantitative estimate on safety effects in terms of reduction of relevant accidents will be determined based on existing literature, results from the pilot tests and expert assessment. The relevant literature covers findings on safety effects of LC safety measures or safety measures implemented e.g. in the road environment. At this step it is important to document all assumptions used to define the estimate so that these assumptions can be updated and modified as new information on road user behaviour is available from future field pilots or other studies. In case no information on long-term effects of the safety measure in terms of prevented accidents is available it is recommended to provide the estimate of the safety measure by using a scale (e.g. 5–10 % reduction of relevant accidents) instead of an exact value.

The estimate of the number of prevented LC accidents with a specific safety measure can be then calculated by combining these two above estimates: the share of relevant LC accidents and the estimated safety effect of the measure.

The focus of the method presented in this chapter is not anymore on HF perspective and therefore its application will be part of WP4 which focusses on quantifying the safety effects of piloted safety measures and on providing inputs for the socio-economic assessment which will be conducted as part of WP5 of the SAFER-LC project.

6.2. Future validation of the Human Factors framework

The first version of the Human Factors assessment tool and its application guide will be used during the project trials to evaluate innovative measures aiming to improve the safety of LCs from safety and human factors point of view. Most of the measures selected or developed within the SAFER-LC project will be tested and further developed under different environments in several test-sites (e.g. laboratory, driving simulator, living lab). The various test-sites available in the SAFER-LC project are a perfect fit for measures at different stages of maturity. Early stage developments can be tested in simulation environments or on controlled test tracks, while more readily developed measures will be evaluated in field pilots.

The information collected in the demonstration phase through the HF assessment tool will allow the evaluation of the developed measures and the drawing of recommendations on human-centred improvements and organizational processes related to the evaluated measures. At the same time, based on the experiences gathered at the test-sites, the proposed HF assessment tool will be validated and improved at the later stages of the SAFER-LC project. For example, the evaluation scales proposed in the HF assessment tool will be adjusted according to the feedback collected from the pilot sites and some criteria may be further refined or excluded.

The main added value of the SAFER-LC project in the context of LC safety is related to the integration of various aspects of LC systems (human, infrastructure, technologies, and management), and the HF methodological framework along with the HF assessment tool presented in this deliverable contribute to the analysis of the human component within this system.
Eventually, the SAFER-LC project will propose a combination of recommended and innovative technical specifications, human-centred low-cost measures and organizational and legal frameworks for implementation. These will be delivered through a toolbox, accessible through a user-friendly interface which will integrate all the project results and solutions to help both rail and road managers to improve safety at level crossings. The HF methodological framework and will be an integrative part of this toolbox. For example, the classification criteria proposed in the HF assessment tool could be used to organise and cross-classify the measures in the SAFER-LC toolbox.

7. REFERENCES

- Bahloul, K., Defossez, F., Ghazel, M. & Collart-Dutilleul, S. (2017). Adding Technological Solutions for Safety Improvement at Level Crossings: A Functional Specification. Procedia-Social and Behavioural Sciences. 48. pp. 1357-1384.
- 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.
- Borsos, A., Gabor, M., Csaba Koren, C. (2016). Safety ranking of railway crossings in Hungary. 6th Transport Research Arena April 18-21, 2016. Elsevier
- Caird, J. K., Creaser, J. I., Edwards, C. J. & Dewar, R. E. (2002). A human factors analysis of highway-railway grade crossing accidents in Canada. Montreal, Canada: Transport Canada.
- Dian, L., Cassidy, T., Bromilow, D. (2011). Product Instructions in the Digital Age.
- Draskóczy, M., Carsten, OMJ. & Kulmala, R. (1998). Road Safety Guidelines. CODE Project, Telematics Application Programme, Deliverable B5.2.
- Edquist, J., Hughes, B., & Rudin-Brown, C. M. (2011). Pedestrian non-compliance at level crossing gates. C-MARC. Perth.
- ERTRAC (2011). ERTRAC Research and Innovation Roadmaps. European Road Transport Research Advisory Council. https://www.ertrac.org/uploads/documentsearch/id7/ertrac-researchinnovation-roadmaps_60.pdf
- European Commission (2011). D4 – Final Report. Action 3.4 – Safety and comfort of the Vulnerable Road User. Brussels. European Commission.
- European Union (2010). Directive 2010/40/EU of the European Parliament and of the Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport.
- Fleishman, E. A. (1975). Development of ability requirement scales for the analysis of Bell Systems jobs. Bethesda, Maryland: Management Research Institute
- Freeman, J., & Rakotonirainy, A. (2015). Mistakes or deliberate violations? A study into the origins of rule breaking at pedestrian train crossings. Accident Analysis & Prevention, 77, 45-50.
- Freeman, J., Rakotonirainy, A., Stefanova, T., McMaster, M., (2013). Understanding pedestrian behaviour at railway level crossings: is there a need for more research? Road and Transport Research Journal, 22 (3), 29–39.
- Fuerstenberg, K. & Boehning, M. (2012). Final Report. D1.2, MINIFAROS project.

- Graab, B., Donner, E., Chiellino, U., Hoppe, M. (2008). Analyse von Verkehrsunfällen hinsichtlich unterschiedlicher Fahrerpopulationen und daraus ableitbarer Ergebnisse für die Entwicklung adaptiver Fahrerassistenzsysteme. In 3.Tagung Aktive Sicherheit durch Fahrerassistenz; 2008; TU München. (Engl.: “Analysis of traffic accidents with regard to different driver populations and results that can be derived for the development of driver assistance systems. Published in the conference proceedings of the 3. Symposium on active safety by driver assistance; 2008;TU München.”)
- Grippenkoven, J. (2017): Wahrnehmung und Verhalten am Bahnübergang. In: Deine Bahn (2), 10–15.
- Grippenkoven, J., & Dietsch, S. (2015). Gaze direction and driving behavior of drivers at level crossings. Journal of Transportation Safety & Security, 4–18.
- Grippenkoven, J., Giesemann, S. & Dietsch, S. (2012). The Role of Human Error in Accidents on German Half-Barrier Level Crossings.
- Hollnagel, E. (1993a). Human reliability analysis: Context and Control, Academic London.
- Hollnagel, E. (1993b). The phenotype of erroneous actions. International Journal of Man-Machine Studies, 39, 1–32.
- Humphrey, W. S. (1988). Characterizing the software process: A maturity framework. IEEE Software 5(2), 73–79.
- Innamaa, S., Rämä, P., Hogema, J., Kroon, L., Heinig, I., Schindhelm, R., Ojeda, L., Servel, A., Visin-tainer, F. & Giosan, I. (2014). Safety impacts of cooperative systems. Deliverable IR4.y of DriveC2X project.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica 47(2), 263–292.
- 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.
- Laapotti, S. (2016) Comparison of fatal motor vehicle accidents at passive and active railway level crossings in Finland. *IATSS Research 40*, 1–6.
- Lerner, N., Ratte, D., and Walker, J. (1990). Driver Behavior at Rail-Highway Crossings (FHWA-SA-90-008). Washington, D.C.: Federal Highway Administration, Office of Highway Safety.
- Lobb, B. (2006). Trespassing on the tracks: A review of railway pedestrian safety research. Journal of Safety Research, 37(4), 359–365.
- Lobb, B., Harré, N., & Suddendorf, T. (2001). An evaluation of a suburban railway pedestrian crossing safety programme. Accident Analysis and Prevention, 33, 157–165.

- 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.
- Michon, J. A. (1985). A critical view of driver behaviour models: what do we know, what should we do? Evans, L. & Schwing, R. (Eds.). Human behavior and traffic safety (pp. 485–520). New York: Plenum Press.
- Naumann, A.; Hurtienne, J.; Israel, J.H.; Mohs, C.; Kindsmüller, M.C.; Meyer, H.A. et al. (2007): Intuitive Use of User Interfaces - Defining a Vague Concept. In: D. Harris (Hg.): Engineering Psychology and Cognitive Ergonomics. Proceedings of the HCI International 2007. Berlin: Springer, 128–136.
- Neth, H.; Sims, C. R.; Gray, W. D. (2006): Melioration dominates maximization: Stable suboptimal performance despite global feedback. In: R. Sun und Miyake N. (Hg.): Proceedings of the 28th annual meeting of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum, 627–632.
- Nilsson, G. (2004). Traffic Safety Dimensions and the Power Model to describe the effect of speed and safety. Bulletin 221. Department of Technology and Society. Lund University. Sweden.
- Näätänen, R. & Summala, H. (1976). Road-user behaviour and traffic accidents. North-Holland / American Elsevier. Amsterdam and New York.
- 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. (1982). Human errors: A taxonomy for describing human malfunction in industrial installations. Journal of Occupational Accidents, 311-333.
- Rasmussen, J. (1983). Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics, 13(3), 257-266.
- 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. (2016): When paradigms collide at the road rail interface: evaluation of a sociotechnical systems theory design toolkit for cognitive work analysis, Ergonomics, 59(9), 1135-1157.
- 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
- Rongfang Liu, R. (2010). Railroad crossing safety − Final report. New Jersey Institute of Technology.

- Rudin-Brown, C. M., French- St. George, M., & Stuart, J.J. (2014). Human factors issues of accidents at passively controlled level crossings. Transportation Research Record, 2458, 96-103.
- Ryan, B., & Kallberg, V-P. (2013). Developing methodology in RESTRAIL for the preliminary evaluation of preventative measures for railway suicide and trespass. In: Dadashi, N., Scott, A., Wilson, J.R., and Mills, A. (eds.) The fourth international rail human factors conference, March, 2013, pp.89–98. London, UK: CRC Press, Taylor and Francis Group.
- Ryan, B., Kallberg, V.-P., Rådbo, H., Havârneanu, G.M., 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. doi:10.1080/00140139.2018.1485970
- SAFER-LC D1.1 (2017). Analysis of level crossing safety in Europe and beyond. Deliverable 1.1 of the SAFER-LC project. http://safer-lc.eu/IMG/pdf/saferlc_20171130_d11_v04_ffe_analysis_lc_safety_europe_and_beyond.pdf
- SAFER-LC D1.3 (2018a). Needs and requirements for improving level crossing safety. Deliverable 1.3 of the SAFER-LC project. http://safer-lc.eu/IMG/pdf/saferlc_20180130_d13_v08_uic_needs_and_requirements.pdf
- SAFER-LC D2.1 (2018b). 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 D4.1 (2018c). Implementation guidelines. Deliverable 4.1 of the SAFER-LC project.
- SAFER-LC D4.2 (2018d). Evaluation framework. Deliverable 4.2 of the SAFER-LC project.
- SafetyCube DSS (2018). https://www.roadsafety-dss.eu/#/methodology
- Salmon, P. M., Lenné, M. G. (2015). Miles away or just around the corner: systems thinking in road safety research and practice. Accident Analysis and Prevention, 74, 243–249.
- Salmon, P, M., Read, G, J, M., Stanton, N, A., Lenne, M, G. (2013). The crash at Kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings.
- 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.
- Searle, A., Di Milia, L. & Dawson, D. (2012). An investigation of risk-takers at railway level crossings. CRC for Rail Innovation. Brisbane.
- Sharit, J. (2006): Human Error. In: Gavriel Salvendy (Hg.): Handbook of Human Factors and Ergonomics. Hoboken, NJ: Wiley, 708–760.

- 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. IVol. 11 (2017) No: 3, 164–172.
- 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., Öztürk, E. & Karahan, G., (2017). Level crossing accidents and factors behind them. Deliverable D1.2 of the SAFER-LC project.
- 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
- Stefanova, T., Burkhardt, J-M., Filtness, A., Wullems, C., Rakotonirainy, A., & Delhomme, P. (2015). System based approach to investigate pedestrian behaviour at level crossings. Accident Analysis and Prevention, 81, 167–186.
- Summala, H. (1988). Risk control is not risk adjustment: The zero-risk theory of driver behaviour and its implications. Ergonomics, 31(4), 491–506.
- Summala, H. (1997). Hierarchical model of behavioural adaptation and traffic accidents. In: Rothengatter, T., Carbonell Vaya, E. (Eds.), Traffic and Transport Psychology: Theory and Application. Elsevier Science, Oxford, pp. 41–52.
- Theeuwes, J. & Godthelp, H. (1993). Self-explaining Roads. In de Kroes and Stoop (Eds). Safety of Transportation. Delft: University Press.
- UNECE (2017). Assessment of safety at level crossings in UNECE member countries and other selected countries and strategic framework for improving safety at level crossings. Geneva. United Nations. http://www.unece.org/fileadmin/DAM/trans/doc/2017/wp1/ECE-TRANS-WP1-2017-4e.pdf
- Vicente, K. J. (1999). Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Mahwah, NJ: Lawrence Erlbaum Associates.
- Walker, G. H., Stanton, N. A., Salmon, P. M., and Jenkins, D. P. (2008). A Review of Sociotechnical Systems Theory: A Classic Concept for New Command and Control Paradigms. Theoretical Issues in Ergonomics Science 9: 479–499.
- Ward, N. J., & Wilde, G. J. S. (1995). Field Observation of advanced warning/advisory signage for passive railway crossings with restricted lateral sightline visibility: An experimental investigation. Accident Analysis & Prevention, 27(2), 185−197.
- 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.

- Wilde, G. J. S. (1982). The theory of risk homeostasis: implications for safety and health. Risk Analysis 2, 209–225.
- Wilde, G. J. S. (1986). Beyond the concept of risk homeostatis: Suggestions for research and application towards the prevention of accidents and life-style-related disease. Accident Analysis and Prevention 18, 377–401.
- Wilde. G. J. S. (1994). Target Risk. PDE Publications. Canada.
- Wilde G. J. S. (1998). Risk homeostasis theory: an overview. Injury Prevention,4, 89–91.
- 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.
- Van Geem, C., Charman, S., Ahern, A., Anund, A., Sjogren, L., Pumberger, A., Grayson, G., & Helman, S. (2013). Speed Adaptation control by Self-explaining Roads (SPACE). 16th Road Safety on Four Continents Conference, Bejing, China 15-17 May 2013.
- Yeh, M., & Multer, J. (2008). Driver Behaviour at Highways – Railroad Grade Crossings: A literature Review form 1990-2006.

ANNEX A: HUMAN FACTORS ASSESSMENT TOOL (VERSION TO PRINT FOR PILOT TESTS)

Name of the measure being assessed	Brief description
....	....

CLASSIFICATION CRITERIA		
Factor	**Brief description**	**Indicator** ***(Tick all the cases that the measure applies to)***
Applicability to different LCs	*Specify the types and characteristics of LCs where the measure can be implemented*	Type of LCs ☐ Passive LCs without any warning devices ☐ 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 devices ☐ Active LCs with traffic lights Characteristics of LCs ☐ 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 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	*Specify the categories of LC users who are targeted by the measure*	☐ All road users MRU ☐ cars ☐ motorbikes / mopeds ☐ trucks / heavy vehicles ☐ buses / coaches ☐ farm / agricultural vehicles ☐ other (specify)………………………… VRU ☐ pedestrians ☐ cyclists ☐ other (specify)…………………………
Adaptation to individual characteristics and conditions of users	*Specify if the measure is specifically targeted at people with the following characteristics or conditions*	Gender ☐ Male ☐ Female Age ☐ children ☐ elderly ☐ all ages Disability ☐ vision loss and blindness ☐ hearing loss and deafness ☐ intellectual disability ☐ reduced mobility ☐ other (specify)………………………… Under influence of ☐ alcohol ☐ drugs ☐ medication Under skill impairing states ☐ fatigue ☐ stress ☐ Risk-seeking personality
Intended effect mechanism	*Specify 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 access to the LC ☐ Reduces the approach speeds of vehicles ☐ Increases the user's 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 ☐ Other (specify) …………………………………

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

Criterion	Brief description
Detectability	The measure can help the LC user detect relevant visual and auditory stimuli, therefore increasing the detectability of the LC or the approaching train

Write down brief descriptions of the expected and/or observed changes in road user's detection of the LC or train as a result of 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				

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 of the LC /or train 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
Identification	The measure can increase safety by helping the LC user identify relevant information in the environment and not be distracted by irrelevant information

Write down brief descriptions of the expected and/or observed changes in road user's identification of relevant information as a result of 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				

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 identification of a possible danger as the user is approaching the LC?**		
Answer modalities	N	The LC user is somewhat confused or distracted by this measure and therefore unable to identify potential dangers related to the crossing of the LC
	0	This measure has no intended influence on the attention or workload of the LC user
	1	
	2	
	3	
	4	
	5	LC users' attention is naturally directed towards identifying a potential danger despite being fatigued, distracted, or under high workload (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
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 expected and/or observed changes in road user's ability to elicit and retrieve relevant safety information as a result of 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				

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 expected and/or observed changes in road user's decisions as a result of 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				

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 help 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 expected and/or observed changes in road user's behavioural execution as a result of 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				

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 influence the safe execution of the approach and crossing behaviour?**		
Answer modalities	N	The LC user's crossing action is disturbed and becomes more difficult when this measure is in place
	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

<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):</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):</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):</td></tr>
<tr><td rowspan="2">Reliability</td><td rowspan="2">The estimated extent to which the users of the LC trust the system and know that it is fail-safe</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):</td></tr>
<tr><td rowspan="2">Usability</td><td rowspan="2">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</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):</td></tr>
</table>