Human reliability assessment under uncertainty - towards a formal method

Subeer Rangra\textsuperscript{a,*}, Mohamed Sallak\textsuperscript{a}, Walter Schön\textsuperscript{a}, Frédéric Vanderhaegen\textsuperscript{b}

\textsuperscript{a}Sorbonne Universités, Université de Technologie de Compiègne, CNRS Laboratoire Heudiasyc UMR 7253, Centre de Recherches de Royallieu CS 60 319, 60203 Compiègne, France
\textsuperscript{b}Université de Valenciennes et du Hainaut-Cambrésis, LAMIH, UMR CNRS 8201, F-59313, Valenciennes, France

Abstract

Humans are and will remain one of the critical constituents of a technological system. The study of human factors is a broad domain with equally varying applications. Furthermore, with the advent of new technologies in safety-critical systems there is always a need to ensure system safety and reliability in accordance with increasingly demanding certification standards. Human reliability is a cause of concern as hardware becomes increasingly reliable and relatively human error is rising in its share of causing an accident. Human Reliability Analysis (HRA) provides a way to quantify the risk associated with a human. This paper presents a discussion on the development of a HRA model for the domain of transportation, rail transport in particular. Railway specific human factors studies are analyzed to identify safety relevant factors in order to create a relevant and relatively applicable Performance Shaping Factor list. This list of factors is compared with railway specific studies to address domain specific concerns, further augmenting it with quantification levels for each. A discussion on our proposition towards the integration of HRA for obtaining human induced system-level risk taking into account uncertainty in data and current work’s positioning in proposed methodology is also included.

© 2015 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of AHFE Conference.

Keywords: Human factors; Human Reliability Analysis; Performance Shaping Factors; Rail transportation

* Corresponding author. Tel.: +33-344-237-922; fax: + 33-344-234-477.
E-mail address: subeer.rangra@utc.fr
1. Introduction

It has been widely reported in rail and road transportation that human error is often the leading cause of an accident. More recently a study [1] states that at least 75% of fatal railway accidents in Europe between 1990-2009 were due to human errors. However, a major Australian study on over 100 rail accident reports concluded that accidents attributed to human errors, were caused by the conditions in which train drivers had to work, indicating driver errors were in fact consequences, not the initial issue causing the accident [2]. Human Reliability Analysis (HRA) involves determining the likelihood of human error. In the early days of development of HRA methods, one of the original aims was to include the human controller in the PRA (Probabilistic Risk Assessment) of nuclear power plants [3]. Human reliability and human error can be defined in terms of the causes of human behavioral dysfunction and/or their consequences for the system. Most HRA methods are risk assessment-based or cognitive model-based methods. They assess or analyze the risks of human or system dysfunction due to human actions [4]. The need for a dedicated method to access human reliability was felt (i) an increasing involvement of humans in transportation accidents while hardware reliability has steadily improved and (ii) the availability of very few methods to measure the risk of human towards the safe operation of the system [5]. Most HRA methods provide a quantitative technique aimed at identifying the probability of occurrence of human error, known as Human Error Probability (HEP) [3]. Human factor experts broadly classify HRA models in to so-called first and second generation models, THERP (Technique for Human Error Rate Prediction) among other task-oriented models are regarded as first generation, and models factoring in the environment and the context in its calculations is considered a second generation model [3]. Usually experts believe human factors (HF) are not to be considered in isolation: environment, human cognitive state and context are also to be taken into account. Performance Shaping Factors (PSFs) offer a simplified way to take into account the context in HRA calculations. More specifically, PSFs allow the consideration of human’s own characteristics along with environment which affect human performance in a negative or positive manner [6]. However, for safety related analyses and subsequent risk calculations the interest is mostly towards their negative impact on human performance. PSFs are often categorized as internal and external (i.e. individual and situational characteristics), direct and indirect i.e. ones having one-to-one relation between magnitude of PSF and the measured quantity and ones whose magnitude is measured multivariate or subjectively [7]. Over the years most PSFs sets have gone through multiple revisions and critiques giving them a refined definition and hierarchical structuring among other classifications.

Rail transportation has multiple entities all contributing towards a safe and efficient transportation system. It is composed of multiple human actors (drivers, signalers, maintenance personnel, operational management) and signaling systems working in a synchronized way. There are various signaling standards in place in different countries; ERTMS (European Rail Traffic Management System) is a relatively new entrant. ERTMS is a major ‘European industrial project’ to enhance cross-border interoperability by creating a single Europe-wide standard for railway signaling. It is composed of the European Train Control System (ETCS), a standard for train control and GSM-R the GSM mobile communications standard for railway operations. Standardized signaling systems such as ERTMS provide a good application domain as it facilitates a wide acceptance of methodologies proposed around it.

This paper aims to present the work done towards a HRA model applicable to transport applications notably railways. Section 2 discusses some of the work done so far in the domain of rail transportation and performance shaping factors. Overview of proposed methodology and the subsequent work done is presented in Section 3 along with the explanation of the data sources used. Some remarks on the work presented in this paper are mentioned in Section 4. Our perspectives along with future work and conclusions are provided in Sections 5 and 6 respectively.

2. PSF and human factors in railways

In the railway domain, ergonomics and human factors are often used interchangeably and over the years have attracted large number of researchers. The review study in [8] indicates an increasing interest of railway stakeholders in human factors. Furthermore, as remarked earlier, a majority of transportation accidents, especially in automobiles and railways [1] are caused by human errors. Human factors at least in the domain of railways, pose a big challenge to analyze, to provide the stakeholders with tools to be able to quantify these factors and eventually
reduce or avoid human errors to make the system operation safer. However, most of the works reported so far show an absence of a concrete HRA model for railways.

A project aimed at developing the methods for human reliability assessment for German railways is focused on the importance of PSFs towards proposing a context related method [9]. They identified some PSFs for the train driver, specific to German national requirements. The work in [10] proposes a straight-forward model of working systems to analyze the influence factors on human performance. The proposed work gives an overview of human factors to a railway engineer dealing with certification requirements. Furthermore, they argue for the applicability of a PSF-based model in safety analysis of a human-barrier interaction. A railway specific PSFs taxonomy (R-PSF) was proposed in [11], it was developed from human factor literature review and validated against findings from railway accident and incident reports and expert opinions. The work in [12] an extension of [11] proposes a HRA model called Human Performance Railway Operational Index (HuPeROI), which considers a weighted PSF-based measure to account for the PSF’s influence. They have not however proposed methodology to calculate HEP and work is still in progress. Their objectives were oriented towards constructing a domain specific PSF with subsequent expert and accident analysis based validation of the said PSF set. Expert opinion is further needed for the quantification of PSF, its relative impact on human error (weightage) and correlations between PSFs to enable HEP calculation. Our work is focused towards HEP quantification. Experimental calculations provide a simpler method where one can observe the specific PSFs needed by breaking down quantifications needed one by one into specific scenarios. It can also be used to substantiate expert opinions. Also, even with the amount of in-depth work performed, the question of interdependencies between the railway PSFs still remains unanswered and by their admission remain ‘an issue of great concern’.

3. Towards a PSF based Human Reliability Analysis model for risk evaluation at system level

Quantification of human errors is but a small part of the overall risk assessment of the system. Taking a holistic approach to human error using a formal framework is needed in order to obtain a numerical value of its system-level risk. A formal framework is required for (including but not limited to): evaluation of environmental effects (PSFs), calculation of HEP and subsequently estimation of the risk at system level.

Fig. 1 gives an overview of methodology we propose. In general the first step will be the creation of HRA model to allow the quantification of human error. The second step will be the estimation of system level risk, much like the intentions of earlier HRA methods used along with PRA studies in nuclear domain. Such a metric will be highly useful to railway authorities and industrial actors alike due to the fact that at present European railway certification standards on one hand require quantitative risk analysis for safety critical systems (CENELEC EN 50129:2003), whereas on the other hand proposes the integration of human factors in it without providing any specific approach (CENELEC EN 50126:1999). Such a methodology will provide engineers and designers to understand human factors affecting the system safety and identify human machine interactions for which improvements are needed to increase system reliability. As visible in Fig. 1, this paper addresses Step 1, Step 2.a and part of Step 2.b of our long

![Fig. 1. Overview of our proposed methodology to access human error at system level.](image-url)
term goals. Step 1 and Step 2.a. are explained in the following sub-sections, viz. the identification of safety relevant activities pertaining to human controller and the identification of PSFs involved in human errors and their quantification.

3.1. Data sources

This section discusses the sources of data used for the creation of PSF specific to railway applications. The European Railway Agency (ERA) started the Human Factors project in 2012 and created a network bringing together human factor experts. The main objective behind this project was to promote and provide a human-centered perspective in the systematic integration of HF at multiple levels e.g. design, development, operation and maintenance in European railways. The final report produced in 2013 [13] consists of the analysis of data from over 16 countries including most EU countries and some non-EU countries such as Japan, USA, and Canada etc. Data was taken from both academic and industrial literature and covered a wide range of job roles. It aimed to gather as globally applicable definitions as possible in order to provide a global perspective. The overall aim of this particular report was to ‘provide a generic and high level view of human factors in railway operations and to identify safety relevant human activities associated with these functions’. Subsequently an expert-opinion based analysis of safety relevant activities on humans involved in railway operations was performed on the raw data gathered by researchers.

The results are presented in a multilevel hierarchy, starting from system level operational goals down to a human’s safety relevant activities. The top level system operational goals are defined as “Purpose/Goals - aim of the socio-technical system and a focus for human efforts”. In total seven such high level goals were identified e.g. maintain safety, provide efficient train service, optimize passenger comfort and journey, minimize environmental impact etc. These goals are then attributed at the second level to ‘human functional goals’, for example for all of the aforementioned Purposes/Goals it is necessary that ‘train movement must be controlled in all operational circumstances’. Subsequently 8 human functional goals were identified each branching from one or more Purposes/Goals. Each ‘human functional goal’ is further broken-down into multiple lower level ‘human functions’. A spreadsheet is produced describing the ‘human function’ under analysis, the context under which said ‘human function’ is executed, and the analysis of ‘safety relevant activities’ associated with it. This analysis includes data on ‘potential for errors’, ‘potential for recovery’, ‘consequences of errors’ and ‘mitigation’ within the system.

Some possible disadvantages of taking into account such a large amount of data are that some factors and safety relevant activities identified might not be applicable to specific technologies such as ERTMS. Also, the analysis spreadsheet does not give clear information on the source of said human function, making it harder to identify the applicability for a specific technology or national rule. Furthermore, all human roles are included when defining human functional goals, making it difficult to separate organizational and maintenance activities from control duties such as driving, signaling etc. The data, by their own admission is heavily influenced by UK-based sources. The reason being UK authorities have more publicly available documents and regulations. This makes it further difficult to identify applicability of their inferences. Starting with a limited yet widely spreading technology like ERTMS might make it more useful for some industrial actors. This study also takes into account all operational circumstances. However, most of the times in railway operation degraded cases and normal operations change the context for a human significantly. Therefore, some activities which are only encountered rarely have been taken into account same as normal modes of train operation. For example in degraded or special cases, workload and responsibility of driver for example in shunting mode increases and the high effect of these factors in that particular scenario might not be applicable to normal operation and vice versa.

The work on analyzing and aggregating PSFs for HRA purposes in [14] provides a fairly exhaustive data-set on PSFs. It gathered PSFs from multiple sources such as Human Event Repository and Analysis (HERA) [15], Information, Decision, and Action in crew Context (IDAC) [16] (a cognitive model) and information from expert workshops conducted to gather expert opinions. This work was done for the nuclear domain and funded by the US Nuclear Regulatory Commission’s Division of Research; it offers a relatively exhaustive set of PSFs, with their definitions and has considerations for their orthogonality. However, with the level of detail and exhaustiveness this work can be considered for generic and non-domain specific cases. Railway relevant information on PSFs to be considered were referenced form [11]. In addition to the R-PSF taxonomy, the subsequent validation activity also serves as a valuable resource.
3.2. A PSF based Human Reliability Analysis model for railways

This section describes the creation of a railway specific PSF-set from the data sources mentioned above. Notably from the ERA study spreadsheet [17] which gives detailed analysis of human activities in terms of safety relevant activities. The focus is on the subsection ‘potential for errors’ and ‘mitigation’, emphasizing on the factors affecting the human. One ‘human functional goal’ was chosen to adhere to our consideration of the train driver ‘To control train movements in all operational circumstances’, which includes nominal and degraded cases. This goal is further divided into 14 human functions. For each one of them we have ‘potential for errors’ listing human errors possible in this context, with a description of context. The descriptions of potential errors and context definitions give an indication as to what factors are responsible.

To give an example of the generation of PSFs from human functions and subsequent description thereof, a simple case is explained. The first entry in the spreadsheet under ‘human functional goal’ we consider the lower level human function ‘Take up control of train movement duties’. The last column in the spreadsheet, the potential for errors associated with this activity states ‘A lack of understanding of the information that is needed to appreciate the status of the system, possibly linked to inattention, memory failure.’ Distraction/Concentration or the absence of attention is a well-defined PSF and also in the rail accident analysis in the work for R-PSF [11] identified as predominant contributing factor that leads to an accident event in railways. Thus, it is identified as a PSF that we should consider. The identification of underlying factor for ‘memory failure’ cannot be identified and more information under mitigation is referred which states ‘improving experience’ along with ‘Protocols for communications and procedures for handovers’, it does give more information but not relevant to the factors responsible for the error, furthermore there is an absence of PSFs directly describing memory failure in both of our PSF sources. Hence, in this case only one PSF is considered. Similarly, for each of the remaining 14 human functions similar strategy was followed and such factors are identified; in some cases where a clear dominating factor fails to emerge we refer the provided definition of context of that human function to arrive at an association. A table is subsequently formulated to concisely define the PSFs list obtained. Since, we have the context represented by the list of PSFs; the next step is to have quantification thereof, to move towards the calculation of a HEP value.

Table 1. Generated PSF list with the definitions considered and assigned quantification levels.

<table>
<thead>
<tr>
<th>Performance Shaping Factor</th>
<th>Definition considered</th>
<th>Quantification levels (SPAR-H [6])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Have the correct knowledge to perform their jobs successfully and safely.</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Continuously train employees to ensure their skills are up to date and relevant.</td>
<td>Nominal</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>Communication</td>
<td>The ability of team members to pass information to each other and a shared understanding of the situation using system status, read-outs etc.</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>On the contrary distraction from what is important.</td>
<td>Nominal</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Poor</td>
</tr>
<tr>
<td>Concentration/Distraction</td>
<td>Attention to task/context – a general awareness of the system state.</td>
<td>High (Concentration)</td>
</tr>
<tr>
<td></td>
<td>On the contrary distraction from what is important.</td>
<td>Nominal</td>
</tr>
<tr>
<td></td>
<td>Advanced</td>
<td>Low</td>
</tr>
<tr>
<td>Experience</td>
<td>The accumulation of information and knowledge gained through interactions with the system and time spent in the work environment.</td>
<td>Advanced</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Sufficient</td>
</tr>
<tr>
<td>Task Load (Workload)</td>
<td>The actual task demand assigned to a person in terms of the number and type of tasks (varying complexity, importance, fault tolerance etc.)</td>
<td>Nominal</td>
</tr>
<tr>
<td></td>
<td>Task load can also be impacted by unplanned or emergency events (a distinction however not made in this work).</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Extra time</td>
<td>High</td>
</tr>
<tr>
<td>Time load (Workload)</td>
<td>Time to the number of tasks; this time perception can affect worker stress beyond the stress of having too many tasks.</td>
<td>Extra time</td>
</tr>
<tr>
<td></td>
<td>Available time to complete a task (currently we consider both detection and completion of the task).</td>
<td>Nominal</td>
</tr>
<tr>
<td></td>
<td>Barely adequate</td>
<td>Inadequate</td>
</tr>
</tbody>
</table>
This activity however poses a challenge, as on one hand it is difficult to accurately measure such subjective factors on human performance and on the other we do not have concrete railway specific studies and expert opinions to provide guidelines. Our approach in this work is to have a coarse-grain PSF quantification such as in SPAR-H [6] and others later substantiate that data from expert opinions and domain specific analysis. Table 1 shows the general definition of the PSFs extracted from ERA study in previous step and the quantification levels for each one of them. One of the reasons this list was kept short was to avoid the problem of inter-PSF relations, as we aim to tackle it at a later step in our work. Furthermore, one of the remarks in [11] after the accident analysis states that around 18 PSFs were responsible for more than 80% of the accidents, identified as one of the initial causes. Hence, it can be safely assumed that an exhaustive PSF list accounting for every possible factor affecting the human is a matter of data gathering and does not make a significant difference for most of the errors. Once a model which provides limited yet reliable quantification results is obtained, more factors can be added to increase the scope and applicability of the model.

4. Remarks and discussion

More often than not domain specific PSFs are extracted from an extensive analysis of accident reports. In such cases concrete information on factors affecting the human is sparse, secondly in most modern human-machine design at some level there are some ergonomic considerations made, rendering analysis from a 20 year old accident on driver conditions obsolete. Also as remarked in [9] the reliability of train driver tasks are considered with respect to its tasks and PSFs need to be identified to those tasks or specific scenarios to be able to obtain a trustworthy quantification. Both of these problems are dealt with by using ERA study as data source for creating PSFs list. Firstly it is relatively recent (releasing the report in 2013 and periodic release of documents after that [17]) keeping up with current regulations and research so the data is relatively relevant for most contemporary railway applications. Secondly the systematic approach of breaking down human functions into safety relevant activities and then identifying factors that might affect those safety relevant activities is useful as a general study. This also creates an association that can be traced back to relevant tasks associated with the PSF in question, allowing experimental setups to correspond to the PSF in question. The railway taxonomy of R-PSF [11] and others railway specific works are referred to determine the relevant PSFs, but there is an absence of clear and detailed domain specific definitions of PSFs hence more work is needed to define them non-ambiguously. It may also emphasized that all accident analysis reports are not created equal; most of them as remarked in [11] are sufficient to extract PSFs but are not enough for validation, requiring further work in accessing the reporting reliability of the authority responsible such as national transportation safety boards, independent national investigation bodies, and railway operators. Training and experience are often considered together in the most models [6]. We try to separate them on the basis of the level of human performance needed. One major difference between experience and training is that experience differs for every member of a crew, but all members undergo the similar type of training. Errors that a human makes on its own and/or a relatively simple case of not following the procedure, e.g. 'incorrect entry of details on control systems' can be considered to be a training issue, whereas in more team based tasks and situations where high level of competency is needed, experience plays a major role.

5. Perspectives and future work

The immediate next step of this work is towards the calculation of HEP as a function of PSFs associated with it. Uncertainty consideration in HEP values is recommended in Good Practices for Implementing HRA report to the US Nuclear Regulatory Commission [18]. A VBS (Valuation Based Systems) [19] based model will be used to calculate the HEP. The nominal HEP values at this stage will be sourced from experimentations or other domains with a reasonable degree of uncertainty. VBS will also aid in addressing uncertainty in the final calculation of system-level risk, as also recommended in [18]. Propagating the uncertainty in HEP values is required to obtain trustable system-level results. Such a method will allow a way to measure the effect of human error on the system reliability; taking a cue from one of the original goals of HRA models to obtain accurate PRA assessments of the system. Human reliability activities often suffer from the absence of data on HEP values, more so in a domain relatively new to human reliability such as transportation. Experimental calculations and expert elicitations are
envisaged; nevertheless absence of data is an issue for HEP calculation. Subsequently, a probability interval can be used for HEP instead of precise values, the use of VBS will allow calculations with probability intervals and their effect on system reliability can be observed nevertheless. Overlap between the definitions of PSFs is not considered in this work. Partly because the number of errors considered and subsequently factors extracted are few and partly due to the definitions thereof we consider, both resulting in an insignificant overlap. An overlap among the definitions of PSF will inevitably result in erroneous quantification of human error (double counting) and remains one of the frequent critiques of PSF-based methods [10]. This issue can be avoided however if we use the model to compare two different human-machine systems from the point-of-view of their error causing/preventing qualities. Such a usage is possible early in the lifecycle of HMI design [20], when comparing different HMI designs to reduce human error with the benefit of lower development costs.

Human behavior is inherently complex and trying to quantify it in a dynamic environment, a risk-analysis perspective on top of that poses considerable challenges. First issue is the very definition of human error, various point-of-views have been discussed over the years. The work in [21] considers the difference between prescribed and performed behavior to account for the violations and added human behaviors intentional or otherwise. Taking into account a flexible human error definition is required in order to obtain accurate data from experimental calculations. Secondly the resilience of humans as well as hardware [22], which is one’s ability to recover from a disturbance, needs to be addressed when we pass from a static definition of human errors to more complex dynamic evolution of human machine interaction in nominal and degraded cases. Defining the behavior of human for our context is required to limit our problem-set and obtain relevant results. Furthermore, addressing these issues is critical to demonstrate the abilities of our model to stand up against the numerous critiques of HRA models and to acceptably quantify human erroneous behavior. Subsequently there is the question of validation of PSFs, in terms of both their presence and quantification values. Work done in [11] initially generated a long list of PSFs from previous propositions, indifferent of application domain. They subsequently consulted experts together with accident analysis reports to validate their PSFs. Expert opinions can be explored along with works such as ERA study to justify the selected PSFs. Work concerning reliability of automobile driver was not covered in present work, however that remains one of our goals to propose a generic model for transportation. This will be considered along with our focus on train driver functions and establishing a minimum common set of PSFs with an automobile driver. To obtain data for nominal HEP calculations we aim to use simulator experiments for which we will employ railway simulators (ERTMS driver simulator and Route Control Center Systems (RCCS)) available at the Université de Technologie de Compiègne. Experimental protocols will be defined as specific scenarios which take into account single or multiple factors. It is fairly difficult to accurately mimic the real context and environment; however a tradeoff can be reached by the use of expert opinions and uncertainty consideration.

6. Conclusions

Human involvement exists at various levels of most public transportation systems and is often concluded to be at fault in majority of accident and incident analysis reports. In increasingly complex and evolving modern technical systems where large resources are allocated towards ensuring system’s operational safety, it becomes necessary to analyze the actions of human operator who directly or indirectly influences system operations. We have presented the first steps taken towards a human reliability analysis model for transportation systems. The work towards a PSF-based HRA model for transportation has been discussed. This HRA model forms a part of the methodology to assess human error at system level; this proposition is also briefly explained. Data on human functions and safety relevant activities thereof have been analyzed from a European Railway Agency study in order to generate a railway specific PSF set. The generated PSF list is then defined and adapted for our application needs referring domain specific studies. Furthermore a quantification level is associated with each PSF to aid in calculation of HEP. Existing problems, the data sources and works needed to address those problems are identified. Finally goals for next steps to arrive at a limited scope but robust HRA model are charted in this paper. At this stage we do not aim for this model to be encompassing all possible human behaviors and provide a strict human error probability, but a method able to measure HEP values with a reasonable degree of uncertainty for the factors frequently observed.
Acknowledgements

This work was carried out and funded in the framework of the Labex MS2T. It was supported by the French Government through the program “Investments for the Future” managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02). This work was also supported by the National Agency for Research, through the project ANR-13-JS03-0007 RECIF.

References