# Developing a Human Performance Railway Operational Index to enhance safety of railway operations

Miltiadis KYRIAKIDIS (Mechanical Engineer, MSc)

October 2013

A thesis submitted as fulfilment of the requirements for the Diploma of the Imperial College (DIC), PhD degree of Imperial College London

Centre for Transport Studies
Department of Civil and Environmental Engineering
Imperial College London, United Kingdom

*This work is dedicated to*
*my FAMILY*

*especially*

*to Ifigeneia and Tasos, my beloved parents*

*to Anna and Ifigeneia, my grandma and my niece, who I so much adore*

*to Nikos, Theofylaktos and Alexandros, my best friends*

*and to Miltiadis and Athina, who I miss*

# DECLARATION

I hereby declare that the entire work presented in this thesis has been personally carried out. Where sources of information or the work of others have been used, they are fully cited and referenced and/or appropriate acknowledgement is given.

.....

Miltos[D] Kyriakidis

.....

Dr. Arnab Majumdar

.....

Professor Washington Yotto Ochieng

'The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work'.

# ABSTRACT

The railway system is a complex network that involves continuous interaction of human operators with technology, procedures and regulations to ensure safe and efficient operations. From an architectural perspective, the complexity of the interactions presents a risk of failure with the consequence that safety incidents and accidents may occur. The common approach to the development of measures for mitigating such occurrences is the retrospective analysis of accidents and incidents; in order firstly to identify, classify and acknowledge the contributing factors and secondly, to suggest mitigation strategies.

Research undertaken globally using retrospective analysis indicates that a large number of railway accidents and incidents are associated with human errors due to degraded human performance. In particular, it has been shown that train operators (drivers, signallers and controllers) account for the majority of accidents and incidents. For example, between 1990 and 2009, at least 75% of fatal railway accidents in Europe were due to excessive speed, signal passed at danger or signalling/dispatching errors.

There has been a significant research effort to examine, identify and understand the factors that affect human performance in railway operations, so as to prevent conditions related to degrade performance and to reduce the probability of human errors. However current methods, developed on the principles of Human Reliability Analysis (HRA), are based on research from other domains, including nuclear, oil and gas, and aviation. Hence, they are not suited to the rail industry and can be difficult to apply reliably to railway specific operations.

Moreover, in the case of the factors that affect human performance, current methodologies have either adopted lists of factors from other domains or slightly modified existing lists, and then applied them to the railway industry. In addition, even in the cases where the lists of factors have been modified, such alterations have been designed on the basis of regional accident and incident analysis. Although the number of factors that influence performance can be claimed to be limited, e.g. fatigue, training, organisational culture and system design, the analysis of only regional occurrences does not provide analysts with a worldwide perspective of the significance of factors on human performance.

Therefore, this thesis addresses the current limitations and proposes a new framework to identify the factors that affect the performance of railway operators, and assess human performance.

In particular, this thesis developed for the first time a novel and comprehensive taxonomy for railway operations, referred to as the Railway-Performance Shaping Factors (R-PSFs) taxonomy. The taxonomy is derived from a variety of sources including: extensive literature review, operators' hierarchical task analysis, and the analysis of global accidents and incidents. Subject matter experts validated the taxonomy. Results identified 43 contributing factors, whilst further statistical analysis indicates that 12 out of 43 factors are responsible for more than 90% of total occurrences regardless of the type of network, responsibility and severity of consequences. Unlike current taxonomies, the framework developed accounts for both the influence of each individual factor and the dynamic interactive influence of the factors due to their mutual dependencies. It is recommended that the R-PSFs taxonomy be used by railway stakeholders to enhance the Safety Management Systems of their organisation. In addition the taxonomy can be used as part of the training program of the organisations in order to inform and engage the railway personnel with respect to the factors that primarily affect their performance. Finally, the taxonomy is recommended for use by the investigator stakeholders to obtain information about the human aspect that may have led to railway occurrences.

This thesis also developed, tested and validated a framework, referred to as the Human Performance (HuPeROI) to enhance safety in railway operations. Based on the 12 largest contributing factors, the HuPeROI is a novel scheme to assess human performance, as function of the various R-PSFs. The HuPeROI for the first time introduces an approach to quantify the impact of each of the factors that affect human performance accounting for all the dependencies amongst those factors. HuPeROI has been developed by integrating the generic concept of two techniques, the Analytic Network Process and the Success Likelihood Index Methodology (SLIM). The former is one of the best known and widely used multi-criteria decision making techniques and was used to evaluate the influence of each R-PSF on operators' performance. SLIM was applied to rate the importance of each of the R-PSFs for different operational actions and finally to estimate the reliability index for these actions. The HuPeROI framework was demonstrated in a case study in three different types of railway operations: regional, high-speed and underground, and helps to define the influence of each individual factor on human performance as well as to indicate the relative likelihoods of different human errors.

Finally, both the R-PSFs taxonomy and HuPeROI can be transferred and used with minor modifications not only in other railway procedures, e.g. maintenance, but also domains, e.g. aviation, maritime and oil.

# ACKNOWLEDGMENTS

This research has been made possible thanks to my supervisors Dr. Arnab Majumdar and Prof. Washington Y. Ochieng who gave me the opportunity and support with all possible means to achieve this goal. My deepest appreciation and gratitude go to them.

I am also grateful to The Lloyd's Register Foundation for funding this research.

This work would not have been completed without the collaboration of several experts in railway and safety domains from various national and international organisations. These include the Swiss Federal Railways (SBB), UK Network Rail, Rail Safety and Standards Board (RSSB), First Capital, First Hull Trains and First ScotRail, Piccadilly Line and London Underground Limited, Norwegian National Railway Administration (JBV), European Railway Agency (ERA), Lloyd's Register Europe and ETH Zurich. A special thank goes to Andreas Hoenger, Ann Mills, Huw Gibson, Stuart Smith, Seamous Scallon, Brian Christie, Stephen O'Brien, Erica Hudson and Gudela Grote. I thank them all for their fruitful collaboration for my PhD research.

At the Centre for Transport Studies many people offered their help at various stages and aspects of my work and I wish to thank them all. Many thanks to Ms. Peri Smith and Mr. Robin Hirsch for their contribution to my work.

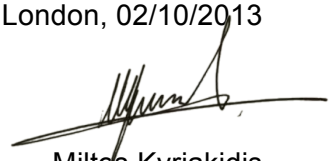
A great thank to all my office mates for the joyful work environment. Especially, I would like to thank Sabine, Bush and Milena.

I would also like to thank Nicole Kalas for reviewing, editing and formatting this thesis.

Finally, I want to express my deepest gratitude to my family, for their constant support and encouragements. I am dedicating this work to them.

*"darauf zu hören, was man machen möchte"*

London, 02/10/2013



Miltos Kyriakidis

# TABLE OF CONTENTS

*Declaration* ..... *5*

*Abstract* ..... *7*

*Acknowledgments* ..... *10*

*Table of Contents* ..... *11*

*List of Figures* ..... *19*

*List of Tables* ..... *22*

*List of Abbreviations* ..... *25*

*Chapter 1 Introduction* ..... *28*

**1.1 Background to the problem** ..... **28**

**1.2 Aim and objectives** ..... **31**

**1.3 Outline of thesis** ..... **31**

*Chapter 2 Fundamentals of the Railway System* ..... *34*

**2.1 Evolution of the railway system** ..... **34**

**2.2 The modern railway system** ..... **37**

2.2.1 Definition of the modern railway system ..... 37

2.2.2 The railway *controlled* components ..... 39

2.2.2.1 Infrastructure ..... 39

2.2.2.1.1 Track ..... 40

2.2.2.1.2 Signalling, command and train control ..... 41

2.2.2.1.3 Electrification ..... 44

2.2.2.1.4 Premises and structures ..... 44

2.2.2.2 Trains ..... 45

2.2.2.3 Personnel ..... 46

2.2.2.3.1 Train driver ..... 46

2.2.2.3.2 Signaller ..... 47

2.2.2.3.3 Controller ..... 47

2.2.2.3.4 Conductor ..... 48

2.2.2.3.5 Station staff ..... 48

2.2.2.3.6 Maintenance personnel ..... 48

2.2.3 The railway *uncontrolled* components ..... 48
2.2.3.1 Third parties ..... 48
2.2.3.1.1 Customers - Passengers ..... 49
2.2.3.1.2 Trespassers ..... 49
2.2.3.1.3 Individuals at Level Crossings ..... 49
2.2.3.1.4 Neighbourhoods ..... 49
2.2.4 The main railways stakeholders ..... 49
2.2.4.1 Train operators ..... 50
2.2.4.2 Infrastructure managers ..... 50
2.2.4.3 Infrastructure maintenance contractors ..... 50
2.2.4.4 Rolling Stock Companies ..... 50
2.2.4.5 Regulators and safety bodies ..... 50
2.2.4.6 Suppliers ..... 51

**2.3 The Railway Operational System Architecture framework ..... 51**
2.3.1 The ROSA - Literature review ..... 51
2.3.2 On-site observations ..... 58
2.3.3 The Railway Operational System Architecture scheme ..... 61

**2.4 Summary ..... 64**

***Chapter 3 Fundamentals of Railway Safety ..... 65***

**3.1 Introduction to railway safety ..... 65**

**3.2 Definitions and description of railway occurrences ..... 67**
3.2.1 The European Railway Agency (ERA) definitions ..... 68
3.2.2 The Federal Railroad Administration (FRA) definitions ..... 69
3.2.3 The Canadian Transport Safety Bureau (CTSB) definitions ..... 70
3.2.4 The Australian Transport Safety Bureau (ATSB) definitions ..... 70
3.2.5 The International Union of Railways (UIC) definitions ..... 71
3.2.6 Accident causes and precursors to occurrences ..... 72

**3.3 Railways safety statistics ..... 73**
3.3.1 Development of railway safety ..... 73
3.3.2 Railway accidents breakdown ..... 79
3.3.3 Contribution of railway constituents to safety ..... 85

**3.4 Summary ..... 89**

***Chapter 4 Human Error - Human Performance Analysis Techniques ..... 91***

**4.1 Concept of operational error in railways ..... 91**

**4.2 Definitions of human error ..... 92**

**4.3 Classifications of human error ..... 94**
4.3.1 The phenomenological approach ..... 94
4.3.2 The cognitive approach ..... 95
4.3.3 Summary of errors classifications ..... 96
**4.4 Models of human error ..... 97**
4.4.1 Person approach ..... 97
4.4.1.1 The Generic Error Modelling System (GEMS) ..... 98
4.4.2 System approach ..... 101
4.4.2.1 The sequential models ..... 102
4.4.2.2 The epidemiological models ..... 103
4.4.2.3 The systemic models ..... 106
4.4.3 Summary of human error modelling ..... 107
**4.5 Human Factors ..... 108**
4.5.1 Definition of Human Factors ..... 108
4.5.2 Railway Human Factors ..... 109
**4.6 Human performance and Human reliability ..... 109**
**4.7 Overview of Human Performance Analysis ..... 110**
**4.8 Synopsis of HPA techniques ..... 111**
4.8.1 First and second generation methods ..... 112
4.8.2 Qualitative and quantitative methods ..... 113
4.8.3 Methods with performance shaping factors ..... 113
4.8.4 Holistic and analytic methods ..... 114
4.8.5 Time dependence ..... 114
4.8.6 Data used ..... 114
**4.9 Performance Shaping Factors ..... 115**
**4.10 Review of Human Performance Analysis Techniques ..... 117**
4.10.1 Railway Human Performance Analysis approaches ..... 118
4.10.1.1 Analysing and Modelling Train Driver Performance ..... 120
4.10.1.2 Driver Performance Modelling to support Railway Safety ..... 121
4.10.1.3 Human Factors in the Railway System Safety Analysis Process ..... 122
4.10.1.4 Railway Action Reliability Assessment (RARA) ..... 124
4.10.1.5 Summary of railway HPA approaches limitations ..... 126
4.10.2 Cross - industry Human Performance Analysis techniques ..... 126
4.10.2.1 THERP (Technique for Human Error Rate and Prediction) ..... 132
4.10.2.2 SLIM (Success Likelihood Index Method) ..... 132
4.10.2.3 ATHEANA (A Technique for Human Event Analysis) ..... 134
4.10.2.4 HEART (Human Error Assessment and Reduction Technique) ..... 135

4.10.2.5 CARA (Controller Action Reliability Assessment) ... 136
4.10.2.6 CREAM (Cognitive Reliability and Error Analysis Method) ... 137
4.10.2.7 HRMS (Human Reliability Management System) ... 139
4.10.2.8 HEPI (Human Error Probability Index) ... 140
4.10.2.9 SPAR-H (Standardized Plant Analysis Risk Human Reliability Analysis) ... 140
4.10.2.10 Summary of cross – industry HPA techniques limitations ... 141

**4.11 Summary ... 143**

***Chapter 5 The Railway Performance Shaping Factors Taxonomy – Literature Review And Initial Development ... 145***

**5.1 The characteristics of R-PSFs taxonomy ... 145**

**5.2 Methodology to develop the R-PSFs taxonomy ... 146**

**5.3 Review of existing PSFs taxonomies ... 149**

5.3.1 Human Factors in the Railway System (TRACEr-Rail) ... 149
5.3.2 The Railway Action Reliability Assessment (RARA) ... 149
5.3.3 Human Factors Analysis Classification System Railways (HFACS-RR) ... 150
5.3.4 Human Error Risk Management for Engineering Systems (HERMES) ... 151
5.3.5 Human Factors in the Railway System Safety Analysis Process ... 151
5.3.6 Technique for Human Error Rate Prediction (THERP) ... 154
5.3.7 Human Error Assessment and Reduction Technique (HEART) ... 155
5.3.8 Human Event Repository and Analysis (HERA) ... 155
5.3.9 A Technique for Human Error Analysis (ATHEANA) ... 156
5.3.10 Controller Action Reliability Assessment (CARA) ... 156
5.3.11 Human Error in Air Traffic Management (HERA-JANUS) ... 156
5.3.12 Cognitive Reliability and Error Analysis Method (CREAM) ... 157
5.3.13 Human Reliability Management System (HRMS) ... 158
5.3.14 Simplified Plant Analysis Risk Human Reliability Assessment (SPAR-H) ... 158
5.3.15 Success Likelihood Index Methodology (SLIM) ... 158
5.3.16 Human Error Probability Index (HEPI) ... 159
5.3.17 Summary of findings ... 159

**5.4 Identification of dynamic R-PSFs ... 160**

**5.5 Augmentation with operators hierarchical task analysis ... 161**

**5.6 Presentation of the R-PSFs taxonomy ... 161**

**5.7 Consolidation of taxonomy with real data ... 164**

**5.8 Data characteristics ... 165**

**5.9 Data requirements ... 165**

**5.10 Data quality criteria ..... 166**
5.10.1 Accessibility ..... 167
5.10.2 Accuracy ..... 167
5.10.3 Completeness ..... 168
5.10.4 Consistency ..... 168
5.10.5 Interpretability ..... 168
5.10.6 Relevance ..... 169
5.10.7 Timeliness ..... 169
5.10.8 Credential ..... 169

**5.11 Background of safety maturity ..... 170**
5.11.1 The Safety Maturity Model (SMM) ..... 171
5.11.2 The SMM results for metro organisations ..... 173

**5.12 Summary ..... 176**

***Chapter 6 The Railway Performance Shaping Factors Taxonomy – Consolidation, Results and Analysis ..... 178***

**6.1 Credential assessment of railway organisations ..... 178**

**6.2 Extracting information from investigation reports ..... 179**

**6.3 Review of initial dataset ..... 182**
6.3.1 Identification of R-PSFs from reports ..... 182
6.3.2 Validation and assessment with SMEs ..... 184
6.3.3 Dependencies amongst the R-PSFs ..... 191
6.3.4 Discussion of findings ..... 195

**6.4 Analysis of the complete dataset ..... 197**
6.4.1 Descriptive statistics ..... 198
6.4.2 The R-PSFs *lite* – combination of R-PSFs for further analysis ..... 206
6.4.3 Examination of relationships between two variables ..... 209
6.4.4 Examination of relationships between more than two variables ..... 218
6.4.5 Multivariate analysis ..... 222
6.4.5.1 Binomial logistic regression ..... 223
6.4.5.2 Ordinal logistic regression ..... 224
6.4.6 Discussion of findings - Recommendations ..... 226

**6.5 Summary ..... 228**

***Chapter 7 The Human Performance Railway Operational Index ..... 230***

**7.1 The HuPeROI background ..... 230**

**7.2 An overview of the multi-criteria decision making methods ..... 231**

7.2.1 Value measurement models ... 233
7.2.2 Goal, aspiration or reference level models ... 233
7.2.3 Outranking models ... 233
7.2.4 Characteristics of the appropriate MCDM approach ... 234

**7.3 The Analytic Network Process ... 234**
7.3.1 The ANP framework - Developing the structure of the decision model ... 235
7.3.2 The ANP pairwise comparisons and relative weight estimation ... 236
7.3.3 The ANP supermatrices ... 239
7.3.4 The ANP final priorities ... 240

**7.4 The integration of ANP and SLIM methodologies – the HuPeROI calculation 240**

**7.5 The implementation of the HuPeROI ... 245**
7.5.1 Signals passed at danger (SPADs) scenario ... 245
7.5.2 Sampling of experts ... 247
7.5.3 Survey methodology – Responses collection ... 249
7.5.4 Questionnaire design ... 250
7.5.5 Preliminary survey ... 251
7.5.6 Completion of survey ... 251
7.5.7 The assessment (weightings) of the R-PSFs ... 251
7.5.7.1 Elements pairwise comparisons – the unweighted supermatrix ... 253
7.5.7.2 Categories pairwise comparisons – the cluster supermatrix ... 257
7.5.7.3 The weighted supermatrix ... 259
7.5.7.4 The limit supermatrix – the R-PSFs final weightings ... 261
7.5.8 The ratings of the R-PSFs ... 263
7.5.9 The HuPeROI calculations ... 265
7.5.10 The change of overall performance under certain conditions ... 267
7.5.11 Possible sources of errors and limitations of the HuPeROI ... 267
7.5.11.1 Possible sources of errors ... 267
7.5.11.2 Limitations of the HuPeROI ... 268

**7.6 Summary ... 269**

***Chapter 8 Analysis and Results of the HuPeROI Case Study ... 270***

**8.1 The sample characteristics ... 270**

**8.2 The weightings of the R-PSFs *lite* categories ... 273**
8.2.1 Influence of the R-PSFs categories on dynamic personal factors ... 275
8.2.2 Influence of the R-PSFs categories on personal factors ... 276
8.2.3 Influence of the R-PSFs categories on organisational factors ... 277
8.2.4 Influence of the R-PSFs categories on team factors ... 278
8.2.5 Influence of the R-PSFs categories on task factors ... 279

8.2.6 Influence of the R-PSFs categories on system factors ..... 280

**8.3 The final priorities of the R-PSFs elements ..... 282**

8.3.1 Distraction – Loss of concentration ..... 283
8.3.2 Fatigue ..... 284
8.3.3 Perception ..... 285
8.3.4 Quality of procedures ..... 285
8.3.5 Safety culture ..... 286
8.3.6 Supervision ..... 287
8.3.7 Familiarity ..... 287
8.3.8 Training ..... 288
8.3.9 System design ..... 289
8.3.10 Workload ..... 289
8.3.11 Communication between employees ..... 290
8.3.12 Quality of information ..... 290
8.3.13 Review of the R-PSFs elements weightings ..... 291

**8.4 The ratings of the R-PSFs elements ..... 294**

8.4.1 The SMEs ratings for “signal detection” ..... 294
8.4.2 The SMEs ratings for “signal interpretation” ..... 296
8.4.3 The SMEs ratings for “executed action” ..... 297
8.4.4 Effect of actions on the SMEs ratings ..... 299

**8.5 The HuPeROI calculations – the relative likelihoods of errors ..... 301**

**8.6 Validation of the HuPeROI findings ..... 305**

8.6.1 Internal validation ..... 305
8.6.2 External validation ..... 306

**8.7 The influence of external conditions on performance ..... 307**

**8.8 Summary ..... 308**

***Chapter 9 Conclusions ..... 310***

**9.1 Revisiting the research objectives ..... 310**

**9.2 Conclusions ..... 311**

9.2.1 The modern railway system and its operational architecture ..... 311
9.2.2 Human Error – Human Performance Analysis Techniques ..... 312
9.2.3 The Railway Performance Shaping Factors taxonomy ..... 312
9.2.4 The Safety Maturity Model ..... 313
9.2.5 The Human Performance Railway Operational Index (HuPeROI) ..... 313

**9.3 Future work and recommendations ..... 314**

**9.4 Publications – conference presentations relating to this work ..... 315**

9.4.1 Publication format: book chapters ..... 315

9.4.2 Publication format: journals – published ..... 315

9.4.3 Publication format: journals – in press ..... 315

9.4.4 Publication format: conference proceedings ..... 316

9.4.5 Conference presentations ..... 316

***References ..... 317***

***Appendices ..... 333***

Appendix I Implementation of the R-PSFs taxonomy ..... 334

Appendix II Overview of contextual factors categories ..... 335

Appendix III Overview of individual contextual factors ..... 339

Appendix IV Definitions of proposed R-PSFs ..... 345

Appendix V The Safety Maturity Model features ..... 360

Appendix VI Report on near-miss at Hanger Lane junction ..... 364

Appendix VII The Swiss questionnaire (English version) ..... 380

Appendix VIII Crosstabulations of reference database ..... 388

Appendix IX Loglinear analysis tests ..... 404

Appendix X Material to support SME's common understanding of R-PSFs ..... 408

Appendix XI The R-PSFs "weighting" and "ranking" questionnaire ..... 411

Appendix XII The ratings of the R-PSFs per human error and classification group ..... 422

# LIST OF FIGURES

Figure 1-1 Thesis overview ..... 33
Figure 2-1 Railway transport system ..... 39
Figure 2-2 Principle of track structure (Esveld, 2001) ..... 40
Figure 2-3 Points and Crossings description (Railway Technical Web Pages, 2012b) ..... 41
Figure 2-4 Signalling blocks protection (Pasquini et al., 2004) ..... 42
Figure 2-5 Driver advisory system architecture (Rail Safety and Standards Board, 2009) ..... 52
Figure 2-6 The top levels of a hierarchical task analysis for UK train driving (Rail Safety and Standards Board, 2009) ..... 54
Figure 2-7 Train driving activities in the UK (Rail Safety and Standards Board, 2009) ..... 55
Figure 2-8 The top levels of a hierarchical task analysis for train signaller ..... 57
Figure 2-9 The top levels of a hierarchical task analysis for train controller ..... 58
Figure 2-11 Part of signalling and controlling system, Piccadilly Line, London Underground ..... 60
Figure 2-12 The Railway Operational System Architecture (adapted from Rail Safety and Standards Board, 2009 and modified) ..... 62
Figure 3-1 Fatal train collisions and derailments per billion train-kilometers in 1990–2011 for the EU-27, Switzerland and Norway (European Railway Agency, 2012c) ..... 74
Figure 3-2 Train fatalities and train accidents with at least 5 fatalities (1990–2011) (European Railway Agency, 2012c) ..... 74
Figure 3-3 US train accidents - incidents per million-train miles and train accidents - incidents fatalities (Federal Railroad Administration, 2012) ..... 76
Figure 3-4 Canadian train accidents and main track accidents per million main track train miles (Transportation Safety Board of Canada, 2012) ..... 77
Figure 3-5 Australian train derailments, collisions and fatalities (Australian Transport Safety Bureau, 2012) ..... 78
Figure 3-6 Types of accidents in Europe based on the ERA (left) and UIC (right) databases ..... 81
Figure 3-7 Types of railway accidents in Canada (left) and Australia (right) ..... 82
Figure 3-8 Types of railway accidents in the US ..... 83
Figure 3-9 The UIC primary causes to accidents ..... 86
Figure 3-10 The UIC second level causes to accidents ..... 86
Figure 3-11 Precursors to the EU train accidents (European Railway Agency, 2011) ..... 87
Figure 3-12 Primary causes to US train accidents (Federal Railroad Administration, 2012) ..... 88
Figure 3-13 Canadian causes to train derailments (left) and collisions (right) (Transportation Safety Board of Canada, 2012) ..... 89
Figure 4-1 Process of operational error in railway operations by (Wickens, 1992) as cited in (Rail Safety and Standards Board, 2004b) ..... 92

Figure 4-2 The continuum between conscious and automatic behaviour (Reason, 1990) ..... 98
Figure 4-3 Unsafe acts framework (Health and Safety Executive, 1999, Rail Safety and Standards Board, 2011) ..... 99
Figure 4-4 Domino model of accident causation and human error exploration (Qureshi, 2007) ..... 103
Figure 4-5 The Swiss Cheese model of human error, adapted from (Rail Safety and Standards Board, 2011) ..... 104
Figure 4-6 The Clapham Junction accident - the Swiss cheese model, adapted from (Rail Safety and Standards Board, 2011) ..... 105
Figure 4-7 The Human Reliability Analysis process ..... 111
Figure 4-8 Classification of HPA techniques (Pyy, 2000, Boring, 2005) ..... 112
Figure 4-9 Review of HPA techniques ..... 118
Figure 4-10 Situational model of driver performance in interacting with AWS (McLeod et al., 2005) 120
Figure 4-11 Model of working system and PSFs (Hammerl and Vanderhaegen, 2009) ..... 123
Figure 4-12 Proposed relation between control mode and reliability (Hollnagel, 1998) ..... 138
Figure 4-13 Relations between CPC score and control modes (Hollnagel, 1998) ..... 138
Figure 5-1 Methodology to develop the R-PSFs Taxonomy ..... 147
Figure 5-2 Detailed framework of the R-PSFs taxonomy development ..... 148
Figure 5-3 The structure of R-PSFs taxonomy ..... 163
Figure 5-4 The proposed safety maturity model ..... 172
Figure 5-5 Maturity score versus reported injuries ..... 175
Figure 6-1 Maturity scores of the integrated company and infrastructure manager ..... 179
Figure 6-2 The identified R- PSFs per type of event from the initial set of reports ..... 183
Figure 6-3 Most frequently identified R- PSFs per type of operator ..... 186
Figure 6-4 The network of interdependences between the categories of R-PSFs ..... 195
Figure 6-5 The methodology for the analysis of the occurrences ..... 197
Figure 6-6 Type of collected reports per country of origin ..... 199
Figure 6-7 Type of train involved per type of event ..... 200
Figure 6-8 Type of responsible personnel per type of event ..... 200
Figure 6-9 Type of location per type of event ..... 201
Figure 6-10 Type of occurrence per type of event ..... 201
Figure 6-11 Associated events to types of occurrences ..... 202
Figure 6-12 Type of immediate causes per type of event ..... 202
Figure 6-13 Type of railway protection system per type of event ..... 203
Figure 6-14 The complete list of the identified R-PSFs per type of event ..... 205
Figure 6-15 The identified R-PSFs for incidents ..... 208
Figure 6-16 The identified R-PSFs for accidents ..... 208
Figure 6-17 The identified R-PSFs for serious accidents ..... 208
Figure 7-1 Classification of the MCDM methodologies (based on Belton and Stewart, 2002) ..... 232
Figure 7-2 Types of components in the network scheme (Saaty and Vargas, 2006) ..... 236
Figure 7-3 The framework to develop and calculate the HuPeROI ..... 243

Figure 7-4 The cluster dependencies for the pairwise comparisons ..... 252
Figure 7-5 The direct and indirect influence of factors categories ..... 259
Figure 7-6 The rating of R-PSFs per type of error ..... 265
Figure 8-1 Comparison of R-PSFs elements weighting by SMEs ..... 291
Figure 8-2 Frequency of the R-PSFs elements in the UK SPAD occurrences ..... 293
Figure 8-3 R-PSFs ratings for signal not detected ..... 296
Figure 8-4 R-PSFs ratings for signal wrongly interpreted ..... 297
Figure 8-5 R-PSFs ratings for action not executed correctly ..... 299
Figure 8-6 The HuPeROI calculations per type of action ..... 301
Figure 8-7 The HuPeROI calculations per type of SME and type of action ..... 303
Figure 8-8 The HuPeROI calculations per train driver, TOC and type of action ..... 304
Figure 8-9 SPADs immediate causes based on UK reports ..... 305
Figure 8-10 SPADs immediate causes based on UK initial reports post 2011 ..... 307

# LIST OF TABLES

Table 3-1 World railway traffic (International Union of Railways, 2011a) ..... 65
Table 3-2 Typical risk classification matrix ..... 67
Table 3-3 Summary of significant accidents and their consequences of the European State Members (European Railway Agency, 2012a) ..... 75
Table 3-4 Summary of significant accidents and their consequences from 19 European Railway Companies (International Union of Railways, 2012) ..... 75
Table 3-5 Summary of US train accidents and their consequences (Federal Railroad Administration, 2012) ..... 76
Table 3-6 Summary of Canadian main track train accidents and their consequences (Transportation Safety Board of Canada, 2012) ..... 78
Table 3-7 Summary of Australian train accidents and their consequences (Australian Transport Safety Bureau, 2012) ..... 79
Table 3-8 Type of accidents as classified per database ..... 80
Table 3-9 Proportion of the total fatalities and serious injuries per person type ..... 83
Table 4-1 Classification of error types according to cognitive stages ..... 95
Table 4-2 The human error classifications ..... 96
Table 4-3 Error causation models ..... 107
Table 4-4: A sample of performance shaping factors ..... 116
Table 4-5 Railway Human Performance Analysis Techniques ..... 119
Table 4-6 Main features of human performance analysis techniques ..... 128
Table 4-7 Comparison of Human Performance Analysis techniques ..... 131
Table 5-1 Review of railway contextual factors taxonomies ..... 152
Table 5-2 Factors influencing human actions in THERP (Kirwan, 1994) ..... 154
Table 5-3 The complete R-PSFs taxonomy ..... 162
Table 5-4 Data quality dimensions review (Dupuy, 2011) ..... 166
Table 5-5 Correlations between maturity score and recorded variables ..... 176
Table 6-1 Information extracted from investigation reports (refer to Appendix VI) ..... 181
Table 6-2 The most frequent identified R-PSFs from the first set of reports ..... 184
Table 6-3 Test for differences between different types of operators for R-PSFs rankings ..... 188
Table 6-4 R-PSFs assessment per scenario ..... 190
Table 6-5 Dependencies amongst R-PSFs ..... 194
Table 6-6 Descriptive statistics of reference database ..... 204
Table 6-7 The combined most frequent identified R-PSFs per type of event ..... 207
Table 6-8 Chi-square associations between R-PSFs lite and other variables ..... 210
Table 6-9 Chi-square summary table for R-PSFs lite and Type of events ..... 211

Table 6-10 Chi-square associations between Type of event and other variables ..... 216
Table 6-11 Examples of loglinear models for three-way contingency tables (Dupuy, 2011) ..... 219
Table 6-12 Pair-wise associations amongst variables ..... 219
Table 6-13 K-Way and Higher-Order Effects for three variables ..... 220
Table 6-14 Three-way associations amongst variables ..... 221
Table 6-15 Variables in the equation for binary logistic regression ..... 223
Table 6-16 Parameter estimates for ordinal regression analysis ..... 225
Table 6-17 The R-PSFs lite table to assist investigation reporting schemes ..... 228
Table 7-1 The Saaty's fundamental scale ..... 237
Table 7-2 Pairwise comparison matrix for Elements 1-3 ..... 238
Table 7-3 The Random Index ..... 238
Table 7-4 The generic structure of an *unweighted* supermatrix ..... 239
Table 7-5 The parent and children R-PSFs elements ..... 252
Table 7-6 Pairwise comparisons for the parent element *communication* ..... 253
Table 7-7 Pairwise comparisons for elements ..... 253
Table 7-8 R-PSFs priorities calculation ..... 254
Table 7-9 The *unweighted* supermatrix ..... 256
Table 7-10 Example of a pairwise comparison matrix for a category of factors ..... 257
Table 7-11 Pairwise comparisons for clusters ..... 257
Table 7-12 The cluster matrix ..... 258
Table 7-13 The weighted supermatrix ..... 260
Table 7-14 The limit supermatrix – final priorities ..... 262
Table 7-15 The rating of R-PSFs for the three type of actions (potential errors) ..... 264
Table 8-1 Characteristics of the sample ..... 271
Table 8-2 Characteristics of the sample after regrouping ..... 272
Table 8-3 Descriptive statistics for the R-PSFs *lite* categories weighting ..... 274
Table 8-4 Effect of R-PSFs categories on Dynamic Personal Factors for the SMEs subgroups ..... 275
Table 8-5 Effect of R-PSFs categories on Personal Factors for the SMEs subgroups ..... 276
Table 8-6 Effect of R-PSFs categories on Organisational Factors for the SMEs subgroups ..... 277
Table 8-7 Effect of R-PSFs categories on Team Factors for the SMEs subgroups ..... 278
Table 8-8 Effect of R-PSFs categories on Task Factors for the SMEs subgroups ..... 279
Table 8-9 Effect of R-PSFs categories on Task Factors for the SMEs subgroups ..... 280
Table 8-10 Descriptive statistics for the R-PSFs *lite* elements weighting ..... 282
Table 8-11 The final R-PSFs weightings means per classification group ..... 283
Table 8-12 Effect of classification groups on weighting distraction ..... 284
Table 8-13 Effect of classification groups on weighting fatigue ..... 284
Table 8-14 Effect of classification groups on weighting perception ..... 285
Table 8-15 Effect of classification groups on weighting of quality procedures ..... 286
Table 8-16 Effect of classification groups on weighting safety culture ..... 286
Table 8-17 Effect of classification groups on weighting supervision ..... 287

Table 8-18 Effect of classification groups on weighting familiarity ..... 288
Table 8-19 Effect of classification groups on weighting training ..... 288
Table 8-20 Effect of classification groups on weighting system design ..... 289
Table 8-21 Effect of classification groups on weighting workload ..... 290
Table 8-22 Effect of classification groups on weighting communication ..... 290
Table 8-23 Effect of classification groups on weighting quality of information ..... 290
Table 8-24 Effect of classification groups on the R-PSFs rating for “signal not detected” ..... 294
Table 8-25 Effect of classification groups on the R-PSFs rating for “signal wrongly interpreted” ..... 296
Table 8-26 Effect of classification groups on the R-PSFs rating for “action not executed correctly” . 298
Table 8-27 Friedman’s ANOVA statistics for the R-PSFs ratings ..... 300
Table 8-28 Comparison of HuPeROIs means between different SMEs groups ..... 302
Table 8-29 The HuPeROI results per type of SME and operational action ..... 302
Table 8-30 Comparison of HuPeROIs means between train drivers from different TOCs ..... 303
Table 8-31 The HuPeROI of signal detection for the classification group operational experience .... 304
Table 8-32 Change in performance under certain conditions ..... 308

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AHP | Analytic Hierarchy Process |
| ANP | Analytic Network Process |
| ATHEANA | A Technique for Human Event Analysis |
| ATM | Air Traffic Management |
| ATP | Automatic Train Protection |
| ATSB | Australian Transport Safety Board |
| AWS | Automatic Warning Systems |
| CARA | Controller Action Reliability Assessment |
| COCOM | Contextual Control Mode |
| CPCs | Common Performance Conditions |
| CREAM | Cognitive Reliability and Error Analysis Method |
| EC | European Commission |
| EPC | Error Producing Condition |
| ERA | European Railway Agency |
| ERADIS | European Railway Agency Database of Interoperability and Safety |
| ERTMS | European Railway Train Management System |
| FRA | Federal Railroad Administration |
| FRAM | Functional Resonance Analysis Method |
| GEMS | Generic Error Modelling System |
| GSM-R | Global System for Mobile Communication |
| HCR | Human Cognitive Reliability |
| HEP | Human Error Probability |
| HEPI | Human Error Probability Index |
| HEART | Human Error Assessment and Reduction Technique |

| | |
|---|---|
| HERA | Human Event Repository and Analysis |
| HERA-JANUS | Human Error in Air Traffic Management |
| HERMES | Human Error Risk Management for Engineering Systems |
| HFs | Human Factors |
| HFACS-RR | Human Factors Analysis and Classification System Railways |
| HMI | Human Machine Interface |
| HPA | Human Performance Analysis |
| HR | Human Reliability |
| HRA | Human Reliability Assessment |
| HRMS | Human Reliability Management System |
| HuPeROI | Human Performance Railway Operational Index |
| IM | Infrastructure Manager |
| JHEDI | Justification of Human Error Data Information |
| LUL | London Underground Limited |
| MAMD | Multi Attribute Decision Making |
| MCDM | Multi Criteria Decision Making |
| MODM | Multi Objective Decision Making |
| NIB | National Investigation Board |
| NTSB | National Transport Safety Board |
| PSF | Performance Shaping Factor |
| RAIB | Rail Accident Investigation Branch |
| RAIU | Railway Accident Investigation Unit |
| RARA | Railway Action Reliability Assessment |
| RCMs | Railway Communication Means |
| R-PSF | Railway Performance Shaping Factor |
| RSSB | Rail Safety and Standards Board |
| TOC | Train Operating Company |
| TRACEr-Rail | Technique for the Retrospective Analysis of Cognitive Errors-Rail |

| | |
|---|---|
| TPWS | Train Protection and Warning System |
| SBB | Swiss Federal Railways |
| SLIM | Success Likelihood Index Method |
| SME | Subject Matter Expert |
| SPAD | Signal Passed at Danger |
| SPAR-H | Standardized Plant Analysis Risk |
| SRK | Skill-Rule-Knowledge |
| STAMP | System Theoretic Accident Modelling and Processes |
| THERP | Technique for Human Error Rate and Prediction |

# Chapter 1 INTRODUCTION

This chapter gives an overall insight into this research by presenting the key role of human performance on the reliability and safety of railway operations. It presents the research problem and its context, and discusses its rationale, objectives, methodology and contributions. The chapter concludes with a summary of each chapter and a generic outline of the structure of this thesis.

## 1.1 Background to the problem

From ancient Greece to the present day, and in particular since the opening of the world's first passenger service in 1825, railways have played a crucial role in world history. Their significant contribution has been widely described in the literature (Berghaus, 1964, Nock, 1975, Ransom, 1981, Wolmar, 2007, Revill, 2012) and includes amongst others: trade proliferation, industrial development, transportation of passengers and goods and the formation of nations, countries and/or continents.

In order to play such a role, railways have radically evolved during their history and many technological achievements have been accomplished. Today's railways, which transport millions of passengers and millions of dollars' worth of goods worldwide every day (Dhillon, 2007), are faster, more powerful and more efficient than ever before. This has required relevant national government, operational and regulatory entities to develop rules, regulations and practices to achieve high performance in terms of efficiency, reliability, quality and safety (Wilson et al., 2007a). Furthermore, whilst railways and their operations have traditionally been the preserve of individual nations, in Europe there is currently a major shift towards the interoperability of railway services across borders that necessitates a very different approach to operations and the rules that govern them.

Railways constitute a complex system that consists of several components and stakeholders. The performance of this system in turn depends on several factors including rail traffic rules, infrastructure and rolling stock reliability, organisational management and human factors (Dhillon, 2007). To assess and improve its performance it is necessary to understand the components of the modern railway system, their dependencies and interactions. Furthermore, it is necessary to understand this for the different types of railway operation, e.g. main line, metros, in order for a railway to be effective.

One performance measure, safety, is of vital importance to operators and to the public. The complexity of the railway system means that accidents and incidents are inevitable. Statistics worldwide indicate a steady general improvement in railway safety in recent years, in terms of the reduced number of accidents and fatalities. Yet, it remains the case a train accident, such as that in Spain's Santiago de Compostela region in July 2013 with a major loss of life, invariably becomes a matter of grave national and international concern.

Statistics also show that the main cause of railway accidents worldwide is human error, which is broadly defined as "*any deviation from the performance of a specified or prescribed sequence of actions*" (Leveson, 2004, p.245). Such an error can involve either third parties, e.g. trespassers, or operational personnel. Although the former are responsible for the majority of accidents and fatalities, they are not within the scope of this research. This thesis concentrates on operational errors involving train drivers, signallers and controllers whom the literature (Dhillon, 2007, Evans, 2011) and statistics indicate that primarily affect the system in terms of safety.

This can be readily seen in some well know train accidents. Possibly the best known human error related railway accident in recent years is the catastrophic Ladbroke Grove train accident in the U.K. (Health and Safety Executive, 2000, Cullen, 2001), which involved a train driver. Initial investigations also appear to indicate that the accident in Santiago de Compostela was primarily due to human error. In addition, a recent study by Evans (2011) shows that at least 75% of the fatal railway accidents in Europe between 1990-2009 were due to operational human errors including exceeding speed limits, signal passed at danger and signalling/dispatching errors.

It would therefore seem that a proper investigation and analysis of operational human errors in the railways, with the subsequent development of mitigation measures, would make an invaluable contribution to the further improvement of railway safety.

However, to classify and subsequently examine human error poses considerable difficulties given numerous classification schemes. Furthermore, since human errors are related to degraded human performance, it is actually human performance that should be assessed. Human performance can be thought of as "*the human capabilities and limitations that have an impact on the safety and efficiency of operations*" (Maurino, 1998). The importance of human performance has been recognised by the U.S. Federal Railroad Administration (FRA) in their analysis of safety data which indicates that during the last decade, consistently more than 30% of the total rail accidents were related to human performance (Federal Railroad Administration, 2012). Therefore, the factors that affect human performance should be

identified and techniques that assess human performance should be reviewed and mitigation measures developed subsequent to this.

Several studies have indeed been conducted to investigate, understand and reduce human errors and subsequently incidents and accidents in railway operations (Wilson et al., 2007a). There has also been a significant increase in the interest in the area of human performance within railway operations (Priestley and Lee, 2008).

These studies are developed on the principles of Human Reliability Analysis building on relevant research from other domains including nuclear, oil and gas, and aviation (Bell and Holroyd, 2009). In this thesis HRA is described as "*the use of systems engineering and behavioral science methods in order to render a complete description of the human contribution to risk and to identify ways to reduce that risk"* (Swain and Guttmann, 1983, Boring, 2005). However, these studies are based on data and scenarios from the domains listed above, and therefore, they are not well suited to the rail industry and can be difficult to apply reliably to railway specific operations (Rail Safety and Standards Board, 2004b).

Furthermore, in analysing the factors that affect human performance in the railways, current methodologies typically adopt lists of factors from other domains, without appropriate definitions being given. In addition, even in the cases where the lists of factors have been modified to become more applicable to railways, such alterations have been designed based on regional accident and incident analysis (Baysari et al., 2009). Subsequently, although the number of factors that influence performance can be claimed to be limited, e.g. fatigue, training, organisational culture and system design, the analysis only of regional occurrences do not provide analysts with a worldwide perspective of the significance of factors on human performance.

On account of the above limitations, this thesis proposes a novel and comprehensive taxonomy for railway operations. This is the first railway performance shaping factors taxonomy developed through extensive literature review, operators' hierarchical task analysis, and analysis of accident and incident investigation reports internationally. The taxonomy is validated by subject matter experts (SMEs) to identify the factors that affect the performance of railway operators and assess human performance.

Building on the taxonomy this thesis also develops a comprehensive framework to quantify the impact of each of the factors that affect human performance. The framework accounts for the dependencies amongst the factors, while it can be also used to compare the relative likelihoods of different human operational errors.

## 1.2 Aim and objectives

Given the background above, the aim of this thesis is to develop a novel approach to assess human performance in order to enhance railway safety. The following objectives have been formulated to achieve this aim:

1. Specify the modern railway system, categorise and describe its components and finally define the operational concept of railway operations.

2. Review the literature on the operational concept of human errors and the models underpin them, including state-of-the-art Human Performance Analysis (HPA) techniques, identifying their strengths and limitations.

3. Determine the contextual factors that affect the performance of railway operators and develop a holistic framework that captures the railway performance shaping factors based on a novel taxonomy.

4. Develop a Human Performance Railway Operational Index (HuPeROI) to:
   (i) assess the influence of each identified contextual factor on the performance of railway operators, accounting for all the dependencies amongst the factors
   (ii) assess the performance of operators for several railway operational scenarios, by comparing the relative likelihoods of different human errors.

## 1.3 Outline of thesis

This thesis is organised into nine chapters. Chapter 2 captures the current railway system, including its main components and stakeholders. Furthermore, it presents the human aspect of the system by introducing the Railway Operational System Architecture (ROSA). The ROSA addresses the duties and responsibilities of railway operators, and identifies the interconnections between operators, signalling systems and communication systems.

Chapter 3 introduces the fundamentals of railway safety. In particular, it presents the historical developments in railway safety, including the analysis of a number of serious accidents fatalities and injuries on the worldwide railway network. The Chapter identifies the people are most exposed to danger in railway operations. Finally, it explores and determines the role of each component of the system to the safety of the network.

Chapter 4 provides the definitions and classifications of human error, and presents the existing methodologies to identify the causes of human error. It introduces human reliability

and human reliability analysis,and explains the link and relationship between the notions of human performance and human reliability. This is followed by an extensive review of several Human Performance Analysis (HPA) techniques and identification of the state-of-the-art in the area of HPA techniques in railway operations. Finally, the strengths and weaknesses of the techniques are described and the latter used to justify the need to develop an alternative approach to investigate human performance for the railway industry.

Having identified the weaknesses of the current HPA techniques, Chapter 5 introduces the Railway Performance Shaping Factors taxonomy (R-PSFs) and its underlying theory. The Chapter presents the approach to create, verify and validate the taxonomy. In particular, this chapter presents the findings from the literature, and introduces the R-PSFs taxonomy. In addition, it describes the different categories of the taxonomy, classifies each of the identified factors and provides precise definitions and examples for each of the factors. Finally, it discusses the characteristics of the required data for the consolidation of the taxonomy, while it also details the data quality criteria.

Based on the findings from the literature, Chapter 6 presents the consolidation of the taxonomy based on analysis of accident and incident reports and consultation with subject matter experts (SMEs). The interdependencies amongst the factors are defined and the most significant contributing factors, which constitute the "*lite*" R-PSFs taxonomy, are obtained. For these factors, any potential associations with the variables that characterise railway occurrences are determined, and the likelihood of different types of occurrences are estimated also.

Based on the R-PSFs taxonomy, Chapter 7 introduces the Human Performance Railway Operational Index (HuPeROI) and its underpinning methodology. This chapter introduces one of the leading multi criteria decision-making (MCDM) methodologies, the Analytic Network Process (ANP), and describes how it is used for the development of HUPeROI in conjunction with the Success Likelihood Index Methodology (SLIM). Based on the qualitative process of dependencies between the R-PSFs from Chapter 6, the impact of each R-PSF on human performance is assessed. For this, pairwise comparisons are performed between the factors that indicate dependencies. The Chapter concludes with the presentation of the pairwise comparison limit supermatrices, which are applied to define the contribution of the R-PSFs to human performance. To verify the methodology, Chapter 8 presents a study that was carried out in collaboration with three U.K Train Operators in order to evaluate three possible causes of a "*Signal Passed at Danger*" – the most typical railway operational error (Pasquini et al., 2004, Dhillon, 2007, Evans, 2011). The analyses of data are discussed and finally the indexes for the three causes are estimated.

The thesis ends with Chapter 9 reviewing the achievements of the thesis together with recommendations for further research. Figure 1-1 presents the overall structure of this thesis.

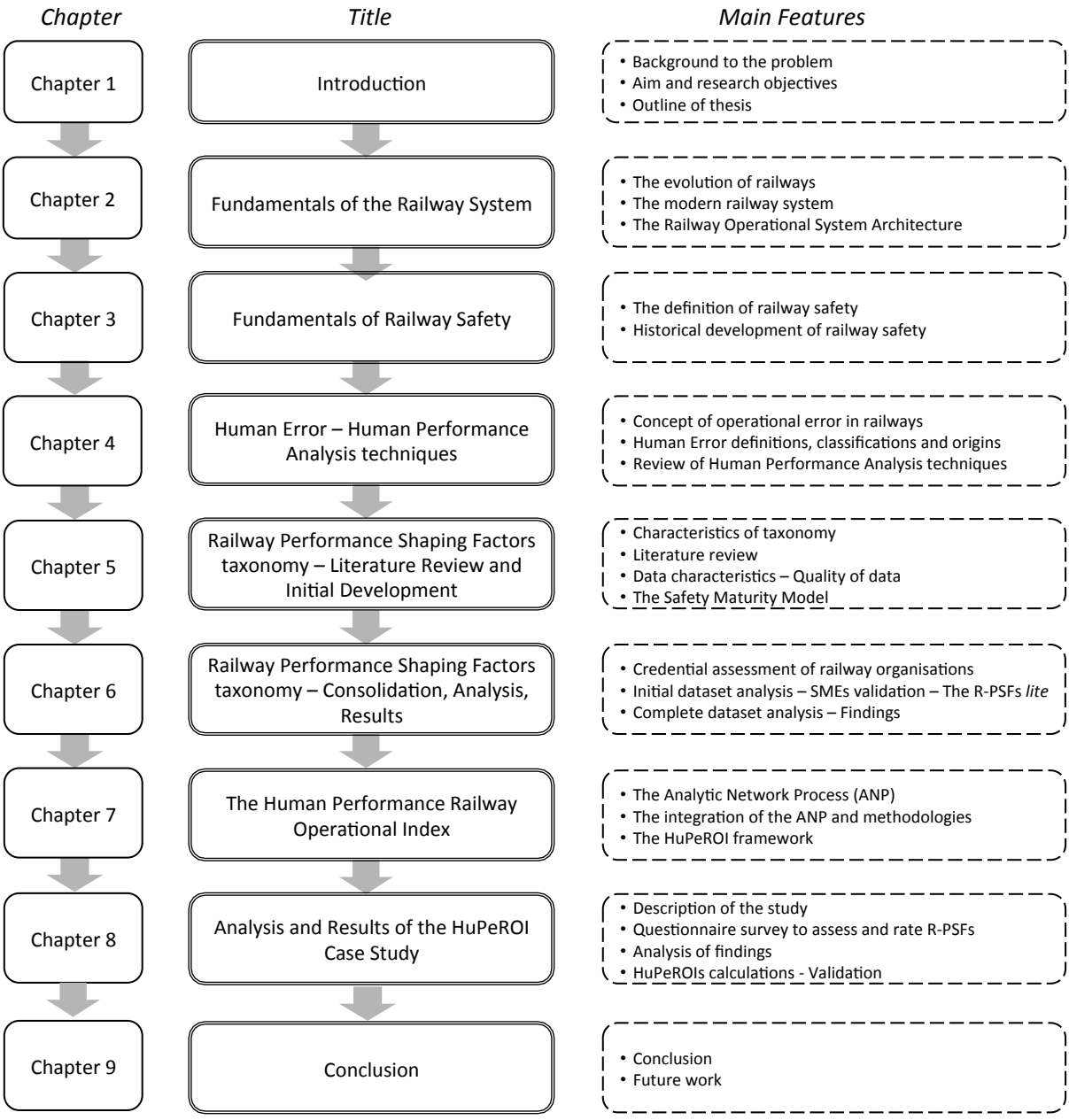



Figure 1-1 Thesis overview

# Chapter 2 FUNDAMENTALS OF THE RAILWAY SYSTEM

One of the principal objectives of this research presented in this thesis is to quantify and evaluate the factors that affect railway operators in the execution of their tasks. This requires a detailed understanding of the railway system, its components and operational framework. This chapter contributes to this in three ways. Firstly, it briefly presents the evolution of the railway system and highlights some of the most significant changes that have occurred over the years. Secondly, it captures the modern railway system, starting with the various definitions, identifies the differences and limitations of these definitions and subsequently proposes a new definition of the railway system. Based on this new definition, the chapter describes the constituents of the system in detail. Thirdly, the chapter introduces and discusses the Railway Operational System Architecture (ROSA), including its specific functional elements and their interactions. The ROSA is used in Chapter 5 to create the taxonomy of factors which influence the performance of railway operators.

## 2.1 Evolution of the railway system

For over 170 years modern railways have played a major role in the global history of transport. Their expansion has not only influenced the development of economies but also the structure of modern societies (Wolmar, 2007). Cities have been created or expanded, nations have been built and countries have come closer (Nock, 1975). Throughout the years many technological feats have been achieved. Faster trains have been designed, e.g. High Speed trains, additional safety countermeasures have been installed, e.g. the Train Protection Warning System (TPWS) and many infrastructure developments have been accomplished, e.g. the Channel tunnel.

The concept of railways can be traced to ancient Greece where heavy carts and two wheeled coaches were run in ruts built into stone paved roads for the transfer of goods (Berghaus, 1964, Wolmar, 2007). However, there was very little evidence of railways found from the period of the Roman Empire (Berghaus, 1964, Guy and Rees, 2011), leading historians to believe that the railways from that period were either destroyed or never found (Guy and Rees, 2011). Similar railways re-appeared in the medieval times and were clearly established from the 15$^{th}$ century (Wolmar, 2007). They were primarily used to move wagons with coal from deep shafts underground to the surface in English, German and Austrian mines (Wolmar, 2007). During this period and until the themed-18$^{th}$ century, the operational

concept of these primitive railways was relatively simple. Mineworkers either pushed or pulled the coal wagons along the tracks to the top of the mines. Then, the loaded wagons were rolled along towards their destination by gravity, while horses, almost the only form of traction on land, were used to haul them uphill when required (Nock, 1975).

The first significant change on railways came with the creation and use of the steam engine, as well as with the evolution of manufacturing techniques resulting in new materials and new types of rails: iron, wrought iron and finally steel rails (Bonnett, 2005). This evolution led to the rapid expansion of railways, and by the 1850's countries with an existing railway network were planning extensions, whereas countries with no railways were looking for means to introduce them (Nock, 1975). At the same time, the first safety regulations to maintain railway operations were introduced (Nock, 1975, Ransom, 1981, Revill, 2012), while considerations regarding the train braking system (automatic braking) was progressively implemented. However, the large expansion of railways led to capacity limitations. At the beginning of railway operations no signalling system was installed, but trains operated on a time interval basis and policemen secured their movements (Wolmar, 2007). This approach had several constrains regarding the flexibility and safety of the system. No trains could be added, rescheduled or delayed on the network, while their safe movement depended almost entirely on the performance of the policemen (Ransom, 1981). Later, with the invention of the telegraph and installation of fixed mechanical signals (semaphores) safety and capacity enhancements were achieved. However, the safety of railways was significantly improved only after the implementation of signal boxes and the interlocking system (Bonnett, 2005).

By the end of the 19$^{th}$ century railways had been transformed into a complex system which in its totality was greater than the sum of its individual parts (Revill, 2012). The system consisted of tracks, locomotives, wagons and coaches. It was operated by train drivers and signalmen and controlled by signalling and telegraph systems. It had extensive civil construction elements including bridges, viaducts and tunnels (Revill, 2012). At the beginning of the 20$^{th}$ century faster (more than 150 km/h) steam locomotives were introduced (Nock, 1975), while the railway network expanded, leading to a significant increase in passenger and freight services. Major improvements in the infrastructure were attained, e.g. new tunnels the Alpines (Nock, 1975, Ransom, 1981), whilst further safety considerations led to the installation of new measures such as the Automatic Train Control (ATC) system (Wolmar, 2007).

During the period between the First and Second World Wars (Nock, 1975), additional improvements on both rolling stock and infrastructure were accomplished, such as the introduction of the first diesel-electric locomotive in Argentina. Compared to steam

locomotives, the diesel-electric could not only run at higher speeds and accelerate faster but could also be operated by only one person. Furthermore, maintenance was faster and cheaper. In terms of infrastructure, the first colour-coded signals were installed, which could be identified more easily and clearly from a longer distance, especially during night operations (Nock, 1975). At the end of the Second World War and for the following fifteen years, railways were characterised by the rapid advance of the diesel-electric locomotive in addition to the successful use of Centralised Traffic Control (CTC), primarily in the USA. While steam locomotives were gradually replaced by the diesel-electric, new faster trains, high-speed trains, were introduced and used by many countries to compete against road and air transportation. Therefore, despite the strong competition of the automobile and aviation industry, the safer and faster railways remained a popular means of transportation and played a substantial role in the post-war reconstruction (Nock, 1975).

In the 1980's and for the following twenty-five years, essential technological and structural changes were accomplished worldwide. Railways became extremely fast, (with high-speed passenger trains travelling at speeds of up to 300 km/h), safer (e.g. the introduction of modern safety systems such as Train Protection Warning System), and more efficient (e.g. fewer delays). In addition, railway systems reduced their environmental impact (e.g. fewer emissions), whilst in some cases they became cheaper compared to other means of transport (Revill, 2012). In contrast, the privatisation of railways in many countries led to rather controversial results with respect to how beneficial it was for both the state and public (Wolmar, 2007).

Going forward, the future of railways, their efficiency and safe performance remains an issue of great concern to all railway stakeholders including train operators, infrastructure managers and maintenance contractors, regulators, suppliers and rolling stock operating companies (Revill, 2012). This has resulted in a number of initiatives including new concepts, and technologies such as, the construction of Maglev trains (Ryall, 2012) and the implementation of the European Rail Traffic Management System (ERTMS). New technologies, i.e. renewable energy resources, are being tested to improve efficiency and to maintain the environmental advantage of railways compared to the other means of transport. In addition, research is underway to improve infrastructure and rolling stock through new materials such as, carbon fibres. Driverless operations have been also investigated and partially implemented by introducing Automatic Train Operation (ATO), primarily in metro services including the metros in London (Railway Technical Web Pages, 2012a), Paris, Madrid, Copenhagen, Singapore and New York (Rumsey, 2009).

Since the introduction of the railway, operational personnel, especially train drivers and signallers have played the most critical role concerning the safety of railway operations (Dhillon, 2007). Therefore, despite the improvements in technology and safety systems, the performance of operators (i.e. human performance) continues to be monitored and analysed in order to improve safety. This is particularly, important in today's rail industry, which has the goal of 24/7 style operations (Dadashi et al., 2013). In this context, the impact of any new developments (technological or procedural, or both) on the performance of operators must be studied in depth in order to ensure the competence and capability of the operators to continue to perform their duties in a safe, secure and reliable manner. Therefore, amongst other studies, the factors that have an impact on the performance of operators must be clearly identified, defined and carefully examined with respect to their contribution to both past and potential railway accidents and incidents. This requires, in the first instance, a thorough understanding of the modern railway system.

## 2.2 The modern railway system

Over the years, although the main concept of operation of the railway system has largely remained the same, new systems have been implemented and entities established, significantly increasing the level of complexity (European Commission, 2004). This section defines the railway system and describes its elements.

### 2.2.1 Definition of the modern railway system

The European Commission (EC) (2004) defines the modern railway system as *"the totality of the subsystems for structural and operational areas, as defined in Directives 96/48/EC and 2001/16/EC, as well as the management and operation of the system as a whole*". For the conventional railway system, the structural subsystems are divided into infrastructure, rolling stock, energy, control-command and signalling, traffic operation and management, while the operational subsystems are separated into maintenance and telematics applications for both passenger and freight trains (European Commission, 2001a). On the other hand, for the high-speed railway system the two subsystems differ slightly. Here, the structural includes infrastructure, rolling stock, energy, control-command and signalling but not traffic operation and management, while the operational subsystem includes maintenance, operations, environment and users. The US Federal Railroad Administration (FRA) (2007), in contrast, considers that the railway system definition should be a narrative statement that *"fully describes, at a minimum, train operations, rolling stock, track configuration, signal systems,*

*infrastructure, and environment*, *while also matching or complementing the system definition included in the railroad's existing system safety program plan*". In addition, the Australian Transport Safety Board (ATSB) (2005) defines railways as "*a guided system designed for the movement of rolling stock, which has the capability of transporting passengers, freight or both on a railway track, together with its infrastructure, and associated sidings"*.

The above definitions describe the modern railways to some extent. However, there are a number of significant differences. For example, the EC definition distinguishes the signalling system from infrastructure, while the Australian definition provides no comparative information. On the contrary, although the FRA definition refers to environmental factors neither the EC nor the ATSB definitions do so. Finally, while the EC definition addresses maintenance, this is not the case for the FRA and ATSB, which primarily describe the system from its operational perspective. However, the main limitations of the above definitions are that only one accounts for the users of the railway system, and none addresses trespassers and the railway environs or neighbourhoods. Therefore, an alternative definition of the railway is proposed below based on Murtagh (2011).

From Murtagh's (2011) the railway system comprises five main components: infrastructure, trains (rolling stock), railway staff, third parties and neighbourhoods. The five components are further classified into technical and human, with the former including infrastructure and trains, while the latter includes railway staff, third parties and neighbourhoods. However, an additional and more useful classification is also introduced, dividing the five components into *controlled* and *uncontrolled*. The *controlled* components are described as "*the components that belong to a railway organisation*[1]*, hence can be controlled, designed and/or trained by the organisation*". These are infrastructure, trains and railway staff. Conversely, third parties and neighbourhoods, the *uncontrolled* components are described as "*the components that, although having an influence on the railway system and contributing to its safe operation, do not belong directly to the railway organisation and are therefore difficult to be controlled, ruled and maintained by the organisation*". Both the *controlled* and *uncontrolled* components are further divided in sub-components and elements, as shown in Figure 2-1. In addition to its *controlled* and *uncontrolled* components, the railway system incorporates a number of stakeholders who have a reasonable interest in safe and reliable railways. The stakeholders have either a direct relation to the railway operational scheme, e.g. train operators, infrastructure managers, or an indirect one, e.g. regulation and safety bodies, rolling stock

[1] Railway organisation is described as any private or public enterprise acting as a railway transport operator, infrastructure manager or integrated company Eurostat et al. (2009).

suppliers and infrastructure maintenance contractors. Finally, to obtain the overall picture of the railway system, the operational and environmental conditions, as well as implemented standards and regulations are taken into account.

Based on the system's description, the dependencies amongst its elements and its components can be identified. These are established either between or within components. For example, within infrastructure the type of track is linked to whether that track is electrified or not. Similarly, with regard to trains, the length of a train is related to its speed limits. On the other hand, train speed is dependent on the track's curve radius.

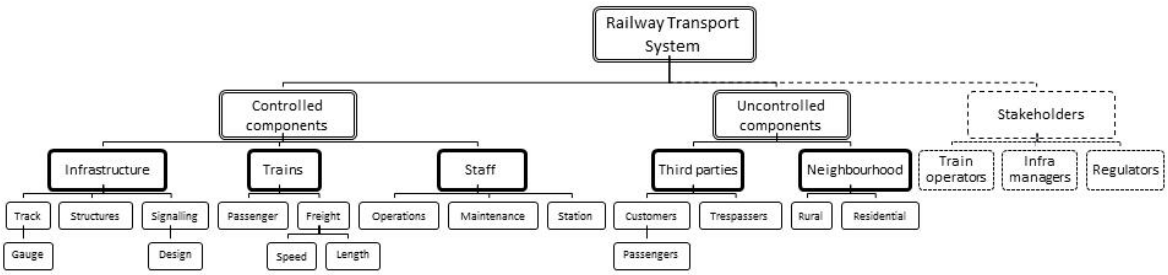



Figure 2-1 Railway transport system

The main components and stakeholders of the railway system are presented in the next subsections.

## 2.2.2 The railway *controlled* components

### 2.2.2.1 Infrastructure

Infrastructure is generally described as "*the system of facilities, equipment and services needed for the operation of an organization*" (International Organization for Standardization, 2005). The EC (2001a) defines railway infrastructure as "*the track, points, engineering structures (bridges, tunnels, etc.), associated station infrastructure (platforms, zones of access, etc.), safety and protective equipment"* necessary for the operation of the system, based on the definitions given in (European Commission, 1996, European Commission, 2001a). The FRA (2007), on the other hand, refers to railway infrastructure as "*the bridges, tunnels, stations, industrial sidings or sites and other fixed objects or facilities along right-of-way*", while the ATSB (2005) as "*the permanent way/area on or about the track*". In order to adapt a more comprehensive definition for railway infrastructure and to bridge the different definitions presented above, railway infrastructure is defined in this thesis as "*the system of*

*facilities, equipment and services necessary for the operation of railways, including tracks, signalling, command and control systems, electrification, premises and structures*".

#### 2.2.2.1.1 Track

A railway track is defined as "*a pair of rails over which the railway vehicles can run*" (Eurostat et al., 2009). Apart from rails, the main structure of tracks comprises sleepers, ballast, fastenings, sub-ballast and subgrade, as shown in Figure 2-2. Depending on the type of traction, track can be either electrified or not, as described in Section 2.2.2.1.3.

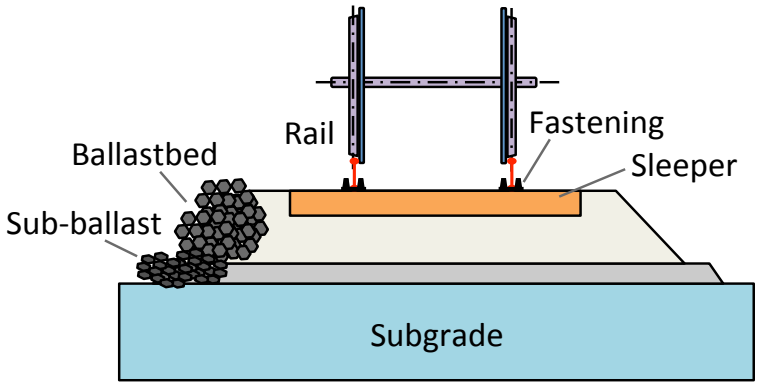



Figure 2-2 Principle of track structure (Esveld, 2001)

The sleepers are the units of tracks that are used to (i) support rails by holding them upright, (ii) maintain the correct distance between rails, (iii) spread wheel loads to the ballast and provide the basis for rail fastenings, (iv) transmit the lateral and longitudinal train forces and (v) isolate rails electrically (Bonnett, 2005). The distance between rails is called gauge and is "*measured between the inside edges of the railheads"* (Eurostat et al., 2009). Today, most railway tracks are at or about 1432 mm gauge (standard gauge) (Bonnett, 2005). Sleepers lay upon the ballast, which is the set of stones/natural rocks used to provide vertical and lateral stability to the track and to drain water from the bottom of sleepers (Bonnett, 2005). In addition, ballast and sub-ballast layers are used to disperse track load in such a way that the subgrade layer can resist the imposed pressure adequately (Bonnett, 2005). Fastenings, on the other hand, are used to connect rails together and the track with the sleepers. Sub-grade, finally, is the term used to describe the soil stratum or embankment soil upon which the track bed is constructed (Bonnett, 2005).

In order to be operational, railways should allow trains to change and be diverted from one track to another. Therefore, along the tracks there are switches (points) and crossings layouts to guide trains to the correct direction (De Fontgalland, 1984, Bonnett, 2005, Eurostat et al., 2009). Switches, the moving part of the layout, guide train wheels to the

correct direction, while crossings, the stable part, allow trains to pass in either direction once the switch has been set, as illustrated in Figure 2-3. Different types of switches and crossings are used depending, amongst others, on the number of lines involved, frequency of use and running line speed. For example, high-speed trains need long switches and crossing layouts, while at stations trains can make tighter turns due to their lower speeds.

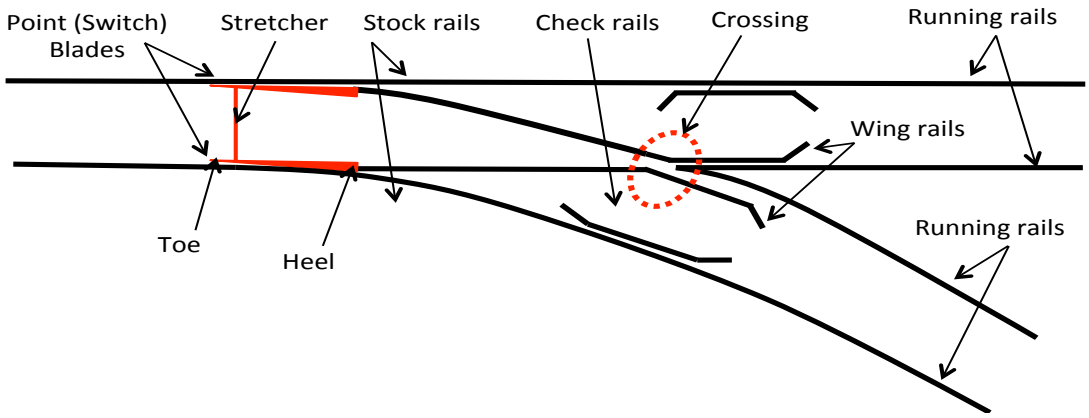



Figure 2-3 Points and Crossings description (Railway Technical Web Pages, 2012b)

#### 2.2.2.1.2 Signalling, command and train control

The importance of signalling and controlling train movements was recognised from the early days of railways. In principal, such systems aim at providing train drivers sufficient time to stop trains before meeting any obstacles ahead, with minimal reduction of line usage (De Fontgalland, 1984). Due to the individual train characteristics, e.g. length, weight and adhesion, braking cannot be immediate. Thus signalling is essential for informing drivers well in advance of the actions that they should take, e.g. reducing train speed or proceeding at normal speed. According to Bonnet (2005) although modern signalling systems may differ not only between but also within countries, they should all have the following objectives: (i) control trains in a safe manner for the conditions ahead; (ii) maintain a safe distance to any train in front or behind; (iii) prevent the setting of conflicting movements; (iv) ensure that switches are locked in the correct position; (v) enable trains to operate at the scheduled speed with a minimum of possible disruptions, without affecting the safety of operations; and (vi) assure that trains always maintain the minimum required time interval between them.

The operational concept of the modern railway signalling is developed on the "*block system*" concept (Pasquini et al., 2004). A route line is divided into blocks, where each is protected by a signal. The operational principle can be summed up to "*no more than one train shall be in a block section on the same line at the same time*" (Hall, 2005). If a Block n is occupied then the signal protecting that block should be set to red, while the preceding signal should

be yellow, as illustrated in Figure 2-4. The yellow signal alerts the driver that the signal ahead is likely to be red, unless the block is cleared during the time required for the train to move between the two signals. For normal running movements different types of signals are used, including, amongst others, the "*two*", "*three*" and "*four*" aspect colour signals. For other specific movements, i.e. shunting operations, small signals are used (ground or dwarf signals). Signals can be found at the start of a block, on the approach to switches, before other signals and level crossings, at stations or ahead of junctions. To ensure that trains have sufficient time to stop, the distance between signals depends on the line speed, gradient of the track and a safety margin specified by the designers (De Fontgalland, 1984, Bonnett, 2005).

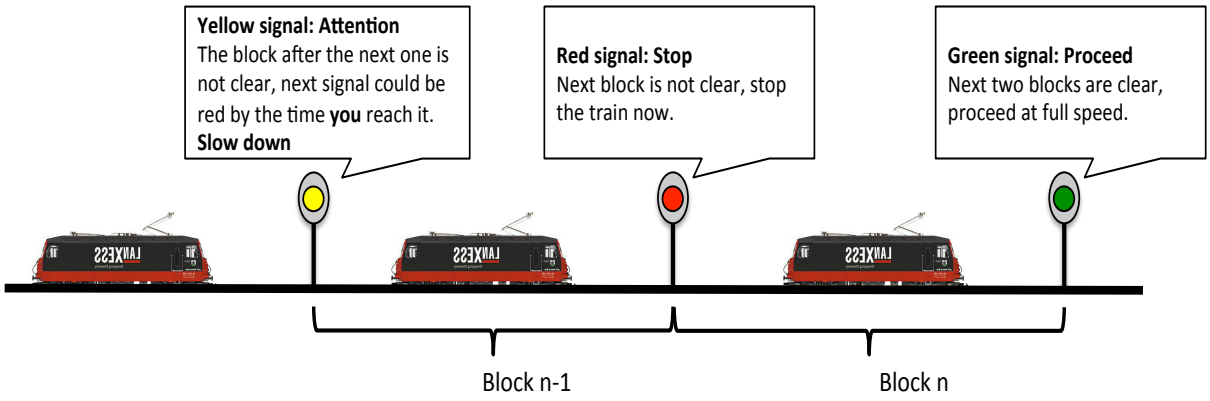



Figure 2-4 Signalling blocks protection (Pasquini et al., 2004)

In addition to signalling, there are many protection systems implemented worldwide, which all primarily aim to warn train drivers of signals and their indications ahead. In addition, many of them are able to stop trains automatically when drivers do not obey signals indications. Some of the best-known protection systems, as described by Hall (2005), are :

- The Automatic Warning System (AWS). This is widely used in the UK and Australia primarily on main line operations for both passengers and freight trains. AWS is designed to warn train drivers of the need to slow down or stop their trains.
- The Train Protection and Warning System (TPWS). Widely used in the UK and Australia, TPWS has been developed as an improvement of the AWS. It is used to reduce the consequences of Signals Passed at Danger (SPADs). The TPWS is designed to capture the probability of a driver acknowledging the AWS warning indicators but failing to apply braking appropriately. Subsequently, if the train passes the signal at danger, the TPWS will force the train to stop regardless of the actions of the train driver.
- The Automatic Train Protection (ATP). ATP is widely applied in the US, UK and Canadian railway operations to ensure that trains operate within the required speed

limits. ATP differs from AWS and TPWS, as it applies braking automatically if it finds that train drivers fail to reduce train speed to the extent required. ATP is very popular on metro services due to the very dense train services.

- The Integra-Signum protection system. This is used in Switzerland and is similar to the TPWS in stopping the train automatically when train drivers either do not acknowledge signals ahead or pass signals at danger.
- The European Train Control System (ETCS). ECTS has been developed to introduce a single train protection system on the European network in order to simplify cross-country and within countries operations. There are three ECTS levels, Level 1, 2 and 3. Depending on the Level, train controls rely on: the conventional line side signals and ATP system (Level 1); the conventional detection systems and information that drivers receive by a secure radio system, referred to as the Global System for Mobile communication-Railway GSM-R (Level 2); on-board calculations of train location and transmission to control centres by GSM-R (Level 3).

With respect to the concept of controlling the train movements, this lies on the "*track circuit*" idea, which is considered as the heart of the modern signalling (De Fontgalland, 1984, Hall, 2005). It is the simplest and probably most efficient method to identify if a train is operating on a particular length of a track and to set up the signals and switches accordingly (Bonnett, 2005). The passage of the wheels over a section activates an electrical circuit between the rails which subsequently turns on the relays and causes the signal and switches to change (De Fontgalland, 1984, Hall, 2005). The track circuit is responsible for a number of essential functions, such as: holding signals in the rear at danger; locking the facing points in the route direction; notifying signallers of the presence of a train; changing signals ahead from danger to proceed aspect; supporting signallers to keep in step with train movements; and finally giving warnings of broken rails or obstacles on the line (Hall, 2005). Finally, in addition to track circuit, the interlocking of signals and switches is another way to enhance safety of railway operations. Interlocking has been used to prevent signallers accidentally clearing a signal before the points are properly set and to clear signals that could lead to a conflicting movement (Bonnett, 2005).

In addition to the signalling and control of trains, the railway communication systems (RCS) are vital for a safe and secure railway operation. Similar to the railway protection systems, the RCS may differ across countries. However, they all aim to provide and maintain secure communication between trains and control rooms along the rail network, as well as provide passengers with essential journey information. Radio is the primary means of communication, while in some cases the use of telephones is permitted (Bonnett, 2005). In

most European countries and others including India, China and Saudi Arabia, national telecommunication railway systems, e.g. the Cab Secure Radio (CSR) and the National Radio Network (NRN) in the UK, are gradually being replaced by GSM-R (Smith et al., 2013). This improves safety and facilitates the interoperability of the railway operations.

#### 2.2.2.1.3 Electrification

Electrification provides trains with electrical energy, which is transferred either by overhead or earthing equipment. The scheme is characterised by its current and voltage. Current can be either direct (DC) or alternating (AC) with a broad variety of voltage (Bonnett, 2005, Eurostat et al., 2009). The main difference between AC and DC systems is the way of transforming and rectifying energy (Bonnett, 2005). While in the case of the DC this takes place at fixed sub stations distributed along the railway lines, in the AC system this occurs in a mobile on board substation (Bonnett, 2005). Most metros, trams and trolleys use the low voltage DC system, while the high voltage AC system is used on the main lines.

Electrification has several advantages compared to diesel traction including: fuel efficiency; reduced environmental pollution; faster acceleration and braking, which is particularly important in the cases of rapid transit; almost unlimited power supply; and lower cost locomotives (Bonnett, 2005, Hoffrichter, 2009). On the other hand, it is characterised by: the necessary high capital cost for its installation; its vulnerability especially in the cases of overhead equipment; and finally its higher maintenance cost.

#### 2.2.2.1.4 Premises and structures

Structures and premises are used to facilitate railway operations. Train stations, level crossings, depots, tunnels, bridges and viaducts are included in this category.

Train stations, according to Eurostat (2009), is a railway premise "*either open or not to the public, usually staffed and designed for one or more operations"*. Such operations include: formation, dispatch, reception and stabling of trains; boarding or disembarking of passengers; loading and unloading of goods; and when open to public, ticketing and information facilities.

Level crossings (or high-way grade crossings for US railways) are considered as "*any level intersection between the railway and a passage, as recognised by the infrastructure manager and open to public or private users. Passages between platforms within stations are excluded, as well as passages over tracks for the sole use of employees*" (European

Railway Agency, 2013a). Level crossings constitute an essential part of the railway system. First, they facilitate the movement of both road vehicles and pedestrian traffic across the railway (Rail Safety and Standards Board, 2004c). Second, as stated in Chapter 3, they significantly affect the safety of railway operations.

Depots are railway premises where trains are stabled, train units are formed and maintenance works are performed. Depots are usually located close to larger train stations. Depending on their functions, they should address the following requirements (Bonnett, 2005): adequate sidings to stable the maximum number of trains that could be in the depot; adequate points and crossings to enable trains movements within the depot; sufficient staff facilities, office, training rooms; adequate equipment for train maintenance work; and safety protection measures, e.g. fire alarms.

Finally, the tunnels, bridges and viaducts contribute significantly to railway operations, e.g. the undersea Channel Tunnel, which connects the UK with continental Europe and the Forth Bridge in Scotland. Such constructions simplify train operations by reducing travelling time.

#### 2.2.2.2 Trains

Trains are described as "*a convoy of railway vehicles moved by a traction railway vehicle or integral motor*, *or a traction railway vehicle travelling alone"* (De Fontgalland, 1984, Eurostat et al., 2009). They are classified based on the type of use primarily into passenger and freight trains. A wide range of rolling stock (all vehicles run on railways) is used nowadays, including: locomotives, passenger carriages and freight wagons, multiple units, metro cars, light rails and trams and rail machines (Bonnett, 2005).

Locomotives are "*traction vehicles which are not intended to carry a payload and have the ability to uncouple in normal operation from a train and operate independently*. There are two main types of locomotives used worldwide (European Railway Agency, 2009):

- The electric, which derives energy from electrified tracks or on board accumulators.
- The diesel, where a diesel engine is the main power source.

Passenger carriages are non-traction railway vehicles which are "*used for the conveyance of passengers and they include saloon cars, sleeping cars, dining and ambulance cars"* (European Railway Agency, 2009, Eurostat et al., 2009). Conversely, freight wagons are "*railway vehicles normally intended for the transport of goods*" (Eurostat et al., 2009). Finally, multiple units are "*tractive railway vehicles constructed for the conveyance of passengers or goods"* (Eurostat et al., 2009). They are distinguished in electric (EMU) and diesel (DMU)

multiple units according to the type of the received power. Metros, another type of passenger trains, are electric trains operated primarily on underground lines, while light trains and trams are also passenger trains used for over ground urban transportation. Rail machines, finally, include types of trains like cranes or tampers, which are used on specific occasions, e.g. to carry ballast.

Trains are characterised based on their technical specifications, such as their axle load, braking system, traction mode, length, speed or size of the wheelbase, which depend on the infrastructure design they operate on.

#### 2.2.2.3 Personnel

The railway industry employs a large number of staff, broadly classified in two categories: administration and operational. The former refers to "*general administration employees, including central and regional management staff (e.g. finance, legal, personnel) as well as the board of directors*" (Eurostat et al., 2009). The operational staff, on the other hand, consists of the frontline personnel, and regional and headquarters staff (Murtagh, 2011). Frontline personnel include train drivers, signallers and controllers, train crew, station staff and maintenance personnel. The central and headquarters personnel, depending on the size of the railway organisation, comprise safety managers, driver train managers, operation competence managers, operation risk managers and the board of directors (Eurostat et al., 2009).

##### 2.2.2.3.1 Train driver

EC (2007) describes train drivers as "*any person capable and authorised to drive trains, including locomotives, shunting locomotives, work trains, maintenance railway vehicles or trains for the carriage of passengers or goods by rail in an autonomous, responsible and safe manner*". Train drivers perform several tasks during their shifts. At the start of each shift, before a train departs, drivers are responsible for conducting all the necessary safety checks, e.g. braking. After a train's departure, they operate the train on the set route, respond to signals, obey speed restrictions and comply with a common set of railway rules and regulations, which are provided by the rail company or the regulators (Rail Safety and Standards Board, 2009). Additionally, drivers have to stop at the stations on the route and keep to the timetable as far as possible. They also stay in contact with control rooms, station staff and on-train colleagues to report any problems on the line or train. A more detailed description of a train driver's duties is provided in Section 2.3.1.

#### 2.2.2.3.2 Signaller

Railway signallers are the operators responsible for the control of the movement of trains on railway networks. Their main duty is to ensure that a train progresses from a point A to B safely and on time. In addition, signallers are in charge of setting the tracks on which trains run and of changing, when required, the signals to inform train drivers either to proceed or stop the operation.

Signallers are responsible for a large number of actions over the period of their shift (Network Rail, 2008b). At the beginning of their shift they study and acknowledge the daily operational schedule. They must also be aware of all train movements, timetable changes, speed restrictions or faults. In addition, they are responsible for planning trains routes and communicating with their co-workers in other control rooms or signal boxes. For example, giving clearance to the previous signal box for trains to enter a section, while waiting for authority from the next signal box before allowing a train to proceed (Roth and Malsch, 1999, Network Rail, 2008b). Signallers are also in charge of sending out appropriate signals and set all the points on the train's route as required. They also warn the proceeding signal boxes that a train is approaching, stay in radio contact with train drivers and maintenance workers, and record details of all train movements (Roth and Malsch, 1999, Network Rail, 2008b). Finally, they have to coordinate access to the railway premises for technicians and engineers who need to carry out maintenance works, to deal with and to resolve any incidents affecting the safe operation of the railway (Garner et al., 2007, Network Rail, 2008b).

#### 2.2.2.3.3 Controller

In addition to signallers, the larger control centres are staffed with controllers who are responsible for the provision of real time integrated management of train operations. This includes the coordination of the supply of information on significant events to the directors of railway organisations (Network Rail, 2008a).

Controllers compile and maintain the daily national operations centre log and review significant incidents for industry-wide transmission. They also update a network's availability display board in order to enable the visual illustration of the fitness of the network. Controllers also review train services, particularly during times of national disruption and liaise with both internal and external controls (Network Rail, 2008a). Additionally, they identify significant incidents, determine the incidents that require statutory reporting and monitor the management of all incidents. They provide leadership and direction to route

control staff, and manage the flow of the necessary information between the frontline personnel of the organisation (Network Rail, 2008a). Finally, controllers lead the daily joint national route control and train operating companies briefing, disseminate the daily weather forecast and identify potential high risk situations (Network Rail, 2008a).

#### 2.2.2.3.4 Conductor

Train conductors are the supervisors of a train crew, responsible for the safety of the non-operational duties of the train (Eurostat et al., 2009). In principal, conductors are responsible for ensuring that trains operate according to the time schedule and that any passengers and cargo are embarked and disembarked safely. Train conductors are also in charge of completing the en-route paperwork and ensuring that trains operate on the relevant safety rules and practices. Finally, they are in control of: train movements when operating in reverse; coupling or uncoupling coaches; assisting with setting out or picking up additional rolling stock; carrying out running repairs and collecting tickets from customers.

#### 2.2.2.3.5 Station staff

The station staff refers to any employee who is allowed to work in the station, e.g. ticketing personnel. Station staff report to the station manager who is the responsible for the day-to-day management of the station (European Commission, 2008).

#### 2.2.2.3.6 Maintenance personnel

The maintenance personnel can be described as any qualified worker who maintains, repairs or renews railway property or constructs new work in relation to the railway property. Maintenance personnel could be either internal, e.g. belong to infrastructure manager or external, belong to certified maintenance contractors.

### 2.2.3 The railway *uncontrolled* components

#### 2.2.3.1 Third parties

Third parties are sub-divided into the customers of the system and trespassers, with the former comprising passengers and freight customers. In addition, third parties also include the people who interact with the railway system at level crossings.

#### 2.2.3.1.1 Customers - Passengers

A railway passenger is defined as "*any person who makes a journey by rail, excluding the members of the train crew*" (Eurostat et al., 2009). Freight customers, on the other hand, are people who use freight railway operations to transfer goods.

#### 2.2.3.1.2 Trespassers

The railway trespassers are broadly described by (Network Rail, 2012b), as those people who "*without permission go on to the railway tracks, the embankments or any other restricted area – for any reason at any time*". Trespassers could be suicidal or people who, for example, take a shortcut across the train tracks or jump off the platform to pick up something that was accidentally dropped.

#### 2.2.3.1.3 Individuals at Level Crossings

Individuals at level crossings refer to both road vehicles occupants and pedestrians who attempt to cross from one side of the level crossing to the other.

#### 2.2.3.1.4 Neighbourhoods

Railways are located and operated within three main types of neighbourhoods: residential, industrial and rural (Murtagh, 2011). Residential areas are described by Murtagh (2011) as "*any districts primarily occupied by private residences*", while the industrial areas are defined as "*any districts occupied by heavy industry buildings and premises*". Finally, rural neighbourhoods are areas "*with relatively low population density and premises, such as the agricultural areas and forests".*

As shown in Chapter 3, the type of railway surroundings may have a significant impact on railway operations in terms of safety.

### 2.2.4 The main railways stakeholders

Railway stakeholders are all persons and organisations with a reasonable interest in train service delivery, including train operators, infrastructure managers, regulators and safety bodies, and finally other suppliers.

#### 2.2.4.1 Train operators

Train operators are described as "*any certified public or private enterprise, which provides railway transportation for passengers and/or goods*" (Eurostat et al., 2009).

#### 2.2.4.2 Infrastructure managers

EC (2004) identifies an infrastructure manager as "*anybody that is responsible in particular for establishing and maintaining railway infrastructure as well as for operating the control and safety systems*".

In the case that an infrastructure manager is also a train operator, e.g. the Swiss Federal Railways (SBB), then this body is called an "*integrated company*" (European Railway Agency, 2009).

#### 2.2.4.3 Infrastructure maintenance contractors

The infrastructure maintenance contractors constitute certified enterprises, which conduct specific maintenance works assigned to them by the infrastructure manager.

#### 2.2.4.4 Rolling Stock Companies

Rolling stock companies (ROSCOs) are the "*owners of the trains, which however, are not eligible to operate on the railway network*". Therefore, they provide train operating companies with trains to cover future development.

#### 2.2.4.5 Regulators and safety bodies

EC (2001b) requires each country to have a regulatory body "*which can be the Ministry responsible for transport matters or any other body, shall be independent in its organisation, funding decisions, legal structure and decision-making from any infrastructure manager, charging body, allocation body or applicant*. *The main task of the Regulatory Body should be to ensure a fair and non-discriminatory access to the rail network and services*".

In addition to the regulators, states across the world have established relevant independent organisations, usually referred to as safety bodies (boards), which are responsible for the investigation of all relevant accidents and incidents.

#### 2.2.4.6 Suppliers

Suppliers are described as "*any certified entity other than train operator or infrastructure manager who manufactures and supplies products to the railway industry*" (European Railway Agency, 2009).

## 2.3 The Railway Operational System Architecture framework

The preceding section has presented in detail the modern railway system including its basic components and elements. Based on this description, this section discusses the operational concept of the system by introducing its functional architecture, which is described as "*a high level decomposition of the overall system into a logical set of functional blocks*" (Subotic, 2007). In addition, the dependencies between the constituents of the operational scheme are briefly illustrated. The functional architecture, referred to as the Railway Operational System Architecture (ROSA), is an integrated system of people, infrastructure, rolling stock, actions and information, developed in an efficient way to achieve the common objective of safe and undisturbed movement of a train from an origin to a destination. The ROSA has been developed from the results of: (i) a literature review; (ii) on-site observations in train driver cabins of mainline and underground train operators and finally (iii) on-site visits to a number of railway control centres.

The ROSA was primarily generated based on the characteristics of the UK main line railway operations. However, it can be transferred to other countries, type of operations and/or services with appropriate modifications. For example, long haul services in the US and Australia operate with two train drivers on-board, while the existence of controllers in control centres worldwide depends on the relevant size of the designated area of responsibility. Signalling, controlling and safety protection systems may also vary. Again, in the case of the US railways there is no TPWS, while on some metro operations, e.g. London, although the AWS is not implemented, the ATP is installed.

### 2.3.1 The ROSA - Literature review

From the literature (Rail Safety and Standards Board, 2009) the architecture of a driver advisory system, as shown in Figure 2-5, was first derived.

It can be seen that the architecture consists of three main components: control centre, train and signalling system. Various forms of data are transferred via and between the components, to informing operators of the status of railway operations. For example, train

drivers continuously receive information on the route timetable and any required adjustments e.g. accelerate or slow down, in order to achieve the scheduled service. Moreover, signallers receive information on their monitors from the signalling and controlling systems on train location and the conditions of the technical equipment. This enables signallers to identify and resolve any potential conflicts.

In addition to the architecture in Figure 2-5, a hierarchical task analysis (HTA) was conducted to define the tasks that operators should perform during their shifts. The HTA was used in order to decompose the complex tasks of operators into a hierarchy of high and low level operations. Such an analysis aimed at identifying the tasks that are more likely to fail due to, for example, poor design, equipment failure and lack of competence (Stanton et al., 2005a).

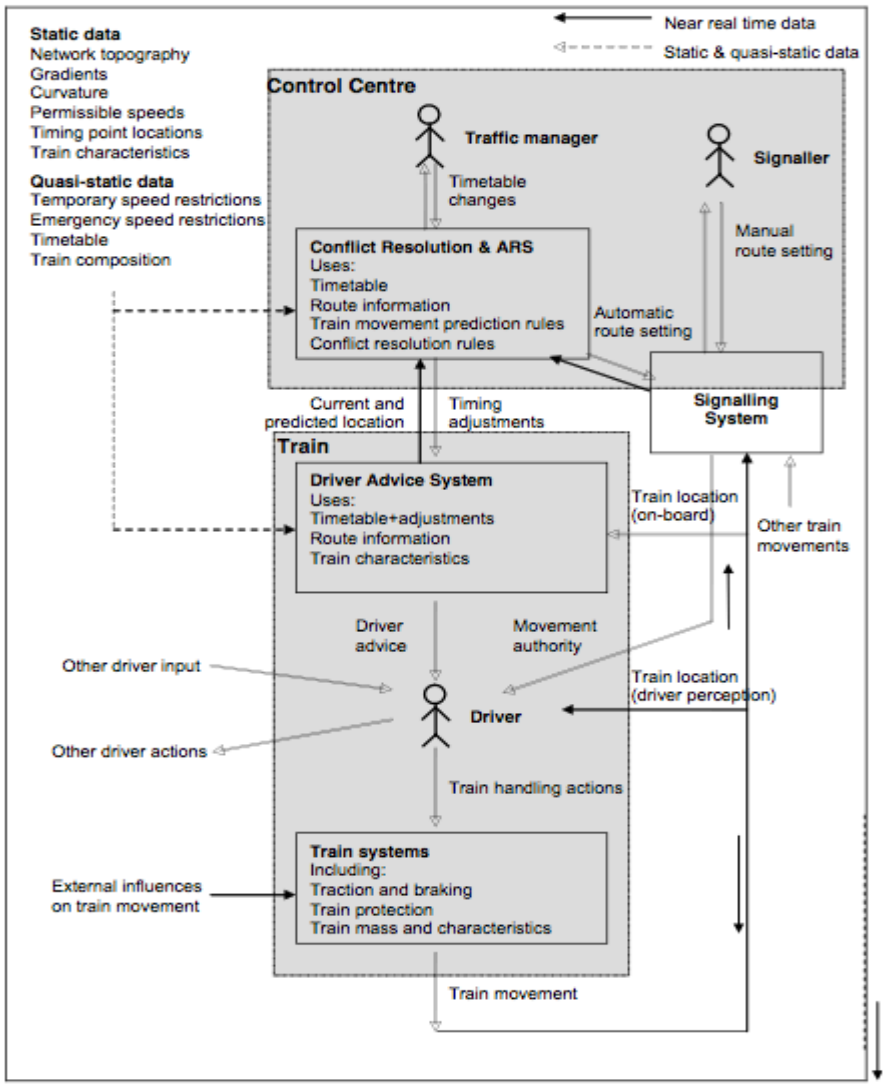



Figure 2-5 Driver advisory system architecture (Rail Safety and Standards Board, 2009)

There are three main phases of the services provided by train drivers: pre-operational, operational and post-operational. The pre-operational phase includes the tasks associated with the preparation of a train for service. The operational phase comprises the main sets of activities of train driving, including performance of activities related to the type of service and response to potential failures during the service. The post operational phase consists of the tasks required at the conclusion of a service. Figure 2-6 provides a high level HTA for all the three phases. In addition, Figure 2-7 presents the activities involved in train driving in the UK. These include: (i) departing from a scheduled stop; (ii) driving towards a scheduled stop; (iii) stopping at the next scheduled stop. All the activities are characterised by two essential goals: maintaining safety and schedule of service. The three activities are further broken down into sub-activities, which in most cases encompass different tasks requiring elaboration. For example, in order to define the driving strategy, train drivers collect and recall from route knowledge the current and future speed targets that apply to the service (Rail Safety and Standards Board, 2009). In addition, they constantly monitor the train speed by collecting information from the speedometer and various sources (e.g. cab noise, and motion). They compare the suggested speed with the actual train speed and subsequently, when the train speed does not match with the required speed, they modify the power or brake control settings accordingly (Rail Safety and Standards Board, 2009).

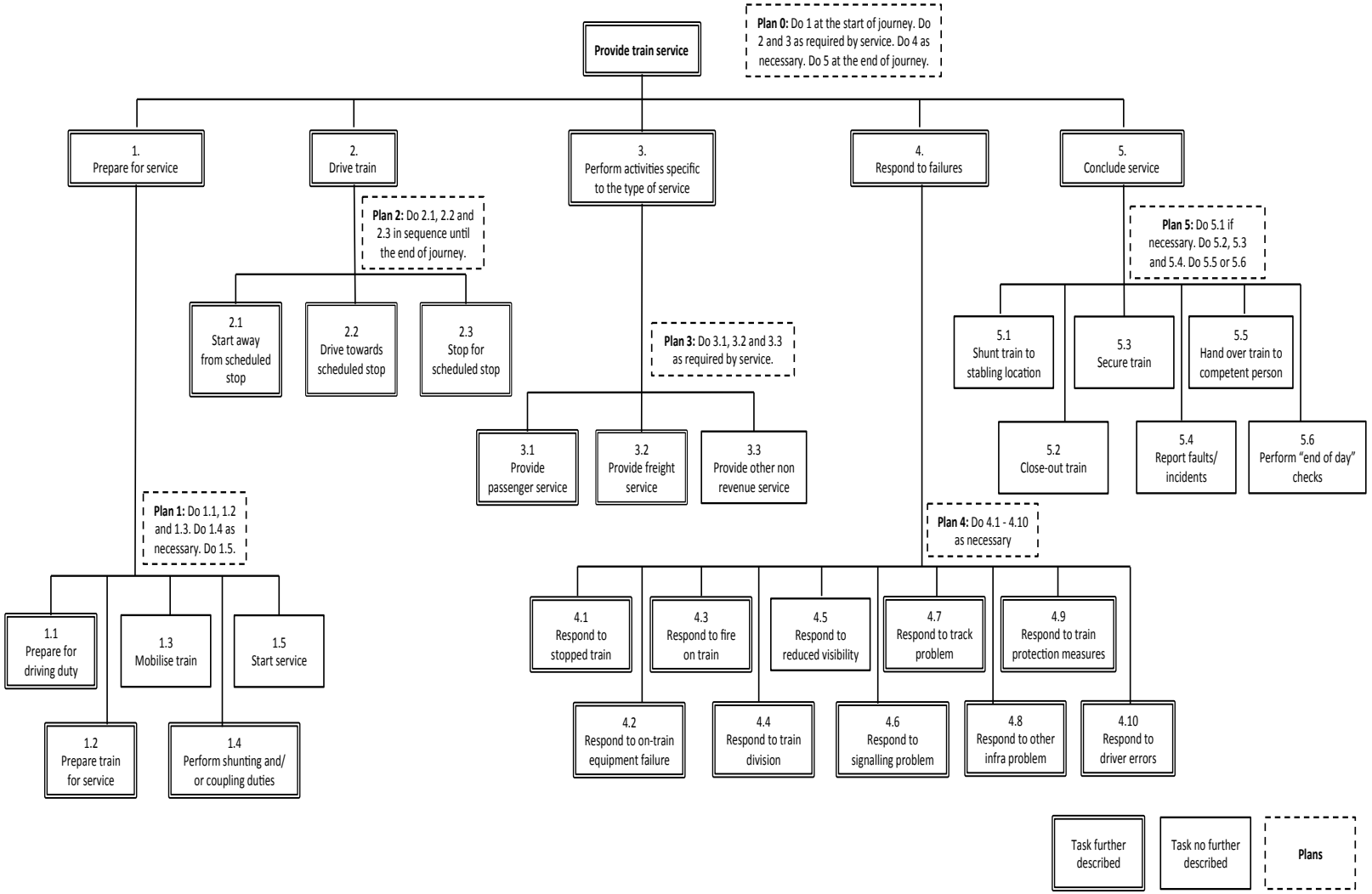



Figure 2-6 The top levels of a hierarchical task analysis for UK train driving (Rail Safety and Standards Board, 2009)

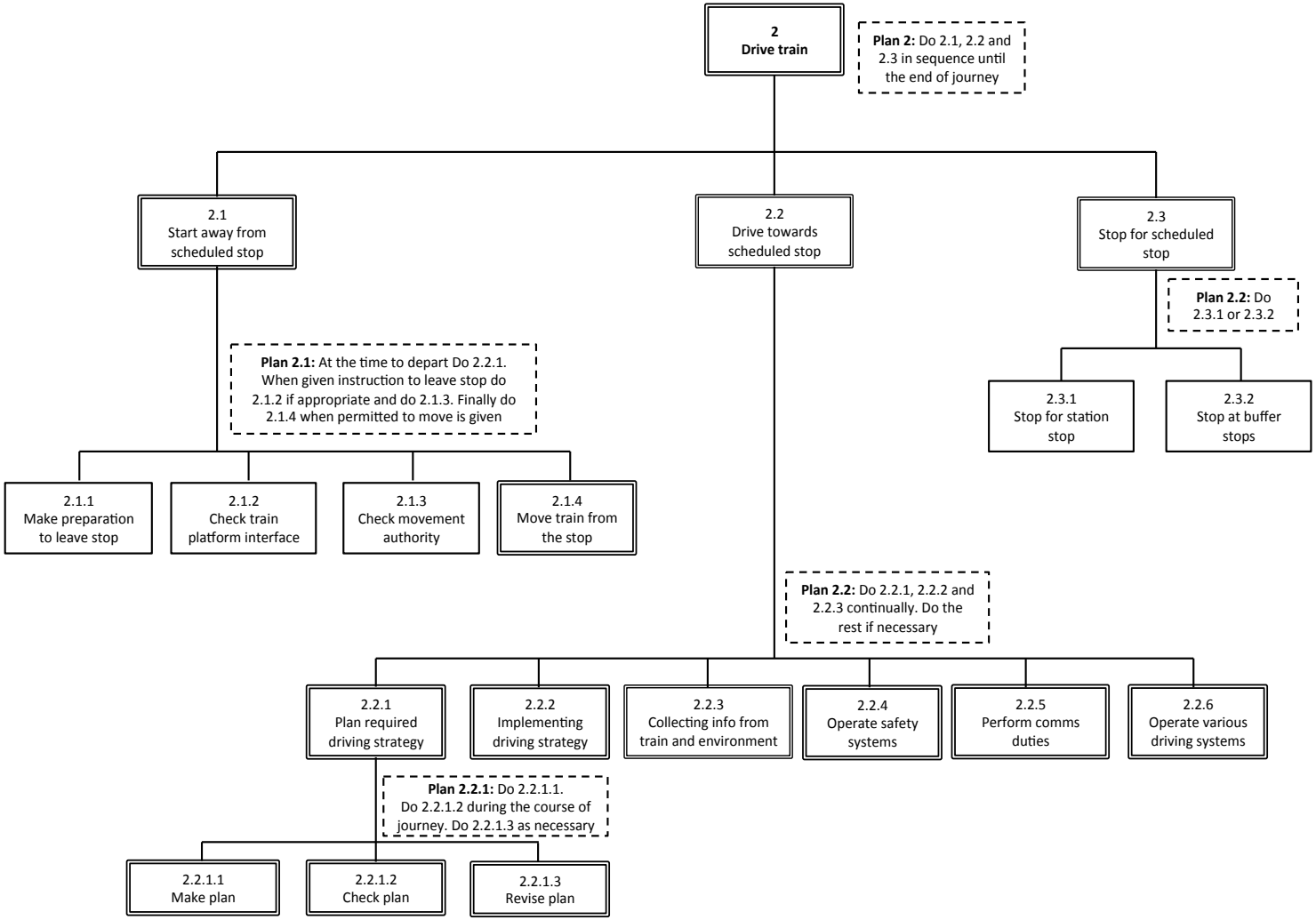



Figure 2-7 Train driving activities in the UK (Rail Safety and Standards Board, 2009)

A similar process was followed for signallers and controllers. The former, as stated in Section 2.2.2.3.2, are responsible for the maintenance of the train traffic along a specific part of the network. The latter ensure train traffic along a larger part of the network by overviewing train services, particularly during times of national disruptions. Based on findings from the literature (Wreathall et al., 2004, Garner et al., 2007), HTA for both types of operators were developed. Figure 2-8 shows HTA for signallers, with their main task defined as "*train traffic control in the assigned area*". This is separated further into: (i) set train routes; (ii) communicate the necessary information; (iii) respond to failures; and (iv) authorise railway works.

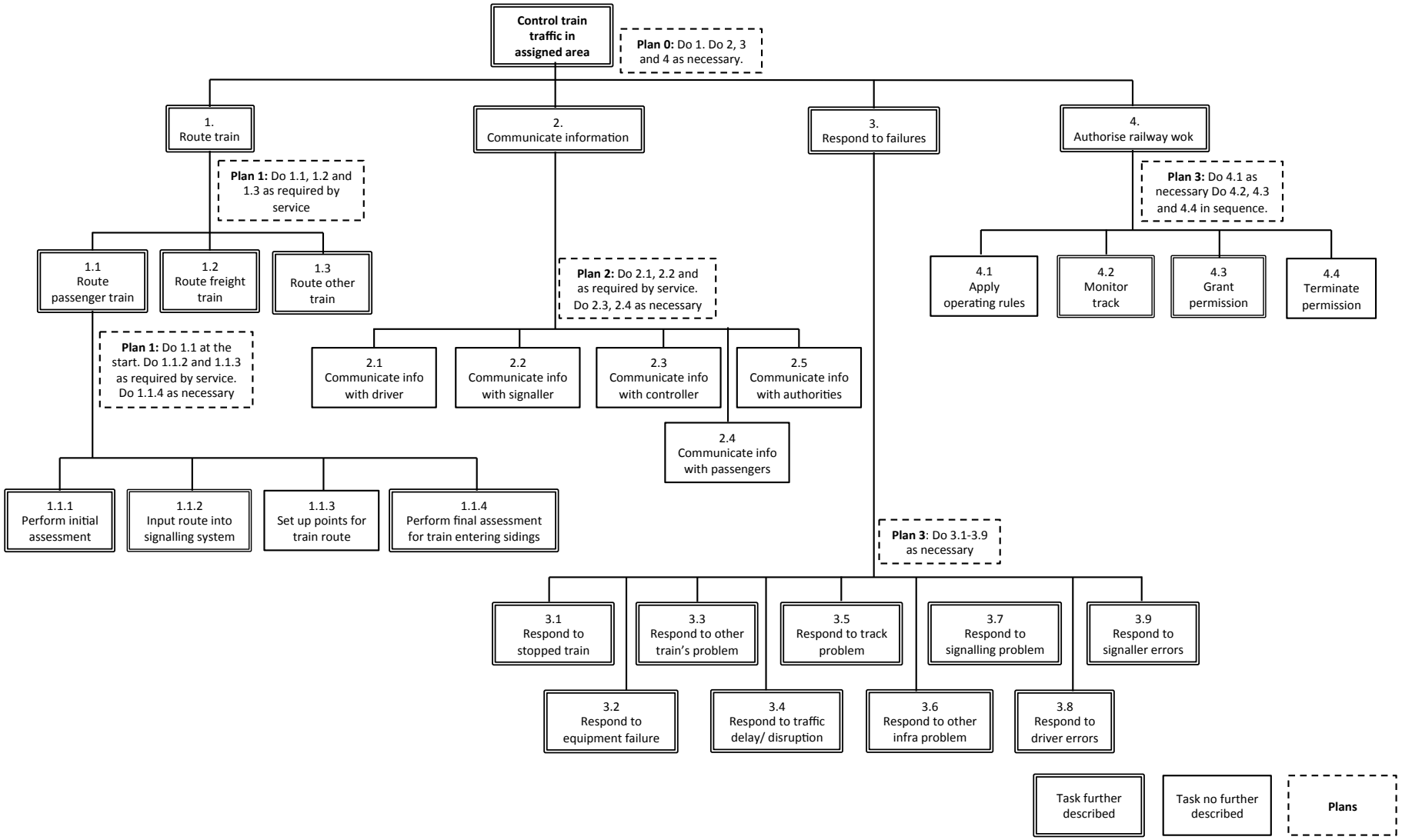



Figure 2-8 The top levels of a hierarchical task analysis for train signaller

Figure 2-9 illustrates in more detail the activities of controllers and indicates their primary goal as "*maintain train traffic*". This is divided into: (i) make an overall plan of train routing; (ii) communicate the necessary information to signallers, controllers, authorities and railway stakeholders; (iii) respond to failures; (iv) plan long term railway works along the network; and finally (v) monitor the management of all types of incidents.

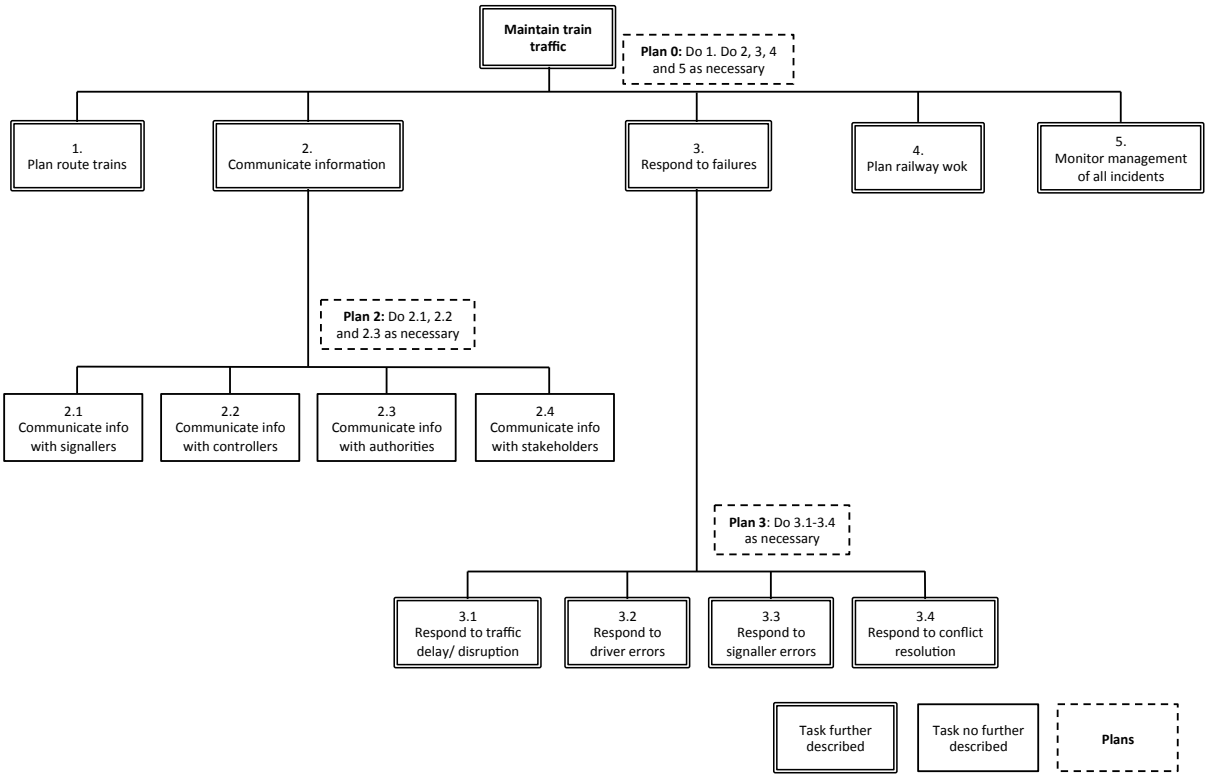



Figure 2-9 The top levels of a hierarchical task analysis for train controller

### 2.3.2 On-site observations

Having identified the tasks of railway operators and conducted the corresponding HTA, on-site observations within train cabins and control centres were performed. As ROSA was primarily developed on the characteristics of the UK mainline railway operations, the aim of the observations was to introduce an architecture that is transferable and could be applied to more than one types of network and between countries.

To satisfy the criterion of transferability to other types of network, observations were conducted in underground metro services. Thus, collaboration was established with the Piccadilly Line of the London Underground Limited. The main reasons for the selection of the Piccadilly Line were: its manual operation; the characteristics of its network, which combines both underground and over ground operations; its good reputation in terms of safety and punctuality; availability and accessibility of its personnel. The Piccadilly Line is

the second longest and fourth busiest line of London Underground, serving 53 stations along the route, with approximately 210 million passengers per year (2012 figures personal communication with Ms. Erica Hudson, LUL Improvements Service Manager).

In addition, to satisfy the criterion of transferability to other countries, observations were conducted in collaboration with a widely recognised European integrated company, the Swiss Federal Railways (SBB). The main reasons for the selection of SBB were: the organisation's excellent reputation for safety; easy and complete access to both trains and control centres; availability and accessibility of SBB personnel and the complexity of its network. The SBB network, although small, is complex and dense. Its complexity is characterised by 804 stations within a 3,139 km network length, almost 347 million passengers and 50 million tonnes of goods transported every year (2010 figures, personal communication with Mr. Andreas Hönger, SBB Safety Manager). The extremely high percentage of its punctuality, combined with the lack of a dedicated freight train line and the multiple maintenance procedures on the network underpin the characteristics of its complexity.

Moreover, observations with additional two UK mainline operators were also performed in order to consolidate the findings from the literature and the HTA. These operators were the First ScotRail and First Hull Trains. The former is the main train operator in Scotland, primarily considered as a commuter service. The latter is the intercity train operator between Hull and London.

Observations at SBB were conducted in the period April-June 2011 on freight, commuter (inter-regio) and intercity passenger trains. In addition, a visit to a shunting yard took place, where trains are secured and stabled after completing service. Observations at FirstScot and First Hull Trains were conducted in December 2012 and February 2013 respectively on stabled trains. Finally, the Piccadilly Line observations took place in March 2013.

Figure 2-10 shows the layout of a typical train cabin of the Piccadilly Line, which has been in use since 1973. Figure 2-11, on the other hand, displays a part of the signalling and controlling system. Both Figures indicate some of the most critical parts of the on board equipment and external infrastructure.

1. Passenger emergency alarm brake override foot switch
2. Speedometer
3. Combined traction and brake controller
4. Passenger doors control
5. Communication handset, radio screen and keyboard
6. Cab audio display unit
7. Selector key

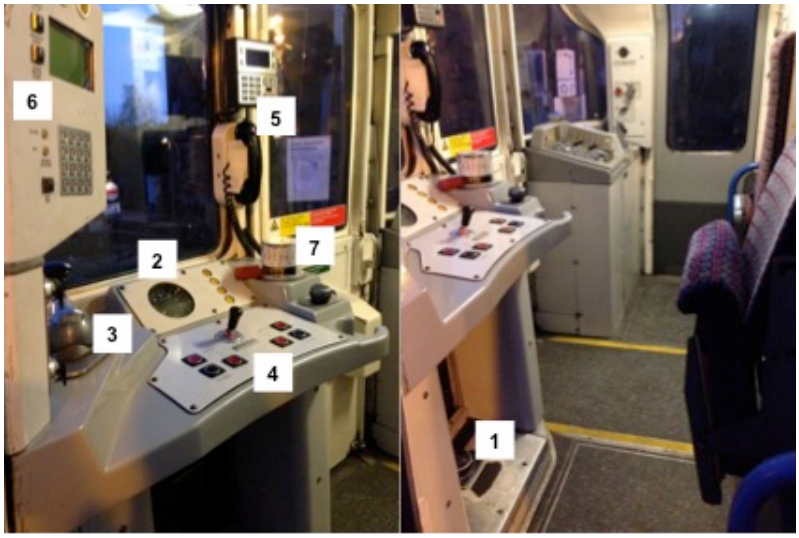



Figure 2-10 Train Cabin, Piccadilly Line, London Underground




Figure 2-11 Part of signalling and controlling system, Piccadilly Line, London Underground

While travelling in train cabins, the author monitored and recorded the tasks and activities of the train drivers, starting from the time they commence the service until the time that they secured the trains. Such activities were then compared with the findings from the task analysis. In addition, the information that train drivers were receiving from the different sources along the route were also recorded and compared to those described in the Figure 2-5.

In addition to train observations, onsite visits were made to two different control centres of the SBB, in May 2011. The first was the master railway control centre located in Zurich airport responsible for monitoring all services in the northern part of Switzerland. The control centre was staffed with both signallers and controllers. The second control centre, located in

Wetzikon of Zurich, is in charge of train movements on a smaller dedicated part of the network and is operated by signallers only.

Observations, in both trains and control centres, confirmed the findings from the literature, in addition to providing further insights on operational procedures and system design. For example, although passenger trains operate on a strict time schedule, this is not the case for freight trains, when running on non-dedicated freight lines. The time planning of such services depends on the passenger train services with the implication that priority is given to a passenger trains over freight trains. The same applies to intercity and commuter trains, where priority is always given to the former. Another example is related to the information train drivers receive along the service. While the SBB train drivers are equipped with personal computers, where not only the route but also stops, speed limits and any unplanned maintenance works are displayed, such equipment is not provided to train drivers on the underground services of the Piccadilly Line. For these operations, train drivers are updated at the depots before starting their duties. During the operation, if something unexpectedly occurs they then contact or are contacted by dedicated signallers via the communication handsets.

### 2.3.3 The Railway Operational System Architecture scheme

The Railway Operational System Architecture (ROSA) has been designed taking into account combined findings from the literature review and on-site observations. As shown in Section 2.3.1, railway operations rely on the coordinated function of operators with the rolling stock and infrastructure. Therefore, the ROSA comprises four main components: (i) the train, and in particular the train cabin, (ii) the control centre, (iii) the signalling and controlling systems and finally (iv) the communication systems. The basic concept of the ROSA represents the interconnections between the two main operators, i.e. train drivers and signallers[2], in conjunction with the on-board train systems, the signalling and control and the railway communication systems. The ROSA is presented in Figure 2-12.

[2] Controllers positions are not found in all control centers.

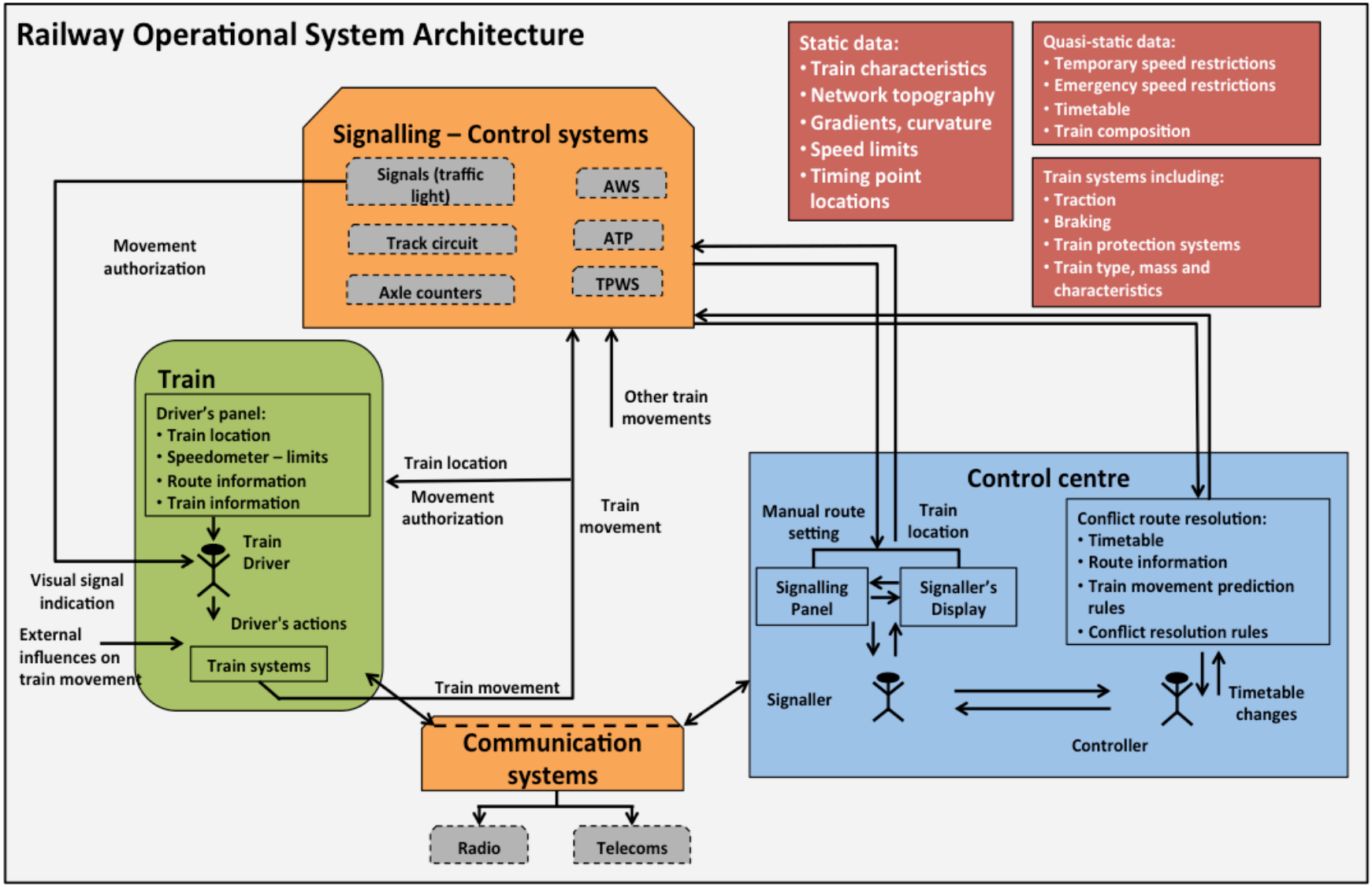



Figure 2-12 The Railway Operational System Architecture (adapted from Rail Safety and Standards Board, 2009 and modified)

As shown in Figure 2-12, the ROSA railway communication means include the legacy conventional radio and modern telecommunication systems (Smith et al., 2013). For the UK main line network, for example, the Cab Secure Radio (CSR) and National Railway Network (NRN) are the main radio systems. On the other hand, the Global System for Mobile Communications-Railways (GSM-R) represents the modern telecommunication system, mandated as standard for the European railways by the European Commission in 1997. With respect to the signaling and controlling of train movements, a number of protection systems are being implemented. Again, in the case of the UK network, apart from the signals (open, at station, at yards), which inform train drivers whether they should proceed, slow down or stop the service, on board and on track safety systems are installed. These include those identified in Section 2.2.2.1.2 (AWS, TPWS and ATP) with additional countermeasures implemented to assure safe train movement. Such measures include axle counters, which are track mounted equipment to count the number of axles on a train along the route and the track circuit (see Section 2.2.2.1.2). Further to the received amount of information from the signaling and controlling systems, an additional set of data flows between the components of the ROSA, informing the operators of the current and overall operational profile. The set of data is classified by the RSSB (2009) as:

- Near real time data, which are delivered within less than 30 seconds (excluding the Piccadilly Line operations).
- Quasi-static data, which are either given at the start of the journey or on a daily basis.
- Static data, which are given when infrastructure, rolling stock or any other changes are made.

Under the assumption of normal operations, train drivers, before starting the operation, are updated by receiving the quasi-static data essential for planning an appropriate driving strategy. The signallers and controllers, on the other hand, who have set the routing of trains on the entire network, inspect and confirm that the services are operated as scheduled. During the journeys or train movements, train drivers constantly receive near real time data and information by the signaling and controlling systems. In addition, signallers and controllers, receive information on their panels with respect to the location of each train along the network. All the exchanged information is indicated in Figure 2-12 by the black arrows. It is shown, for example, that a train driver receives information from the signals while the panel is being updated based on the location of the train relative to the locations of the other trains on the network. Further, signallers and controllers receive also information from the signaling and controlling systems. They use this information to maintain the traffic on the network or perform any necessary changes, aiming at resolving any potential conflict between the trains.

It is apparent from the ROSA that the safety of railway operations depends on the reliability and quality of the technical systems, as well as on the performance of the relevant operators. Again, assuming normal operations where no technical failures are reported, it is sound to claim that either the actions or inactions of operators or any external to the operational scheme factors, e.g. trespassers, can compromise the safety of operations. For example, a train driver who does not acknowledge the received information may maintain the train's speed above the permitted limits, leading to a signal passed at a danger (SPAD) (Dhillon, 2007). However, even under the assumption of technical failures, the performance of operators remains significant with respect to the safe operation of the system. For example, in the case of a failure of the signalling system or on a train, depending on the size of the disruption, new routings may be required, which could be complicated to set and therefore, any wrong actions from the side of signallers or controllers may jeopardise the safety of the system (Dhillon, 2007). Therefore, railway organisations should define and further examine the causes and factors that have an impact on the performance of railway operators, in order to enhance performance of their employees.

## 2.4 Summary

This chapter started by a review of the evolution of railways over the years. This was followed by the description of the modern railway system and a detailed analysis of its main components, elements and stakeholders. Based on this description, this chapter then introduced the operational concept of the railway system, from the perspective of operators, referred to as the Railway Operational System Architecture (ROSA). The ROSA displays the operational constituents, and identifies the interactions between operators, equipment and transmitted information. In addition, ROSA has facilitated the identification of a number of issues that should be taken into account concerning the safety of railways operations, including the reliability of infrastructure and rolling stock as well as the significant influence of human performance. The next chapter investigates the contribution of each component of the railway system to the overall safety of railways.

# Chapter 3 FUNDAMENTALS OF RAILWAY SAFETY

The previous chapter presented the modern railway system and its components, which were classified as *controlled* and *uncontrolled*. Furthermore, it detailed the operational framework of the system by introducing its architecture. The main characteristics of the operational concept of railways were identified together with the interactions and dependencies amongst its components.

Following the description of the system and its operational framework, this chapter completes the description of the context of this thesis by addressing the relevant fundamentals of railway safety. It starts with the relevant definitions of railway safety, accidents, incidents, fatalities and injuries. The chapter justifies the need for research on the contribution of the human operators to railway safety through evidence from safety reports from a variety of sources including the European Railway Agency (ERA), US Federal Railway Administration (FRA), Australian Transport Safety Board (ATSB), Transportation Safety Board of Canada (CTSB) and the International Union of Railways (UIC).

## 3.1 Introduction to railway safety

Railways are a vital means of transport, moving millions of passengers and a vast amount of goods worldwide, as shown in Table 3-1. The safety of railway operations is therefore an issue of international concern.

Table 3-1 World railway traffic (International Union of Railways, 2011a)

| **Passenger – km (billions)** | **2006** | **2007** | **2008** | **2009** | **2010** |
|---|---|---|---|---|---|
| Europe* | 642.1 | 642.6 | 659.7 | 624.1 | 611.9 |
| Africa | 61.5 | 61.6 | 62 | 62.2 | 62.3 |
| America | 12.8 | 13.3 | 14 | 13.5 | 12 |
| Asia and Oceania | 1,646 | 1,788.6 | 1,950.9 | 2,012.0 | 2,079.3 |
| World | 2,362.4 | 2,506.1 | 2,686.6 | 2,711.8 | 2,765.5 |
| **Tonne – km (billions)** | **2006** | **2007** | **2008** | **2009** | **2010** |
| Europe* | 2,646.6 | 2,813.6 | 3,103 | 2,411.4 | 2,603.8 |
| Africa | 142.2 | 139.2 | 138.4 | 137.1 | 139.2 |
| America | 3,519.5 | 3,540.2 | 3,513.8 | 2,973.2 | 3,076.1 |
| Asia and Oceania | 2,872.6 | 3,095.9 | 3,452.7 | 3,466.2 | 3,462.0 |
| World | 9,180.9 | 9,588.9 | 10,207.9 | 8,987.9 | 9,281.1 |

| Length of lines (km) | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|
| Europe* | 349,458 | 348,788 | 349,000 | 353,747 | 370,370 |
| Africa | 52,159 | 52,400 | 52,482 | 52,299 | 50,275 |
| America | 385,272 | 389 863 | 386,773 | 383,079 | 375,774 |
| Asia and Oceania | 221,788 | 222,645 | 221,827 | 224,151 | 224,205 |
| World | 1,008,677 | 1,013,696 | 1,010,082 | 1,013,276 | 1,020,624 |

*including Russian Federation and Turkey

Railways are considered to be one of the safest modes of travel, as expressed by the number of passenger fatalities. For example, on the European (EU-27, CH, NO) mainline network for the period 2008-2010, 188 passenger fatalities were recorded over 1,200 billion passenger-km (European Railway Agency, 2012c), while the registered fatalities of road accidents were more than 100,000 (European Commission, 2013). Nevertheless, railway transportation is prone to structural risks, such as technical failures, human failures and vandalism. When railway accidents occur, e.g., the recent fatal derailment in Santiago de Compostela of Spain (Topham, 2013), society is distressed and debates on the safety of railways arise. Hence, railways safety is a matter of great concern for all relevant stakeholders including, operational, regulatory and governmental entities, as well as the academic community and the public, all having the goal of high performance in terms of reliability, quality and safety (Wilson et al., 2007a).

According to Collins English Dictionary (Collins English Dictionary, 2009, p. 1442) safety is defined as "*freedom from danger or risk of injury*". However, railways, similar to any systems consisting of technical and human components, cannot be completely free from either danger or risks (Huang, 2009). Therefore, many researchers tend to associate safety with accident prevention and define safety as "*as few accidents as possible*" (Huang, 2009).

In the railway industry safety is defined as the "*freedom from unacceptable risk of harm*" (CENELEC, 1999), where unacceptable risk is identified by the risk classification matrix shown in Table 3-2. This matrix is recommended in the European Standards EN 50126-129 (1999, 2001, 2003), and used in the railway industry because it provides a quick and easy method to represent risk (Bearfield et al., 2012). It consists of a two-dimensional matrix with the Y-axis representing the frequency or likelihood of a hazard occurring. The X-axis, on the other hand, shows the severity or consequences of the hazard. The final risk, which is classified as low, medium and high, is indicated as a combination of the frequency and consequence of the hazard. Based on such frequencies and severities associated to certain cases and specific scenarios, each band of the matrix can be quantified.

Table 3-2 Typical risk classification matrix

| | | Severity Band | | | |
|---|---|---|---|---|---|
| **Frequency Band** | | 4 Catastrophic | 3 Critical | 2 Marginal | 1 Minor |
| A | Frequent | High | High | High | Medium |
| B | Probable | High | High | Medium | Medium |
| C | Occasional | High | Medium | Medium | Low |
| D | Remote | Medium | Medium | Low | Low |
| E | Improbable | Medium | Medium | Low | Low |
| F | Incredible | Medium | Low | Low | Low |

The safety of the railway system is affected by any failures associated with its components and elements. Therefore, railway organisations and safety authorities monitor, record and investigate the failures caused by the *controlled* and *uncontrolled* components of the system, in order to identify the principle causes that lead to railway occurrences. Since the number of serious and fatal accidents is relatively small, less serious events are also investigated. These are used as indicators of incidents, i.e. precursors to accidents, which under different conditions could lead to an accident.

Railway safety can be measured in terms of the number of: (i) accidents, incidents, fatalities and injuries, (ii) fatalities per train kilometre, (iii) accidents per train kilometre and (iv) fatalities and injuries per passenger kilometre. While (i) provide the most obvious indicators of safety, (ii)-(iv) support the identification of safety trends and comparisons between the railway organisations as well as the different modes of transport. However, the definitions used by safety authorities to describe the different types of railway occurrences and their consequences are not always consistent. For example, the European definitions clearly define the concept of serious accidents, whereas the US definitions do not classify the different types of accidents. This has the potential to result in differences in statistical outcomes and their interpretation. However, as shown in the next section, regardless of the differences in definitions, national and international safety authorities follow similar guidelines with respect to the occurrences that must be reported.

## 3.2 Definitions and description of railway occurrences

This section presents the definitions of railway occurrences and their consequences as given by five international railway administrators and authorities: the European Railway Agency (ERA), the US Federal Railroad Administration (FRA), the Canadian (CTSB) and

Australian (ATSB) Transport Safety Boards and, finally, the International Union of Railways (UIC). In addition, the events that must be reported are identified.

### 3.2.1 The European Railway Agency (ERA) definitions

The European legislation (European Commission, 2009) requires every Member State to report to the ERA any serious or significant accident that occurs. The National Investigation Bodies (NIB) are responsible to investigate and report all serious accidents, whilst the National Safety Authorities (NSA) are required to report all significant accidents, as determined by the European Commission's (EC) definitions.

According to the EC (2004) railway occurrences are classified as:

1. **Serious accidents** refer to **"***any train collision or derailment of trains, resulting in the death of at least one person or serious injuries to five or more persons or extensive damage to rolling stock, the infrastructure or the environment and any other similar accident with an obvious impact on railway safety regulation or the management of safety; "extensive damage" means damage that can immediately be assessed by the investigated body to cost at least EUR 2 million in total*".

2. **Accidents** describe any "*unwanted or unintended sudden event or a specific chain of such events which have harmful consequences; accidents are divided into the following categories: collisions, derailments, level-crossing accidents, accidents to person caused by rolling stock motion, fires and others*".

3. **Incidents** are "*any occurrence, other than accident or serious accident, associated with the operation of trains and affecting the safety of operation*".

In addition, the EC (2009) defines **significant accidents** as "*any accident involving at least one rail vehicle in motion, resulting in at least one killed or seriously injured person, or in significant damage to stock, track, other installations or environment, or extensive disruptions to traffic. Accidents in workshops, warehouses and depots are excluded. Significant damage is defined as equivalent to EUR 150 000 or* more, and *extensive disruption to traffic with tracks blocked for more than 6 hours. Suicides and suicide attempts are excluded".* This broad definition covers a wider range of events than serious accidents, which by legislation should be reported.

Finally, the EC (2009) defines a **fatality** as "*any person killed immediately or dying within 30 days as a result of an accident, excluding suicides*", while a **serious injury** is defined as

"*any person injured who was hospitalised for more than 24 hours as a result of an accident, excluding attempted suicides*". All other injuries are considered minor injuries.

### 3.2.2 The Federal Railroad Administration (FRA) definitions

The FRA (2003, 2011) requires that "*all accidents and incidents resulting in injury or death to an individual or damage to equipment or a roadbed arising from the carrier's*" are reported. These occurrences include *"train collisions, derailments, and other events involving the operation of on-track equipment and causing reportable damage above an established threshold; impacts between railroad on-track equipment and highway users at crossings; and all other incidents or exposures that cause a fatality or injury to any person, or an occupational illness to a railroad*". The occurrences are further classified into:

1. **Train accidents** refer to "*any collision, derailment, fire, explosion, act of God, or other event involving the operation of on-track equipment (standing or moving) that results in total damages to all railroads involved in the event that is greater than the current reporting threshold (e.g. the threshold for 2008 was $8,500) to railroad on-track equipment, signals, track, track structures and roadbed*".

2. **Train incidents** are events "*involving the movement of on-track equipment that results in a reportable casualty but does not cause reportable damage above the threshold established for train accidents*".

3. **Non-train incidents** are events "*resulting in a reportable casualty, but does not involve the movement of on-track equipment nor cause reportable damage above the threshold established for train accidents*".

4. **Other accidents** describe *"other than train accidents or crossing incidents that cause physical harm to persons*". Such events primarily define accidents caused by trespassers.

In addition, the FRA considers **fatalities** as "*any death of one or more persons instantly or within a month after the event occurs*". On the other hand, **injuries** are described as "*any harm to a person resulting from a single event, activity, occurrence ort exposure of short duration*".

### 3.2.3 The Canadian Transport Safety Bureau (CTSB) definitions

According to the Canadian regulations (Transportation Safety Board of Canada, 2012), **railway occurrences** should be reported by the accident investigation authorities if they describe "*any accident or incident associated with the operation of rolling stock on a railway or any situation or condition that the Safety Board has reasonable grounds to believe could, if left unattended, induce an accident or incident*".

The CTSB (2012) defines **accidents** as "*any event resulting directly from the operation of rolling stock where:*

1. *A person sustains a serious injury or is killed as a result of: being on board or getting off the rolling stock or coming into contact with any part of the rolling stock or its contents.*
2. *The rolling stock is involved: in a grade-crossing collision; in a collision or derailment and is carrying passengers in a collision or derailment and is carrying dangerous goods or is known to have last contained dangerous goods the residue of which has not been purged from the rolling stock; sustains damage that affects its safe operation; causes or sustains a fire or explosion, or causes damage to the railway, that poses a threat to the safety of any person, property or the environment*".

In addition, **incidents** are defined as "*any occurrence resulting directly from the operation of rolling stock, where: a risk of collision occurs; an unprotected main track switch is left in an abnormal position; a railway signal displays a less restrictive indication than that required for the intended movement of rolling stock; an unprotected overlap of operating authorities occurs; a movement of rolling stock exceeds the limits of its authority; there is runaway rolling stock; any crew member whose duties are directly related to the safe operation of the rolling stock is unable to perform the crew member's duties as a result of a physical incapacitation that poses a threat to the safety of any person, property or the environment; any dangerous goods are released on board or from the rolling stock*".

Finally with respect to the consequences, **serious injuries** are defined as "*any injury that is likely to require admission to a hospital*", whilst neither **fatalities** nor **other injuries** are defined explicitly.

### 3.2.4 The Australian Transport Safety Bureau (ATSB) definitions

The "Memorandum of Understanding" (Australian Transport Safety Bureau, 2005) between the ATSB and the Australian safety regulators defines the reportable **railway occurrences**

as "*any accident or incident involving a railway train or other railway vehicle operated on rails, whether in motion or not, or other event on railway property affecting the safety of persons and property, including: train collisions, train derailments, fires, explosions, acts of God or other events; slips, trips and falls on trains or railway infrastructure*".

According to the Australian Infrastructure and Transport Administration (2009) **accident** is described as "*an investigable matter involving a transport vehicle where: (a) a person dies or suffers serious injury as a result of an occurrence associated with the operation of the vehicle; or (b) the vehicle is destroyed or seriously damaged as a result of an occurrence associated with the operation of the vehicle; or (c) any property is destroyed or seriously damaged as a result of an occurrence associated with the operation of the vehicle*". Any other occurrence is an **incident**.

**Fatalities** describe "*any death within 30 days of a railway occurrence, from injuries sustained in that occurrence*", while **serious injuries** are characterised by "*any person admitted to hospital as the result of injuries sustained in a railway occurrence*". Finally, any injury, not followed by admission to a hospital, is considered as a **minor injury** (Australian Transport Safety Bureau, 2005).

### 3.2.5 The International Union of Railways (UIC) definitions

The UIC acknowledge and adopt the EC definitions in their entirety as given in Section 3.2.1. However, amongst the definitions, the declaration of significant accidents is in accordance with the definition given by the Commission Regulation (EC) N° 1192/2003 (2003). According to this "*significant accident means any accident involving at least one rail vehicle in motion, resulting in at least one killed or seriously injured person, or in significant damage to stock, track, other installations or environment, or extensive disruptions to traffic. Accidents in workshops, warehouses and depots are excluded*".

Given the various definitions, it can be seen that the safety authorities do not describe the types of railway occurrences and their consequences in a consistent manner. For example, the EC definitions indicate three severity levels of accidents, while the remaining definitions provide two levels. In addition, although the EC, Canadian and Australian definitions account for serious injuries, the US definitions do not provide such a classification. Nevertheless, it can be argued that the differences in the definitions are mainly related to the severity classifications rather than to the content. Thus, the EC definitions are adopted and used in this thesis, as it is considered that they provide the more detailed and well distinguished definitions in terms of the events severity.

### 3.2.6 Accident causes and precursors to occurrences

The ERA (2013b) specify that railway occurrences can be caused by: (i) infrastructure failures, e.g. signals, switches and crossings or electricity supply lines failures; (ii) rolling stock failures, such as failures in the braking systems, wheels or axles; (iii) operational failures, i.e. human performance and human factors; (iv) third parties, including level-crossing events, occurrences caused by rail passengers or unauthorised personnel. In addition to the causes, ERA monitor precursors to accidents by annually reporting all the cases that involve broken rails, track buckles, signals passed at danger (SPAD), wrong-side signalling failures, broken wheels and broken axles.

Similarly, the FRA (2013) indicate that railway occurrences are caused by: (i) human factors, e.g. use of brakes, signals observations and use of switches; (ii) track defects, e.g. track geometry, roadbed and switches; (iii) signal defects, e.g. signal and communication codes; (iv) equipment defects, e.g. brakes, wheels, locomotives, axles; and (v) miscellaneous causes, such as environmental conditions, loading procedures or unusual operational conditions.

Likewise, the CTSB (2012) and the ATSB (2005) acknowledge technical failures, human failures, third parties behaviours and miscellaneous events as the most common causes of railway occurrences, while only the ATSB detects animal activities as an additional cause. In addition the ATSB report SPADs, track irregularities and load irregularities as precursors to accidents.

Finally, the UIC (2012) identifies that railway occurrences can be caused by internal or/and external causes. The latter are further subdivided into third parties and environmental conditions, whilst the former into human factors and railway sub-systems. Furthermore, the UIC monitors the number of SPADs, broken rails, broken wheels or axels, overheated axle boxes and dangers due to over-speeding as precursors to accidents, in particular train derailments.

Given the various classifications, it can be seen that all the authorities acknowledge the same factors as the causes of the several railway occurrences. On the other hand, only three of them monitor and account for precursors. Thus, for the purposes of this thesis the ERA classifications are adopted, in order to be in agreement with the definitions adopted for the severity of railway occurrences.

## 3.3 Railways safety statistics

Based on the definitions in the previous section, this section discusses the development of railway safety, and investigates the types of accidents and their causes. In addition, it examines the contribution of each of the components of railway system to the overall system safety. The investigation relies on railway accidents and fatalities records, as these are both difficult to conceal and must be reported within the reporting scheme of all organisations.

For the contribution of components in the overall railway safety, data were extracted from the safety performance reports of ERA, FRA, ATSB, CTSB and UIC for the period 2006-2011. The ERA data include those from all EU Member States, Norway (NO) and Switzerland (CH), while the UIC data are from 19 European railway entities, including the UK's Network Rail, the Swiss Federal Railways, the ProRail in the Netherlands and the Deutsche Bahn in Germany. The FRA and Australian data include information from all railway companies that operate in the US and Australia, while the Canadian data represents only railways under federal jurisdiction.

### 3.3.1 Development of railway safety

European data on train derailments and collisions in Figure 3-1, from the ERA (2012c) database shows that the overall railway safety performance of the EU Member States, NO and CH has progressively improved by up to 70% since 1990. In addition, Figure 3-1 presents the 5-year moving average statistic for safety, which suggests that despite the positive long-term trend over the past two decades, safety improvement has been slowing down since 2005. This can be exlained by the inability of the current safety and protection systems to provide any further significant improvement in safety.

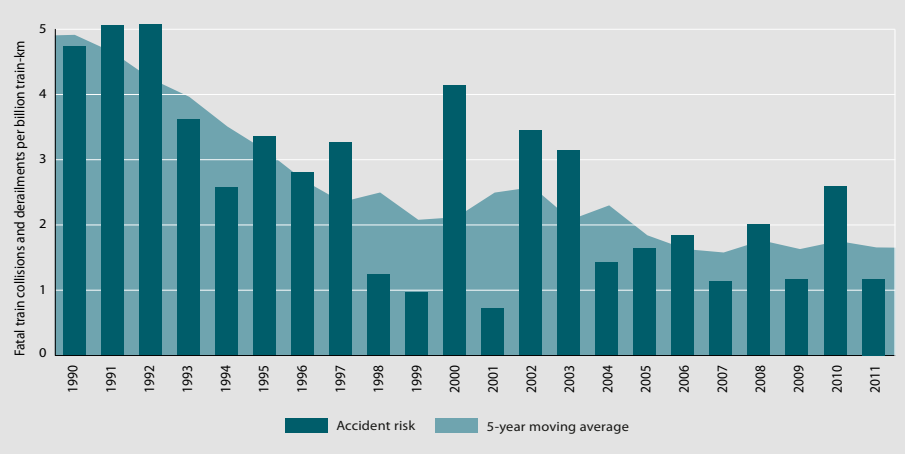



Figure 3-1 Fatal train collisions and derailments per billion train-kilometers in 1990–2011 for the EU-27, Switzerland and Norway (European Railway Agency, 2012c)

Furthermore, the same data show a downgrade in the number of railway accidents that resulted in more than five fatalities. Figure 3-2 presents data derived from accidents at level crossings, fire in rolling stock and accidents involving persons hit by rolling stock in motion, in addition to train collisions and derailments. It is observed that the average number of eight accidents per year during the 1990s has been reduced to five in the 2000s, while the number of fatalities caused by those accidents has dropped accordingly. Considering that only three serious accidents occurred in 2010 and one in 2011, it can be claimed that the downgrade is substantial.

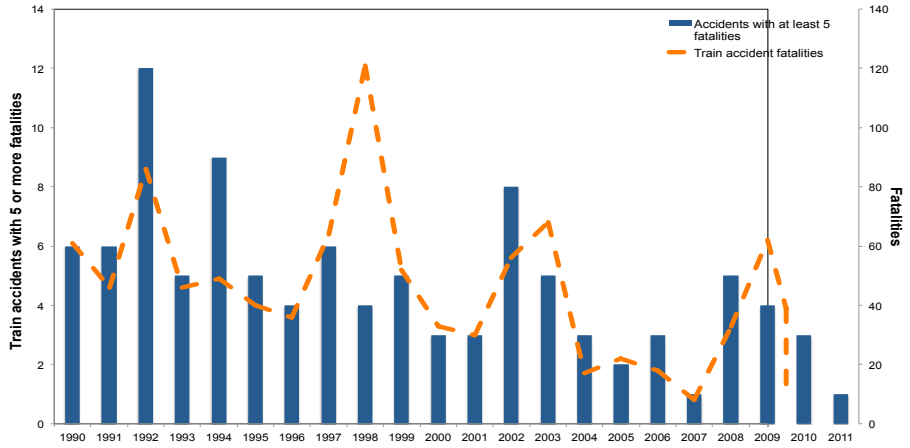



Figure 3-2 Train fatalities and train accidents with at least 5 fatalities (1990–2011) (European Railway Agency, 2012c)

Table 3-3 presents a summary of significant accidents in Europe and their consequences, for the period 2006-2011. The large difference in the number of significant accidents in 2006 and the following years is due to a change in the reporting procedures in some of the EU State members (European Railway Agency, 2012c). Nevertheless, the trend of all reportable variables both in absolute numbers as well as per million train kilometres, indicate significant

improvement in safety of railway operations, in particular with respect to the number of recorded fatalities. Similar findings and positive trends result from the UIC (2012) database (Table 3-4). This was expected as both the ERA and UIC databases to a great extent use data from the same sources.

Table 3-3 Summary of significant accidents and their consequences of the European State Members (European Railway Agency, 2012a)

| ERA database | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|
| Number of significant accidents | 5,057 | 3,831 | 3,744 | 3,073 | 2,421 | 2,377 |
| Number of fatalities | 1,471 | 1,519 | 1,480 | 1,391 | 1,265 | 1,188 |
| Number of serious injuries | 1,459 | 1,375 | 1,383 | 1,114 | 1,241 | 1,037 |
| Significant accidents per million train-km | 1.26 | 0.92 | 0.88 | 0.75 | 0.60 | 0.57 |
| Fatalities per million train-km | 0.37 | 0.36 | 0.35 | 0.34 | 0.31 | 0.29 |
| Serious injuries per million train-km | 0.36 | 0.33 | 0.33 | 0.27 | 0.31 | 0.25 |
| Number of million train kilometres | 4,011.66 | 4,184.45 | 4,243.61 | 4,101.84 | 4,065.06 | 4,159.55 |

Table 3-4 Summary of significant accidents and their consequences from 19 European Railway Companies (International Union of Railways, 2012)

| UIC data* | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|
| Number of significant accidents | 2,256 | 2,215 | 2,178 | 2,016 | 1,935 | 1,812 |
| Number of fatalities | 1,230 | 1,320 | 1,194 | 1,244 | 1,049 | 985 |
| Number of serious injury accidents | 2,031 | 2,074 | 1,977 | 1,815 | 1,662 | 1,574 |
| Significant accidents per million train-km | 0.58 | 0.56 | 0.53 | 0.52 | 0.57 | 0.46 |
| Fatalities per million train-km | 0.32 | 0.33 | 0.29 | 0.32 | 0.31 | 0.25 |
| Serious injuries per million train-km | 0.52 | 0.52 | 0.48 | 0.47 | 0.49 | 0.39 |
| Number of million train kilometres | 3,881.1 | 3,948.4 | 4,078.8 | 3,855.4 | 3,392.6 | 3,980.0 |

*NRIC (Bulgaria) ended its contribution in 2010, MAV (Hungary) did not provide data for 2008 and HZ (Croatia) started its input in 2010

Similar to the findings in Europe findings, the FRA (2012) data suggest that the safety of the US railways has also gradually improved (Figure 3-3). The number of train accidents has progressively dropped by 68% since 1990, and the number of fatalities by 46%. This is explained by the numerous and effective technological improvements and company-wide

safety programs that were implemented within the US railways. The former include, track and equipment improvements, e.g. welded rails and hot box detectors. The latter include remote control operations, personnel fatigue countermeasures and interactive video training programs (French, 2008).

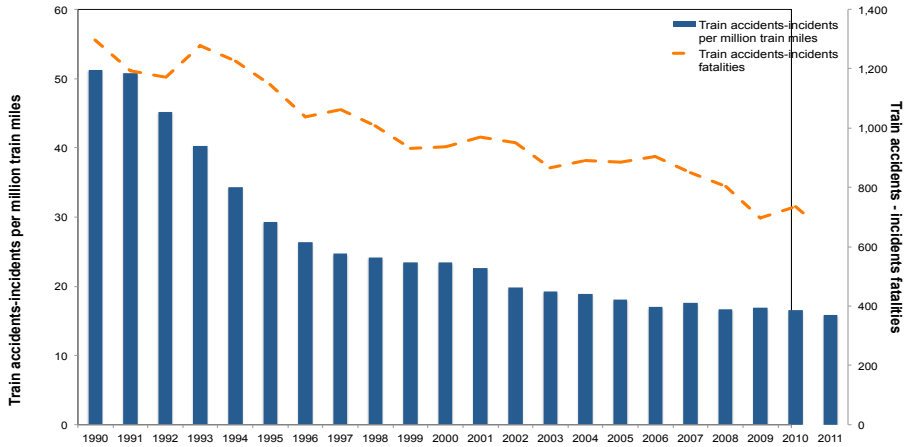



Figure 3-3 US train accidents - incidents per million-train miles and train accidents - incidents fatalities (Federal Railroad Administration, 2012)

The US trends for the number of accidents, fatalities and injuries for the period 2006-2011 are given in Table 3-5. Two main observations can be made. Firstly, the number of recorded accidents in the US is approximately five times higher than in Europe. However, this is due to the differences in the definitions, where the FRA does not distinguish significant from non-significant types of accidents. Secondly, the number of US million train kilometres is almost half of the Europe's, as due to its large size the US railway network is less dense.

Table 3-5 Summary of US train accidents and their consequences (Federal Railroad Administration, 2012)

| FRA database | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|
| Number of accidents | 13,803 | 13,936 | 12,958 | 11,263 | 11,628 | 11,434 |
| Number of fatalities | 903 | 851 | 803 | 696 | 736 | 698 |
| Number of injuries | 8,797 | 9,669 | 9,064 | 8,037 | 8,372 | 8,333 |
| Accidents per million train-miles | 16.97 | 17.56 | 16.74 | 16.86 | 16.50 | 15.93 |
| Fatalities per million train-miles | 1.11 | 1.07 | 1.04 | 1.04 | 1.04 | 0.97 |
| Injuries per million train-miles | 10.81 | 12.18 | 11.71 | 12.03 | 11.88 | 11.61 |
| Number of million train miles | 813.58 | 793.62 | 774.05 | 667.97 | 704.83 | 717.6 |

Similar to the European and US findings, the overall safety of Canadian railways has also improved. The CTSB data (2012) show reductions of 23% and 13% in total train accidents since 2002 and 2006 respectively. In addition, the largest proportion of train accidents occurs on non-main track operations. These typically result in minor consequences and take place during switching operations at speeds of less than 10 mph. Therefore, focus should be on the main track accidents. As shown in Figure 3-4, slight improvement is observed with regard to main track accidents per million main track train miles (MMTTM). Although the main-track accident rate increased by 18% (excluding level crossings and trespasser accidents) compared to 2010, the long-term rate is down by 1.5% from the ten-year average of 2011. On the other hand, considerably higher improvement is observed on the number of fatalities, which gradually dropped by 20% since 2002.

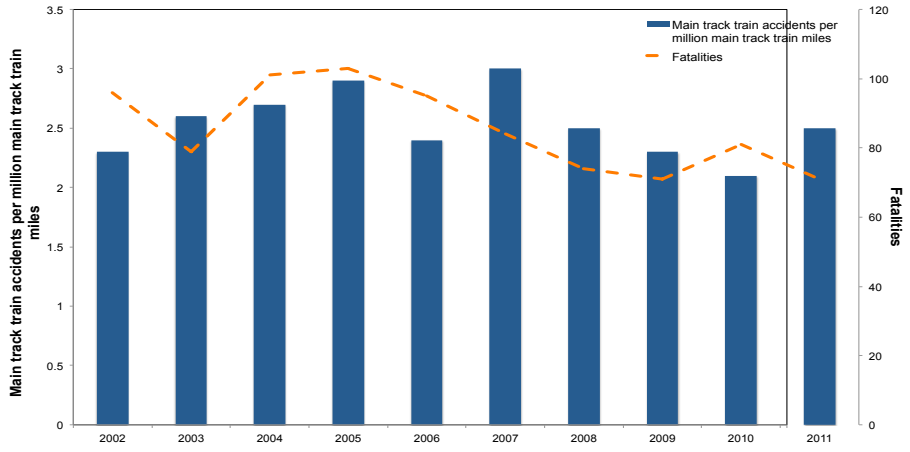



Figure 3-4 Canadian train accidents and main track accidents per million main track train miles (Transportation Safety Board of Canada, 2012)

Table 3-6 displays Canadian trends for the period 2006-2011. It is worth mentioning the low number of fatalities and injuries on the Canadian railways compared to the European and US. This is explained by the nature of the Canadian network, which is primarily used to transfer goods. Therefore, when train accidents occur, they usually involve freight trains and subsequently involve either the train crew or level crossing users.

Table 3-6 Summary of Canadian main track train accidents and their consequences (Transportation Safety Board of Canada, 2012)

| CTSB database | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|
| Number of accidents | 209 | 257 | 209 | 169 | 161 | 192 |
| Number of fatalities (all tracks) | 95 | 84 | 74 | 71 | 81 | 71 |
| Number of serious injuries (all tracks) | 71 | 58 | 64 | 50 | 62 | 51 |
| Main track accidents per MMTTM | 2.41 | 3.04 | 2.52 | 2.34 | 2.07 | 2.45 |
| Fatalities per MMTTM (all tracks) | 1.09 | 0.99 | 0.89 | 0.98 | 1.04 | 0.91 |
| Serious injuries per MMTTM (all tracks) | 0.82 | 0.69 | 0.77 | 0.69 | 0.80 | 0.65 |
| Million main track train miles (MMTTM) | 86.9 | 84.5 | 83.1 | 72.2 | 77.6 | 78.3 |

Finally, data from the ATSB (2012) indicate a slight improvement on the safety performance over the last decade, as measured by the number of train derailments, collisions and fatalities.

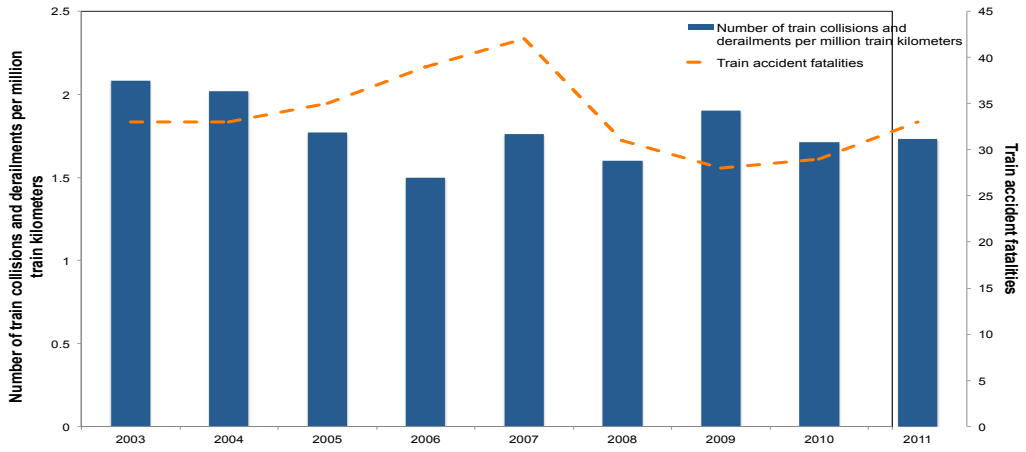



Figure 3-5 Australian train derailments, collisions and fatalities (Australian Transport Safety Bureau, 2012)

Table 3-7 presents the Australia trends for the last six years. Again, it can be observed the low number of fatalities and serious injuries, despite the high number of accidents per million train miles. Similar to the Canadian network, this observation is explained by the nature of the Australian network, which is primarily used to transfer goods.

Table 3-7 Summary of Australian train accidents and their consequences (Australian Transport Safety Bureau, 2012)

| ATSB database | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|
| Number of accidents | 804 | 766 | 878 | 842 | 796 | 808 |
| Number of fatalities | 39 | 42 | 31 | 28 | 29 | 39 |
| Number of serious injuries | 135 | 183 | 114 | 91 | 38 | 66 |
| Accidents per million train miles | 8.93 | 8.37 | 9.37 | 9.25 | 8.73 | 8.37 |
| Fatalities per million train miles | 0.43 | 0.46 | 0.33 | 0.31 | 0.32 | 0.40 |
| Serious injuries per million train miles | 1.50 | 2.00 | 1.21 | 1.00 | 0.42 | 0.68 |
| Number of million train miles | 90 | 91.47 | 93.75 | 91.05 | 91.22 | 96.53 |

In summary, from the high level analysis of safety performance of European, US, Canadian and Australian railways, it is clear that safety has improved. It is also clear that the current safety protection measures are quickly approaching their maximum capabilities. Therefore, the railways industry must continue their efforts to improve safety by investing in the most vulnerable and fundamental areas: infrastructure, rolling stock, employees' competence and public awareness. This requires a detailed understanding of railways accidents to identify and prevent the causes that lead to such events.

### 3.3.2 Railway accidents breakdown

Railway accidents are classified according to their characteristics, as shown in Table 3-8. Despite the differences in the classifications all the schemes include the same types of events. These are train derailments, train collisions, train collisions with obstacles, accidents to persons hit by trains in motion, accidents at level crossings, fire in rolling stock, and miscellaneous events. Based on Table 3-8, Figures 3-6 to 3-9 present the contribution of each type of accident to the total train accidents for all databases.

Table 3-8 Type of accidents as classified per database

<table>
<tr><th colspan="5">Type of accidents as classified per database</th></tr>
<tr><th>UIC*</th><th>ERA*</th><th>FRA</th><th>CTSB**</th><th>ATSB</th></tr>
<tr><td>Train derailments</td><td>Train derailments</td><td>Train derailments</td><td>Train derailments** (all)</td><td>Train derailments</td></tr>
<tr><td>Train collisions</td><td>Train collisions and collisions with obstacles within clearance gauge</td><td>Train collisions</td><td>Train collisions** (all)</td><td>Train collisions</td></tr>
<tr><td rowspan="2">Train collision with obstacle</td><td>LC accidents, including accidents involving pedestrians at LC</td><td>Train collision with obstacle</td><td>Level Crossing accidents</td><td>Train collisions with other Rolling Stock</td></tr>
<tr><td rowspan="3">Accidents to persons caused by RS in motion - excluding suicides</td><td rowspan="4">Level Crossing accidents</td><td>Employee/passenger accidents</td><td>Train collision with infrastructure</td></tr>
<tr><td>Individual hit by train</td><td>Trespassers accidents</td><td>Person hit by train</td></tr>
<tr><td>Individual falling from train</td><td>Accidents involving track units</td><td>Train collision with road vehicle</td></tr>
<tr><td>Fire in rolling stock</td><td>Fire in rolling stock</td><td>Fire/ Explosions</td><td>Level crossing collision with road vehicle</td></tr>
<tr><td>Electrocution by overhead line or third rail</td><td rowspan="2">Other types of accidents</td><td rowspan="2">Other types of accidents</td><td rowspan="2">Other types of accidents</td><td rowspan="2">Level crossing hit person</td></tr>
<tr><td>Accident involving dangerous goods</td></tr>
</table>

*ERA and UIC databases refer to significant accidents, FRA, CTSB and ATSB to the reportable accidents as defined in Section 3.2
** CTSB data distinguish derailments and collisions on main and non-main track occurrences

Figure 3-6 shows the European findings based on the ERA and UIC databases. It is clear that in Europe on average 77.3% of the 20,503 occurrences (European Railway Agency, 2012a) involve accidents to persons caused by rolling stock in motion and accidents at level crossings. The former, excluding the cases of suicides, refer to all the accidents caused by unauthorised individuals, i.e. trespassers, who occupy rail tracks or premises and interfere with the railway operations. The accidents at level crossings, on the other hand, involve all level crossing users, i.e. pedestrians and road vehicles occupants. Based on the ERA data, level crossing accidents have consistently decreased in the last five years. This is due to relevant campaigns and safety programs within the EU Member States (European Railway Agency, 2012b). Finally, train derailments, train collisions and fire in rolling stock although proportionally low cannot be neglected. 1,542, 1,519 and 560 accidents respectively were recorded with more than 80 fatalities (European Railway Agency, 2012a).

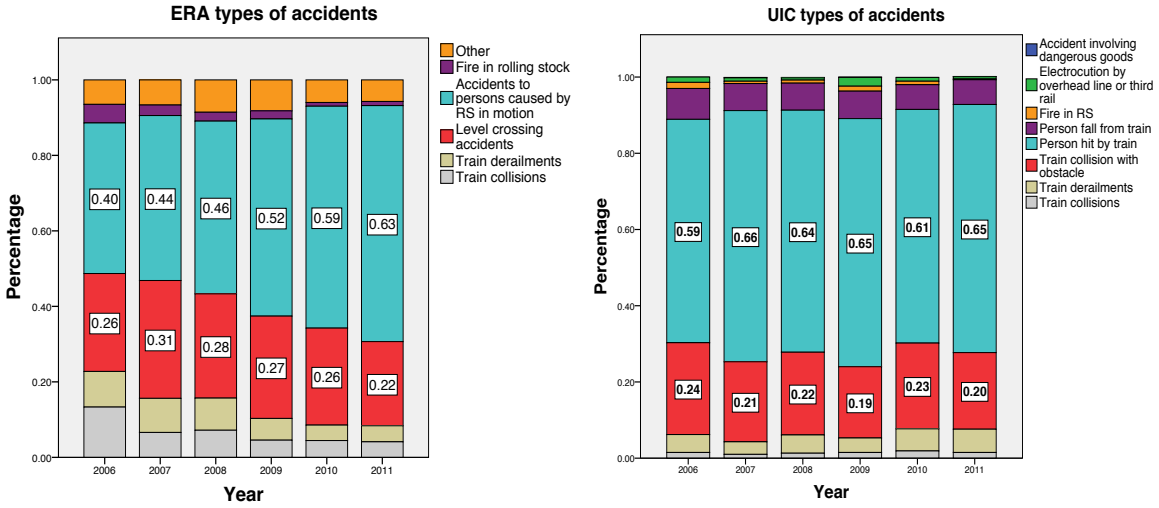



Figure 3-6 Types of accidents in Europe based on the ERA (left) and UIC (right) databases

From the ERA (2012b) data, the total reported number of fatalities for the period 2006-2011 was 8,314, 61.7% and 27.8% of which involved trespassers and level crossing users respectively. Passengers and railway personnel account for 4.5% and 2.6% respectively. Both passengers and railway personnel fatalities are usually caused by accidents to persons by rolling stock in motion and train collisions (European Railway Agency, 2012c).

The UIC data (2007, 2008, 2009, 2010, 2011b, 2012) show that train collisions with obstacles and persons hit by train account for more than 80% of all 11,133 accidents. Third parties were involved in 93% of the total of 7,580 fatalities, while passengers and railway personnel account for 4% and 3% respectively. Regarding the 6,391 serious injuries on the other hand, figures differ with approximately 71% accounting for third parties, 20% for passengers and 9% for railway personnel.

In contrast to the European findings, the Canadian (2012) data in Figure 3-7 (left) show that the majority of 7,012 train accidents are related to train derailments. As presented in Figure 3-7, train derailments account on average for 58% of all railway accidents, while the level crossing accidents represent about 17% of the occurrences. Finally, train collisions and trespassers accidents follow with 8.5% and 7.5% respectively.

Australian data (2012) in Figure 3-7 (right), on the other hand, show that train derailments and collisions with infrastructure, i.e. buffer stops, represent about 65%, of the total 2,447 accidents. Again, 26% of the occurrences involve events at level crossings and individuals hit by trains (trespassers and passengers). Finally, train collisions amounted to less than 10% of all events, with more than 9% taking place on non-main tracks, e.g. while trains are preforming shunting manoeuvres.

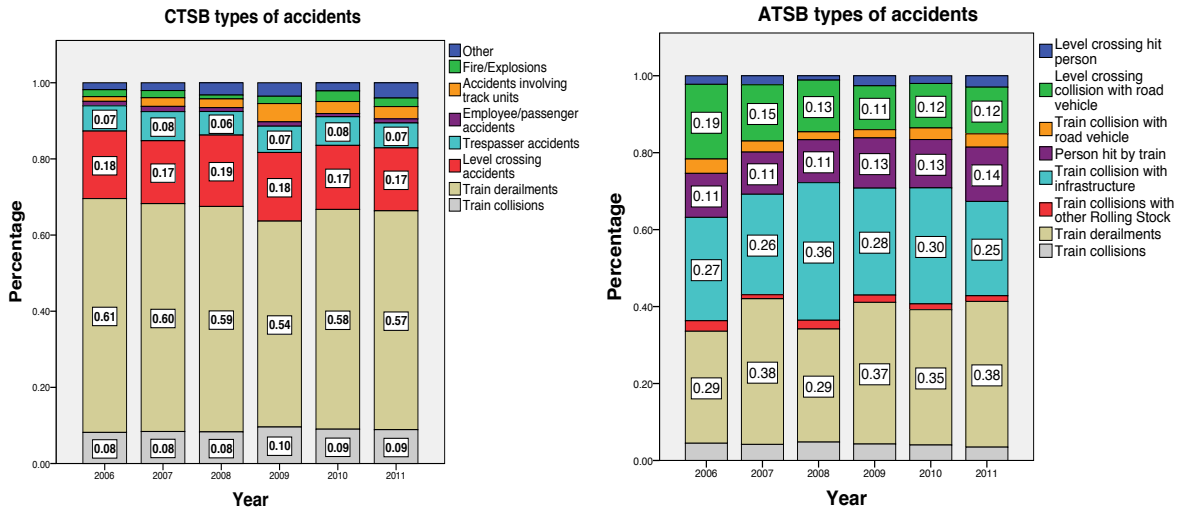



Figure 3-7 Types of railway accidents in Canada (left) and Australia (right)

Despite the differences in the proportions of the types of accidents compared to the European data, similar findings are obtained for fatalities and serious injuries. Specifically in Canada, between 2006 and 2011, about 95% of 476 fatalities involved trespassers (accounting for 63.66%) and level crossing users. Only 2.5% were associated with railway staff and less than 1% with passengers. In addition, from 356 reported serious injuries more than 95% were related to level crossing and trespassers accidents. 78.1% of the serious injuries involved third parties, 18.8% railway employees and 3% passengers.

The ATSB data do not provide explicit information on the type of persons killed or seriously injured.

Finally, from the FRA (2012) data, level crossing and other accidents account for 85% of the total 72,147 occurrences. These accidents are usually caused by trespassers and level

crossings users. The contribution of train derailments is also a crucial factor for US railway safety, explaining 14% of the events. Additionally, level crossing and other accidents explain 98.7% of 4,687 fatalities, while they are also responsible for 98% of 52,272 injuries. In contrast, train collisions and derailments are responsible for 1.3% and 2% of the fatalities and injuries respectively. Finally, 57.8% of 4,687 fatalities were due to trespassers, less than 1% to passengers and about 2.4% to employees on duty. The remaining approximately 38.8% represents level crossings user fatalities. On the other hand, trespassers are responsible only for 4% of the total of injuries, while the corresponding proportion for employees on duty and passengers were up to 52.4% and 15.12% respectively. A summary of fatalities and injuries ratio per type of person is illustrated in the Table 3-9.

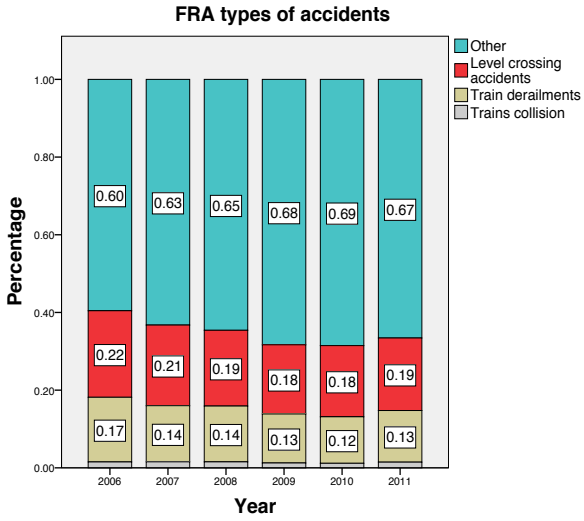



Figure 3-8 Types of railway accidents in the US

Table 3-9 Proportion of the total fatalities and serious injuries per person type

| Period 2006-2011 | | ERA | UIC | FRA | CTSB | ATSB |
|---|---|---|---|---|---|---|
| Fatalities (%) | Employees | 2.59 | 2.81 | 2.48 | 2.31 | n.a. |
| | Passengers | 4.57 | 3.78 | 0.91 | 0.84 | n.a. |
| | Level crossing - Trespassers | 89.59 | 93.4 | 96.59 | 95.59 | n.a. |
| Serious injuries(%) – Injuries for FRA | Employees | 7.32 | 9.31 | 52.44 | 18.82 | n.a. |
| | Passengers | 21.39 | 19.59 | 15.12 | 2.81 | n.a. |
| | Level crossing - Trespassers | 59.29 | 71.1 | 32.42 | 78.09 | n.a. |

Comparing the results from the five databases, two main observations can be made. Firstly, the number and type of people who either die or get injured due to railway occurrences are

related to the characteristics of the network in terms of usage. Secondly, regardless of the network characteristics, most fatalities involve trespassers and level crossing users.

Concerning the first observation, it is acknowledged that the European railway network is used to an equal extent to transport passengers and goods. In the EU Members States, Norway and Switzerland, about 406.75 billion passenger-km and 308.62 billion tonnes-km were moved in 2010 (UIC (2011a)). In contrast the US, Canadian and Australian railway networks are used primarily to transport goods. Therefore, it is not surprising that in the US only 6.518 billion passenger-km were registered in 2010, while the corresponding value for freight operations rose to 2,468.738 billion tonnes-km. The corresponding figures for Canada are 1.346 billion passenger-km and 322.741 billion tonnes-km. Finally, the relevant figures for the Australian railway network were 1.5 billion passenger-km and 64.172 billion tonnes-km. Therefore, while in Europe about 75% of accidents (International Union of Railways, 2012) involve passengers trains, the opposite is the case for the US, Canada and Australia. Subsequently, the number of passenger fatalities in Europe, which are primarily caused by train collisions or train collisions with obstacles, is six times higher compared to the US. Likewise, the number of passengers seriously injured in Europe is considerably higher than in Canada. However, similar findings cannot be extracted with confidence for the US railways, since the FRA does not distinguish the types of injuries.

With regard to the second observation, it is reasonable to claim that trespassers fatalities constitute the great majority of total fatalities in railway operations. Trespassing actions are more common amongst young males, while they are also strongly associated with vandalism acts (see also Section 2.2.3.1.2) (Forsdike et al., 2007). For example, the FRA (2008) identifies the average trespasser *"as a 38-year-old Caucasian male under the influence of alcohol and/or drugs, with a median household income of $36,000, while more than 25 percent of these individuals did not graduate from high school"*. Therefore, it could be argued that the type of neighbourhood could be used as precursor for the cases of trespassing and for vandalism actions on railway premises. Clearly, research is required to confirm such an argument. In addition, especially for the dense European railways proximate to residential areas, trespassing actions are common with individuals who either attempt to take a shortcut or walk very close to the railway tracks and subsequently are hit by moving trains.

Secondly, concerning the level crossing accidents, it is acknowledged that railway networks are usually sited close to road networks, and in some cases close to residential areas. Therefore, a large number of level crossings exist along the tracks to accommodate the transfer of car vehicle users and pedestrians. Particularly in Europe, circa 120,000 level crossings, or five per 10 line-km, are operational currently (European Railway Agency,

2012b). Due to the fact that only 38% of level crossings are equipped with automatic protection systems, level crossing users are vulnerable to accidents. However, it should be noted that although the proportion of level crossing accidents of the total accidents in the European countries is similar across the Member States (European Railway Agency, 2012c), not all the countries record the same number of fatalities. This is explained by the different number of level crossings in each country. Similarly, the proportionally large number of level crossings compared to the train line miles in the US and Canada (about 210,000 (Association of American Railroads, 2013), and 27,600 (Railway Association of Canada, 2011) respectively together with the fact that many are located in rural areas, increase the number of accidents due to inattention of their users.

Although the consequences of trespassing and level crossing accidents primarily affect trespassers and level crossing users, the literature shows that train drivers who have met such accidents are likely to face post-traumatic stress disorder. Subsequently, this may affect their abilities to conduct their duties in future (Briem et al., 2004, Lunt and Hartley, 2004). Therefore, railway organisations worldwide run campaigns to inform and educate people about the dangers of trespassing and careless use of level crossing. It is also indisputable that both level crossing (Rail Safety and Standards Board, 2004c) and trespassing accidents (definition in Section 2.2.3.1.2) are caused primarily by third parties. Hence, neither type of accidents constitutes an occurrence caused by the railway operators. Subsequently, it is out of the scope of this research to investigate them further. Hence, the rest of the thesis is focused on the remaining types of accidents.

### 3.3.3 Contribution of railway constituents to safety

Having identified the types of accidents and their proportion of the total number of accidents, this section investigates the contribution of each railway component to railway occurrences.

European data from the UIC (2007, 2008, 2009, 2010, 2011b, 2012) databases (Figure 3-9) show, that for the past six years, third parties were responsible on average for about 76% of the train accidents. The contributions of the "*controlled*" components of the system were associated with human factors at 17% and railway sub-systems at 4.33%, while the remaining 2.67% was not defined.

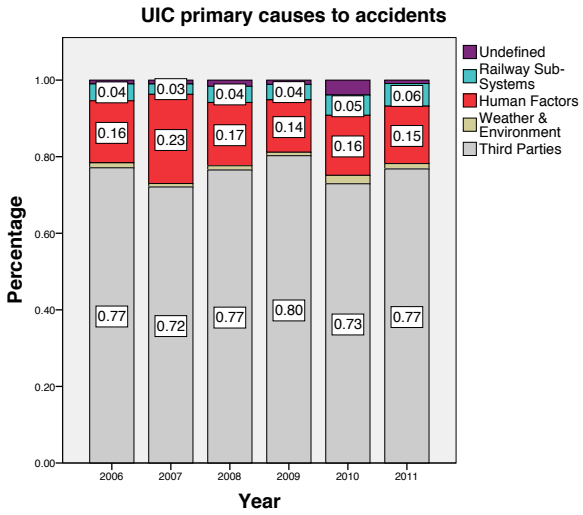



Figure 3-9 The UIC primary causes to accidents

A further analysis of the railway sub-systems, Figure 3-10 (left), indicates that infrastructure and rolling stock inconsistencies cause most accidents. It also shows that within the infrastructure the most common causes are broken tracks and wrong-side signalling failures. The former is associated with train derailments, and the latter with both train collisions and derailments. The most usual rolling stock inconsistencies, on the other hand, are associated with broken axles and broken wheels, which may both lead to train derailments.

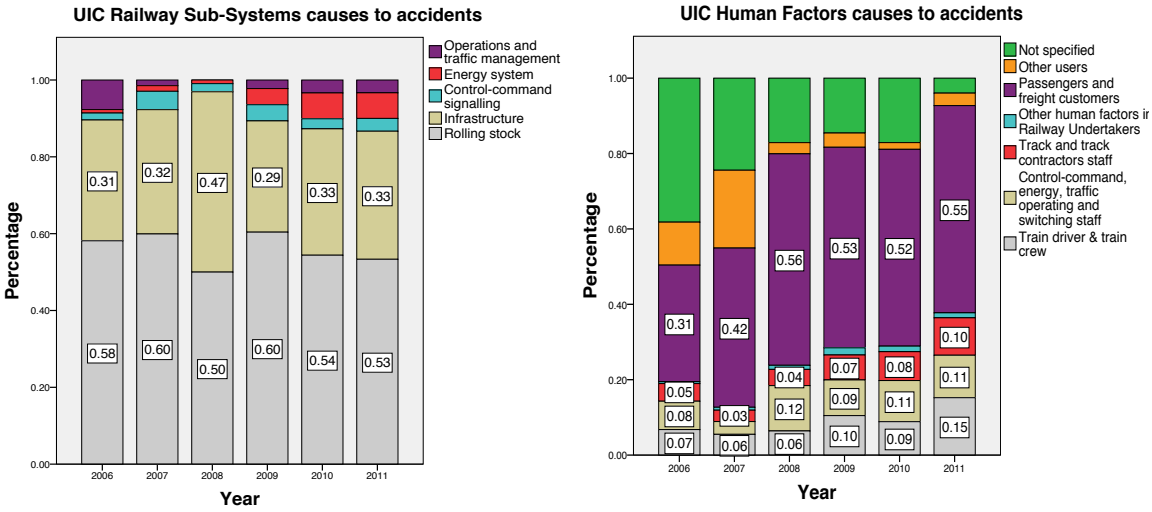



Figure 3-10 The UIC second level causes to accidents

The causes of human factors are shown in Figure 3-10 (right) with passengers and freight customers contributing most to accidents. However, these types of accidents are considered as *third party* accidents, and are therefore outside the scope of this thesis. Figure 3-10 (right) shows that operators, i.e. train drivers and signallers, have a significant effect on the safety of the railway system. For the period from 2006 to 2011, operators contributed on

average to 18% of human factors accidents and to approximately 3% of total accidents, resulting in more than 330 accidents. In addition, the literature shows that the largest number of serious accidents leading to a high number of fatalities and financial losses, have been attributed to operators (Dhillon, 2007).

Moreover, the ERA (2011) data (Figure 3-11) indicates that SPADs and broken rails are the most common precursors to accidents. The data shows that the accidents caused by broken rails although resulting in large financial losses are not extensive with respect to human losses. This is in contrast to accidents caused by SPADs (e.g. the Ladbroke Grove accident in the UK in 1999, the Chatsworth trains collision in Los Angeles in 2008 and the train crash in Hordorf in Germany in 2011) with extensive human losses. In particular, the Ladbroke Grove accident initiated a debate about the role of human performance and human factors in railway safety. Rail organisations have started to investigate and examine in detail the contribution of railway staff to the safety of the system. Many studies have been published (Wilson et al., 2005, Wilson and Norris, 2005, Wilson and Norris, 2006, Wilson et al., 2007a, Wilson et al., 2007b), in addition to the establishment of many Human Factors groups including the RSSB and Network Rail HFs.

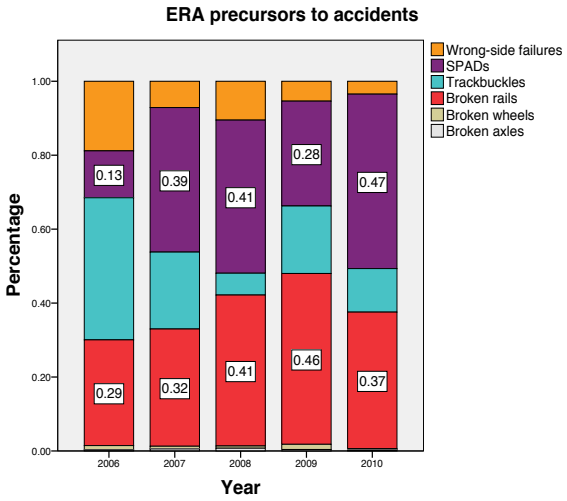



Figure 3-11 Precursors to the EU train accidents (European Railway Agency, 2011)

In addition, FRA (2012) data (Figure 3-12) show that in the past six years human factors consistently contributed to 33% of the total train accidents. Again, train drivers and signallers were identified as the operators with the most significant impact on safety. In the case of the drivers, failure to comply with restricted speed and failure to control train over shoving (shunting) movements are the most representative examples. With respect to signallers, on the other hand, authorisation for trains to enter an already occupied line, and the change of switches before a train runs through lines, have been reported as the most typical errors.

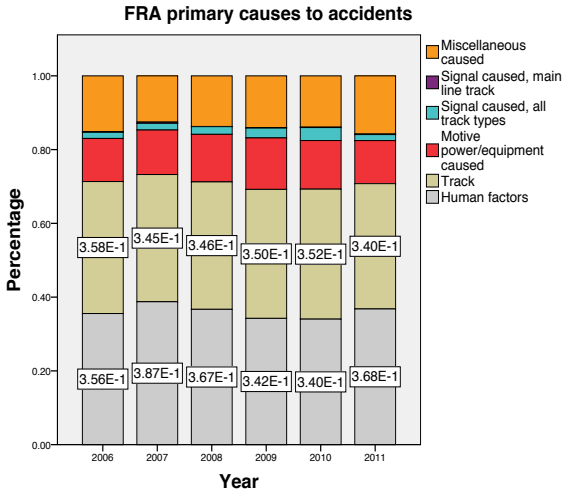



Figure 3-12 Primary causes to US train accidents (Federal Railroad Administration, 2012)

From the CTSB (2012) data, it was shown in the previous section that train derailments constitute the majority of train accidents. It is clearly shown in Figure 3-13 (left), that human actions have caused most of the derailments. Amongst the actions the majority is related to operator incompetence specifically in protecting and securing the railways, as well as maintaining the required speed and using the on-board equipment (Transportation Safety Board of Canada, 2012). Even more profound are the results with respect to train collisions, Figure 3-13 (right). Again, data shows that operator incompetence in protecting and securing the railways, as well as using the on-board equipment is responsible for nearly all the identified actions. (Transportation Safety Board of Canada, 2012).

The analysis so far has focused on OECD countries. However, similar data from India (representing developing countries) show similar trends. In particular, the data show that approximately 59% of the total railway accidents for the period 1996-2006 were attributed to failures of railway personnel. On the other hand, 25.5% of the occurrences were caused by other than railway employees' failures and 6.5% by equipment failures (Agarwal, 2006).

Based on the findings presented in this section, it is clear that further analysis of the impact of railway operators on safety of railway operations is required. Such an analysis will support railway organisations to direct resources more efficiently towards the development of smart and effective solutions to reduce operator errors.

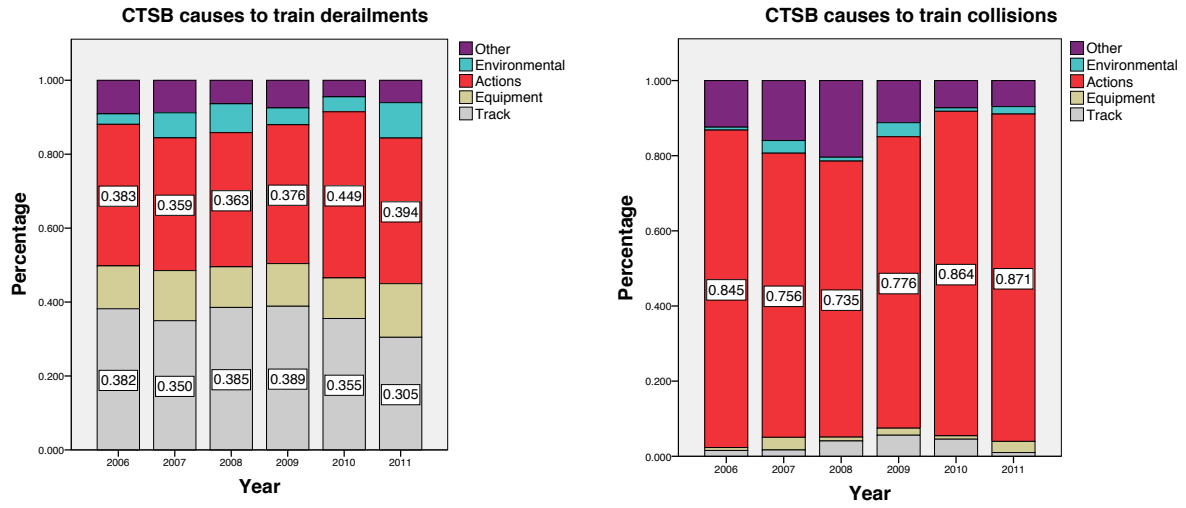



Figure 3-13 Canadian causes to train derailments (left) and collisions (right) (Transportation Safety Board of Canada, 2012)

## 3.4 Summary

This chapter summarised the history of the development of railway safety in Europe, US, Canada and Australia and explores the types and causes of accidents. To this end, firstly a high level analysis of the recorded injuries and fatalities was undertaken, from which the categories of people who are most affected by incidents and accidents associated with railway operations were determined. Subsequently, the causes of these were scrutinised in greater detail. Overall, the analysis showed that the *uncontrolled* components, i.e., third parties comprising trespassers or level crossings users, are involved in the majority of railway accidents. Although their contribution to the railway safety overall is significant, occurrences caused by such *uncontrolled* components are violations external to the system. As the focus of this thesis is on *controlled* components, occurrences, associated with *uncontrolled* components are henceforth not considered further.

With respect to the operational concept of the controlled components of the railway system it was shown that human actions are responsible for nearly 80% of the total accidents attributed to the *controlled* components of the system. Out of this 80%, more than 60% of the occurrences refer to railway personnel, and in particular to train drivers, signallers and controllers. It was shown from the analysis that most of the human failures are associated with signals passed at danger (SPADs), non-compliance with speed restrictions, failure to secure operations, wrong authorisation to trains and change of switches before trains run through.

In addition to the findings from the data analysis, the literature (Andersen, 1999, Rail Safety and Standards Board, 2004d, Dhillon, 2007, Wilson et al., 2007a, Evans, 2011) confirmed that most of human failures occur while trains move from one point to another. The most common failures are SPADs, ignoring train speed limits and signalling or dispatching inconsistencies (no technical failures included), confirming that train drivers, signallers and controllers have a significant impact on the network in terms of safety (Dhillon, 2007).

Therefore, a critical issue within the railway community is to identify and explore the reasons that lead operators to commit errors. These are described in detail and analysed in the remainder of this thesis, starting in the next chapter with a discussion of the types of human errors and the methodologies employed to understand their origins, capture and quantify their impacts.

# Chapter 4 HUMAN ERROR - HUMAN PERFORMANCE ANALYSIS TECHNIQUES

Chapters 2 and 3 identified human factors as having a significant role to play in the safety of railway operations. This chapter reviews the existing techniques for the analysis of human performance. This chapter first presents the concept of the railway operational error. Subsequently, the different types of human errors are classified, and the models used to identify the origins of human errors are described. After having identified the origins of human errors, the relevant terms of human factors, human reliability and human performance are then introduced. This is followed with a critical review of the most common human performance analysis techniques employed in the railway and cross-industry domains. This review identifies the strengths and weaknesses of the existing techniques and defines requirements of a technique suitable for railway operations.

## 4.1 Concept of operational error in railways

Based on the Railway Operational System Architecture (ROSA), the framework used to classify the tasks of train drivers, signallers and controllers can be divided into three phases: (i) receiving and assessing information, (ii) making decisions and (iii) executing actions. Each phase is prone to errors, which are referred to as operational errors. The concept of railway operational errors is similar to what Wickens (1992) describes as the information processing stage, which is a cognitive process consisting of the three fundamental phases of input, processing and output. In the first phase, also captured in the ROSA, operators receive information via sensors, usually in the form of visual or auditory stimulus, such as signal indications. In this phase operators detect, perceive and acknowledge information. In the second phase the perceived information is processed. This entails interpretation, judgement and decision making regarding the most suitable cognitive/behavioural response/action. During this phase recall or recognition of related stored information within working or long-term memories may also be required. Finally, in the output phase, the selected response is applied according to a predetermined plan. Within the context of information processing stages, all phases are associated with attention resources, which indicate the performance of operators. Hence, it is usually assumed that the higher the attention resources (performance), the safer the execution of the phases.

Figure 4-1 illustrates an example of the process in its entirety. The example refers to a train driver, who receives a signal at danger and should act according to a predetermined plan. As can be seen, the input phase comprises the sensory processing and perception, that is, the driver detects the signal, reads and acknowledges its indication. During the second phase the driver processes the information, deciding to stop at the signal. Finally, in the output phase, the decision to stop at the signal is executed by applying the brakes. In addition, Figure 4-1 shows how attention resources and memory, both working and long-term, support the three phases.

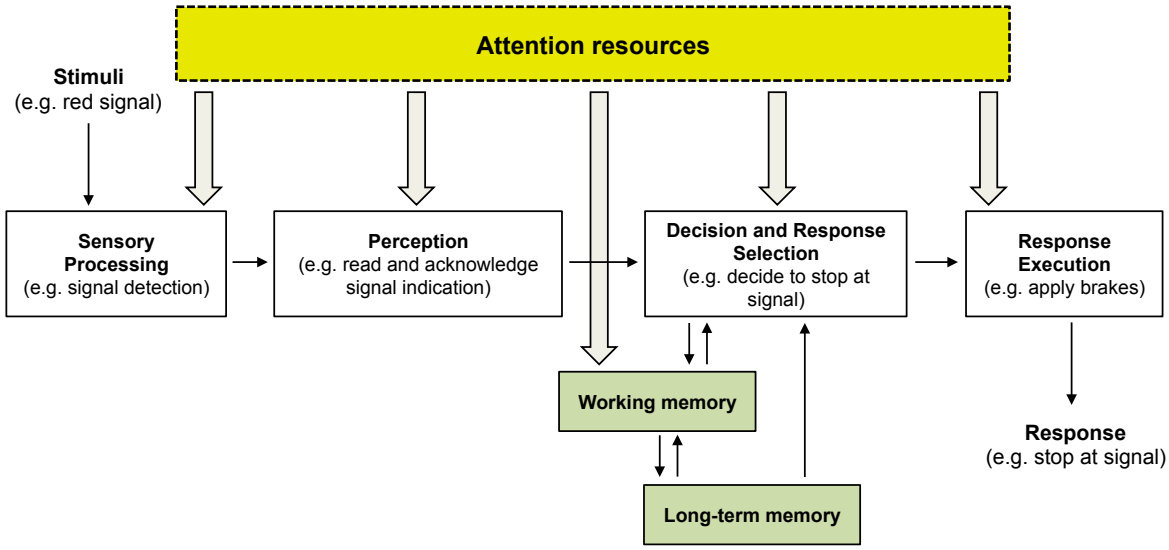



Figure 4-1 Process of operational error in railway operations by (Wickens, 1992) as cited in (Rail Safety and Standards Board, 2004b)

However, each of the three phases is prone to errors. For example, at the input stage the train driver may not detect the signal, while at the processing phase the signal may be misinterpreted. Finally, in the output stage, the train driver may not apply the brakes appropriately, resulting in a signal passed at danger. The different types of relevant human errors and their origins, within the operational concept of railways, are reviewed and presented in the next section.

## 4.2 Definitions of human error

Over the last thirty years a lot of progress has been achieved in understanding human errors (Health and Safety Executive, 1999). According to Leveson (2004, p.245), the term human error usually describes "*any deviation from the performance of a specified or prescribed sequence of actions*".

A psychological definition, given by Reason (1990, p.9) indicates human error as "*a generic term to encompass all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some change agency"*. Matthews et al. (2000, p.147), on the other hand, define human error from the perspective of a system as "*any inappropriate or undesirable human decision or behaviour that violates one or more system tolerance limits"*.

A more comprehensive definition given by Swain and Guttmann (1983, p.2-7) and adopted later by EUROCONTROL (2007, p.1) is that human error is "*any member of a set of human actions or activities that exceeds some limit of acceptability, i.e. an out of tolerance action (or failure to act) where the limits of performance are defined by the system"*. Finally Gertman et al. (2004, p.xxvii), determine human error as *"an out-of-tolerance action, or deviation from the norm, where the limits of acceptable performance are defined by the system. These situations can arise from problems in sequencing, timing, knowledge, interfaces, procedures, and other sources"*.

Although a variety of definitions exist, it can be concluded that human error is generally associated with failure to deliver the expected outcome, deviations from norms and violation of a system's limits of tolerance. Nevertheless, such variety confirms Hollnagel's (1998, p.23) earlier observation that "*practically everyone who has published work on the topic of human fallibility in a work context has attempted, at one time or another, to provide a technical definition of the concept of "error". The net result of this effort has been a clear demonstration that it is* ***extremely difficult*** *to provide a precise definition, despite the fact that the term "error" has a relatively simplistic meaning in everyday life*".

Due to the lack of a commonly agreed definition of human error, this thesis introduces two new definitions. Both are developed based on existing definitions, as described above, and define human error more narrowly from the perspective of railway personnel and railway operations[3]. In the context of railway personnel, human error is "*any inappropriate or undesirable action or behaviour from railway personnel, which compromises the safe operation of the system and jeopardises the system's limits of acceptability, where the limits of performance are defined by the system and regulations".* Based on the definition for railway operations, the railway operational human error is defined as "*any inappropriate or undesirable action or behaviour of a train driver, signaller or controller, individually or as team, which compromises the safe operation of the railway and increases the probability of*

[3] "*any train movement from one point to another including shunting operations"*

*an incident/accident"*. This definition of railway operational error is used throughout the remainder of this thesis.

## 4.3 Classifications of human error

Although *errare humanum est ("*to err is human"*)*, not all humans or all errors are the same. The literature indicates different ways to classify human errors, of which the phenomenological and cognitive approaches are considered to be predominant (Matthews et al., 2000).

### 4.3.1 The phenomenological approach

The phenomenological approach separates human errors into *omission* and *commission* errors. According to this theory, errors are described as they are observed, i.e. actions have either not been performed or have been performed inadequately (Meister, 1964, Swain and Guttmann, 1983, Matthews et al., 2000). In line with this, omission errors define actions which, although they should have been performed, have not been carried out, e.g. the failure of a train driver to test the braking system before starting the service. Errors of commission, on the other hand, comprise actions which have been executed poorly, including actions performed imprecisely, at the wrong time or in the wrong sequence (Kirwan, 1994), e.g. a signaller changes points before train passes.

Furthermore, commission errors are sub-divided into *extraneous acts*, *sequential errors* and *time errors* (Matthews et al., 2000). Extraneous acts refer to situations where an action is likely to prevent the achievement of a particular goal, e.g. a train driver using a mobile phone while approaching a station and consequently failing to stop within the designated area. Sequential errors indicate actions executed out of sequence, e.g. a train driver departing from a station and checking the appropriate signal only after departure. Finally, time related errors are actions that were performed either too early or too late or outside a permitted period, e.g. signaller does not inform train driver about a signalling failure in time.

This phenomenological classification is a simple and straightforward scheme to distinguish human errors. However, due to its simplicity, the model focuses only on the outcomes of an error rather than on its causes. Hence, it does not provide any insights for probable underlying causes that may have contributed to the error, such as weaknesses in the system design, inadequate supervision, procedures or training. Subsequently, the available

information on the causes of error is inadequate for developing efficient remedial measures (Kariuki, 2007).

### 4.3.2 The cognitive approach

The main feature of cognitive approach, which is very similar to the operational concept of human errors in railways, is the assumption that errors are associated with the stages of cognition in conceiving and executing a task sequence (Reason, 1990). According to Reason (1990) there are three cognition stages in a task sequence. These are: (i) planning, which refers to both the identification of a goal and to the sequence of the selected actions (plan) to achieve that goal, (ii) storage, which refers to the storage in memory of a selected plan until the time it is executed and (iii) execution, which refers to the implementation of the stored plan by carrying out the selected actions. Each of the cognitive stages is further associated with a basic type of error, known as slips, lapses and mistakes, as shown in Table 4-1.

Table 4-1 Classification of error types according to cognitive stages

| Cognitive stage | Basic type of error |
|---|---|
| Planning | Mistakes |
| Storage | Lapses |
| Execution | Slips |

Slips and lapses are described as "*errors which result from some failure in the execution and/or storage stage of an action sequence, regardless of whether or not the plan which guided them was adequate to achieve its objective*" (Reason, 1990, p.9). Specifically, slips which occur during the execution stage are described as *actions not as planned* errors (Reason, 1990, Health and Safety Executive, 1999). They indicate errors where the intention is correct but the execution of the required action is false, e.g. a signaller activating the wrong switch. Lapses, on the other hand, are errors associated with the memory storage phase, primarily characterised by missed actions and omissions. They are referred to also as *forgetting state* errors (Reason, 1990, Matthews et al., 2000, Kariuki, 2007) and happen when an individual fails to perform an action because of inattention and/or memory failure, e.g. a train driver forgetting to check the signal before departing the station.

Finally, mistakes, which occur in the planning stage, are defined as any "*deficiencies or failures in the judgemental and/or inferential process involved in the selection of an objective or in the specification of the means to achieve it, irrespective of whether or not the actions*

*directed by this decision-scheme run according to plan*" (Reason, 1990, p.9). By their nature, mistakes are not only more dangerous than slips and lapses, but also more difficult to detect. Mistakes occur when people wrongly execute a task in the belief of doing it correctly. They are related to individuals' mental capabilities to plan, assess available information, make and/or take decisions and evaluate consequences. For example, applying a rule in a situation where it is not appropriate, because it is often used and seems appropriate for the current situation (Rail Safety and Standards Board, 2011).

Although the cognitive classification provides a constructive way to distinguish human errors, it is not always clear and straightforward to identify whether an error is to be classified as a slip or a mistake. For instance, there are cases of erroneous behaviour that contain elements of mistakes and slips, when operators execute wrongly a selected task based on inadequate assessment of the state of the system (Reason, 1990). Consequently, the final error classification leads to inconclusive and misleading findings, which provide insufficient feedback for remedial measures. Additionally, similar to the phenomenological approach there is no information causes of error.

### 4.3.3 Summary of errors classifications

The phenomenological and cognition human error classifications have been widely used to identify the different types of human errors. Both classifications focus on analysing errors from the perspective of an individual. Their findings are used to (i) distinguish and better understand the different types of human error and (ii) statistically analyse human error occurrence in a variety of contexts. Table 4-2 provides a summary of the concept, strengths and weaknesses of both approaches.

Table 4-2 The human error classifications

| Models | Concept | Pros | Cons |
|---|---|---|---|
| Phenomenological | ▪ Human errors are described just as they are observed, i.e. omission & commission | ▪ Simple and straightforward classification to distinguish human error | ▪ Focus on the outcome rather than the cause of errors<br>▪ No insights concerning any underlying factors that contribute to an error |
| Cognitive | ▪ Human errors are related with the different cognitive stages in conceiving and executing an action | ▪ Constructive approach to classify human error<br>▪ Associates human error with cognition stages | ▪ Not always clear how to distinguish slips and mistakes<br>▪ Insufficient feedback for remedial measures<br>▪ No insights concerning any underlying factors that contribute to an error |

As shown in Table 4-2, the main drawback of both classifications is that they do not account for any underlying or hidden contextual factors that may have affected the performance of frontline personnel and subsequently contributed to the error's occurrence. Therefore, although they explain the types of human errors as the final outcome of a broader context, they do not provide any information about the origins of the errors within a specific context. To address this limitation, human error models were developed and implemented. These include Heinrich's domino theory (1931), Reason's generic error modelling system and Swiss cheese (1990) and more recently Leveson's STAMP (2004) and Hollnagel's FRAM (2012).

## 4.4 Models of human error

The contribution of human error classifications is substantial in the identification of the nature of human errors. However, they do not provide conclusive information or capture the overall picture of why and under which circumstances an error occurs. Since the majority of serious accidents in railways are caused by human errors, it is important to seek means to identify and define the causes of such errors. Reason (2000) proposed two distinct ways to consider human errors, the person and system approaches. Each has its models of error causation and provides different theories of error management. Both approaches have been widely used in the process of understanding human error causations in the railway industry.

### 4.4.1 Person approach

The person approach focuses on the *unsafe acts*[4] of frontline personnel, primarily caused by aberrant mental processes such as distraction, loss of concentration, memory lapses, poor motivation or decision making skills (Reason, 2000). It is expressed by the Generic Error Modelling System (GEMS) created by Reason (1990), based on the Skill-Rule-Knowledge (Rasmussen, 1982) classification of human performance and the cognitive error mechanisms (Reason, 1990).

[4] According to Reason (1990) "u*nsafe acts*" are perpetrated by system operators and are directly linked to an accident/incident.

#### 4.4.1.1 The Generic Error Modelling System (GEMS)

The GEMS is a conceptualised framework used to identify the origins of human error types, which in large part stems from the Skill-Rule-Knowledge (SRK) classification of human behavior introduced by Rasmussen (1982).

The SRK framework is the first of its kind that successfully combined psychology and engineering into a single approach. The SRK refers to the degree of conscious behavior of an individual during the execution of their tasks. It comprises three hierarchical levels, as illustrated in Figure 4-2, each representing a discrete level of behavior.

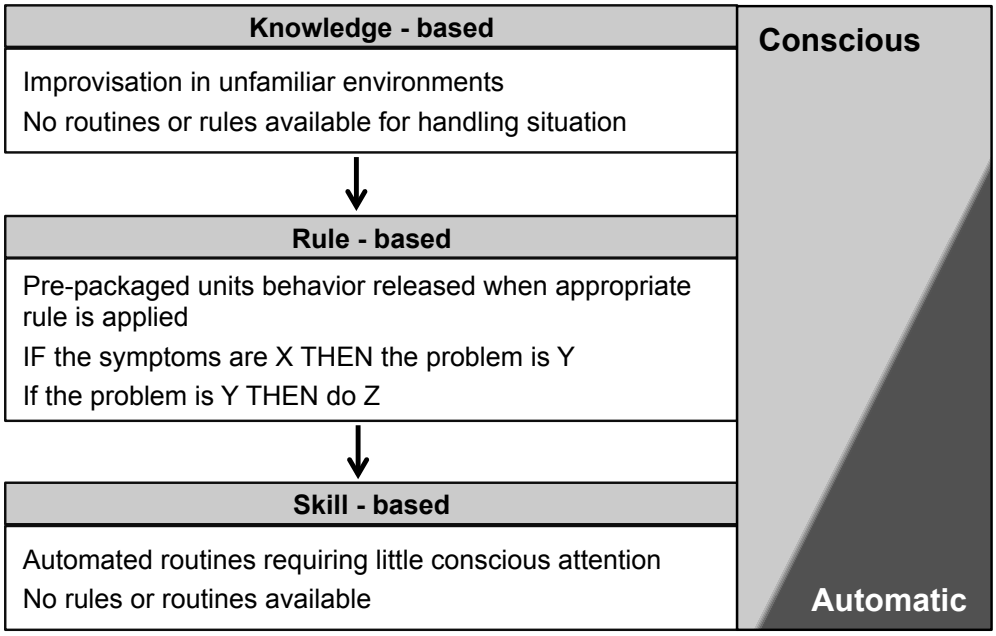



Figure 4-2 The continuum between conscious and automatic behaviour (Reason, 1990)

The lowest level, skill-based, refers to almost unconscious (automated) behavior, where well-practiced actions are instinctively executed. The intermediate level, rule-based, indicates performance controlled by stored rules. At this level, both the acknowledgment of the necessity to apply the rules and the cognisance of selecting and implementing the appropriate rules are entailed. Finally, the knowledge-based mode refers to tasks that for their execution require a highly conscious state of mind. This type of behavior typically occurs either under abnormal conditions or during a learning process and therefore, demands a thorough understanding of the functions of the system and its characteristics. (Reason, 1990, Kirwan, 1994, Matthews et al., 2000, DiMattia, 2004, Kariuki, 2007, Kyriakidis, 2009).

With the integration of the SRK based framework and the cognitive error mechanisms (see Section 4.3.2), Reason (1990) generated the Generic Error Modelling System (GEMS). The

GEMS allows analysts to: (i) understand the origins of human errors, (ii) comprehensively understand the nature of errors, (iii) expect and predict the time and conditions in which a specific type of error may occur, (iv) understand better the transition from the skill- to knowledge-based behavior and (v) assess the cognitive determinants within *complex sociotechnical systems*[5] (Kirwan, 1994).

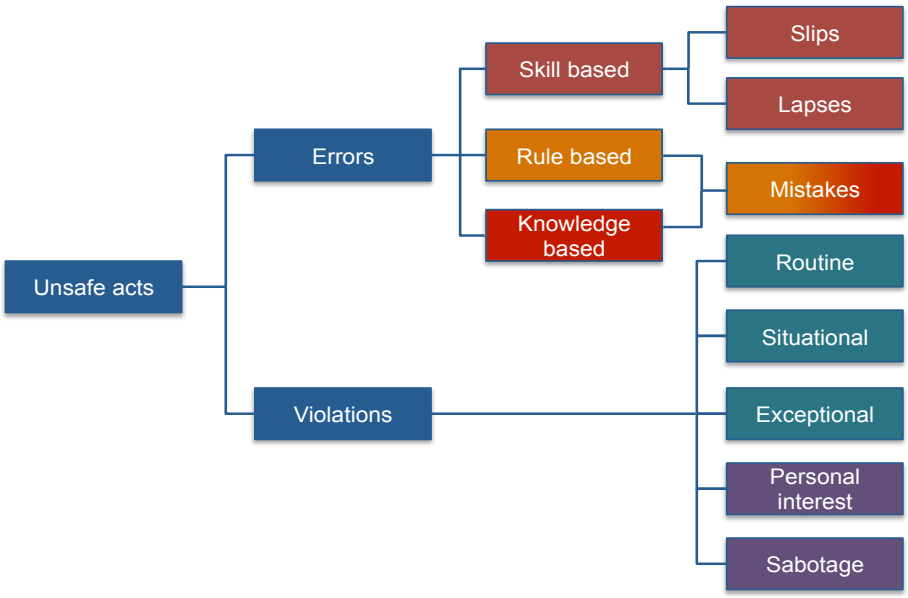



Figure 4-3 Unsafe acts framework (Health and Safety Executive, 1999, Rail Safety and Standards Board, 2011)

The GEMS focuses on a person's *unsafe acts*, which are classified as: errors and violations. Within the GEMS concept, the former generally refer to any "*unintentional deviation from the accepted standards*", while the latter to any "*deliberate deviations from rules, procedures, instructions and regulations*" (Health and Safety Executive, 1999). Moreover, it considers human errors as the combination of behavioural and cognitive errors. Therefore, skill-based errors are associated with slips and lapses. Such errors happen when the tasks are very familiar, their execution is almost automatic (reflexive) and therefore, do not require a lot of attention. On the other hand, mistakes are associated with the rule- and knowledge-based errors. Rule-based mistakes happen when the behaviour of an operator is dependent on either memorised rules or their familiarity with procedures. Hence, there is strong tendency to either misuse well-defined rules or implement inappropriate procedures, e.g. authorising a train driver to pass a signal at danger. By contrast, knowledge-based mistakes arise under unfamiliar or unusual conditions, when no rules or procedures exist to support and handle

[5] A complex sociotechnical system is defined as "*the complex interaction between humans, machines and the environmental aspects of the work system*" (Vanderhaegen, 2001, Cacciabue, 2004b). Nowadays this interaction is true of most enterprise systems, including railways.

that specific situation, e.g. coping with a major signalling failure (Reason, 1990, Kirwan, 1994, Health and Safety Executive, 1999, Rail Safety and Standards Board, 2004e).

With respect to violations, the UK Health and Safety Executive (HSE) (1999) divide these into routine, situational and exceptional. The Rail Standards and Safety Board (RSSB) (2011) have later highlighted two more categories, the personal interest and sabotage violations. A routine violation refers to any rule or non-compliance with procedure or disregard that over time becomes a standard process within a work group. Such violations happen when the rules are believed to be too restrictive, unnecessary or out of date, e.g. a shunter getting on and off the engine while it is running (Rail Safety and Standards Board, 2011). In the case of situational violations, rules are not followed due to work pressure including time pressure, understaffing or adverse weather conditions. They usually occur under extraordinary conditions where the work situation makes the rules difficult or impossible to follow. For instance, a train driver continuing a shunting operation after losing visual contact with the person in charge of the movement (Health and Safety Executive, 1999, Rail Safety and Standards Board, 2011).

Exceptional violations, on the other hand, occur in new, critical or emergency situations, when it is believed that the risk of breaking a rule would be less of a risk taking than sticking to the rule during the execution of a task. For instance, a train driver exceeding speed limits due to an emergency situation in one of passengers coaches (Health and Safety Executive, 1999, Rail Safety and Standards Board, 2011). The personal interest violations include any deliberate disregard of rules that leads to a personal benefit. In these situations a task is executed in an unsafe manner because it is faster and considered beneficial to do so, e.g. driving a train above speed limits in order to return home earlier. Finally, sabotages refer to violations due to acts of vandalisms, which are executed deliberately and consciously to cause harm and damage (Health and Safety Executive, 1999, DiMattia, 2004, Rail Safety and Standards Board, 2011).

The GEMS has been broadly applied to several domains, including aviation, railways and offshore operations to identify and mitigate human error (Shappell and Wiegmann, 2000, Khan et al., 2006, Rail Safety and Standards Board, 2011). However, the model has its limitations. Firstly, human errors are seen as random and independent events, while frontline operators are considered incorrectly as the only cause of errors. However, it is often the case that the *best people* make the *worst errors* (Reason, 2000). As a result, the possibility of individuals with an overall good performance making an error should not be overlooked. In turn, investigation should consider the overall context of error's occurrence. Secondly, human errors tend to fall into repeated patterns. The same conditions can produce similar

errors regardless of the characteristics of the involved personnel, which again indicates the significance of studying the overall context of an error's occurrence. Finally, due to the fact that the models based on the cognitive approach ignore the underlying causes of errors, limited insights with regard to the origins and propagation of human error are provided. Consequently, this negatively impacts the development of remedial countermeasures (Kariuki, 2007). To address these issues, the system error causation models were derived. Such models are the most widely used for investigating the origins of human errors.

### 4.4.2 System approach

As discussed in the previous section, the person approach considers human errors as random and independent events caused exclusively by frontline personnel (Health and Safety Executive, 1999). However, it has been proven that this theory is no longer valid (Leiden et al., 2001). Analyses of railway accidents and incidents for the past 30 years, for example, show that in most of the cases frontline personnel errors do not contribute in isolation to unwanted safety outcomes. Instead, they coincide with additional contributing factors stemming from the operational environment. These factors include technical and organisational failures, such as, e.g. in the Clapham Junction (1998) accident, Ladbroke Grove (1999) trains collision in London and Chatsworth (2008) trains collision in Los Angeles (Health and Safety Executive, 1999, Evans, 2011, Federal Railroad Administration, 2012).

Therefore, the system approach to study the origins of human errors was developed. The approach is expressed by Reason's (1997, p.126, 2000, p.768) statement that "*errors should be considered as consequences rather than causes, while their origins are not found so much in the perversity of human nature as in "upstream" systemic factors*". The system approach focuses on the conditions in which individuals work with the aim to suppress accidents/incidents by establishing defences that prevent errors or mitigate their effects. Subsequently, human errors are studied as the result of a series of events that eventually lead to an accident/incident and are investigated within the broader prism of accidents causation (Reinach et al., 2007).

In the literature, a multitude of models are available to assist in this search. These models are necessary to facilitate accident/incident investigations by accounting for all the factors and conditions (including but not restricted to human errors), apparent or not. A critical review of the numerous models that have been developed and implemented across several industries shows that they all rely on one of the sequential, epidemiological and systemic theories (Hollnagel, 2004, Dupuy, 2011, Salmon et al., 2011). The three theories are in turn explained in the following sections.

#### 4.4.2.1 The sequential models

Characterised by Heinrich's (1931) domino accident causation model, sequential models explain accidents as the result of a series of discrete events, including human and component failures that occur in a particular order (Hollnagel, 2004). They are based on the cause-effect concept, where a sudden undesirable or unexpected event initiates a sequence of upcoming events, until the occurrence reaches its final state, i.e. the *accident* (Hollnagel, 2004, Qureshi, 2007, Dupuy, 2011).

According to the model, human errors can be identified, as shown in Figure 4-4, in the *social environment*, *fault of person*, and *unsafe acts*, domino pieces. The first refers to human errors at management level, while the second describes errors caused by factors such as fatigue, lack of training, competence, and attitude. The *unsafe acts*, are those errors made by the frontline personnel just before an accident occurs, e.g. train driver exceeds speed limits leading to train's derailment. The dominos are arranged in such a way that once the first in the series falls, it sets off a chain reaction, causing the entire row of upended dominos to fall. Within this context, the model provides a constructive framework to identify human errors not only as "*unsafe acts*" caused by the frontline personnel but also as errors that can arise at the higher levels of hierarchy within an organisation. Nevertheless, due to its nature, the model considers human errors as the result of *root cause* events. Thus, an error would be prevented if the single event that causes that error were removed, e.g. a tired train driver is released from duty.

A limitation of the sequential models is that it is not always clear how to determine the unique *root* event in the sequence that started the *domino effect* (Dupuy, 2011). Therefore, the models have limited power in complex systems. This is explained by the fact that in such systems human errors are considered as the result of combining multiple factors, hidden and apparent, that occur together in space and time (Qureshi, 2007). Subsequently, since the sequential models focus only on the root causes of an error, they fail to capture any other factors that may contribute to such an error. Hence, the epidemiological accident models, characterised mainly by Reason's Swiss Cheese model (1990), were introduced to address such limitations.

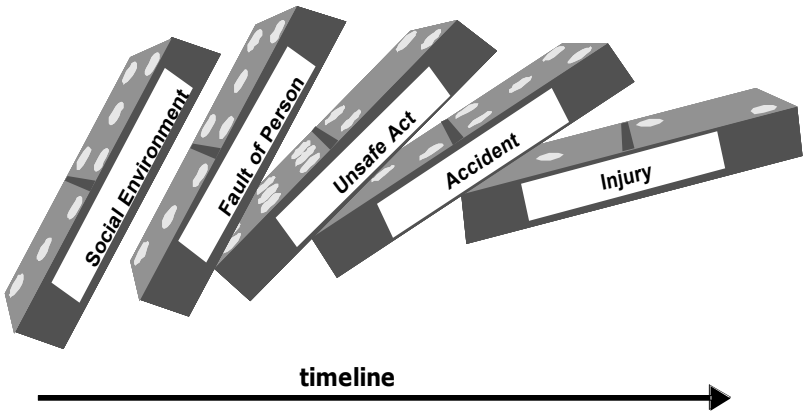



Figure 4-4 Domino model of accident causation and human error exploration (Qureshi, 2007)

#### 4.4.2.2 The epidemiological models

The epidemiological models provide an analogy between the events leading to accidents and the spreading of a disease, as the result of a combination of factors, hidden or apparent, that happen together in space and time (Qureshi, 2007). Therefore, when considering the human contribution to accidents, the active (apparent) and latent (hidden) types of errors should be taken into account. The former are associated with the actions of frontline personnel, while the latter are more likely to occur with personnel whose activities are not related to the direct function, execution and control of the system, e.g. maintenance personnel, designers, or local and headquarters management (Reason, 1990).

The Swiss Cheese developed by Reason (1990) is probably the most well-known epidemiological model. The model comprises a number of layers (cheese slices), as shown in Figure 4-5, which represent the defensive mechanisms of a system. In a perfect world the defensive layers would be flawless and thereby successfully preventing any potential accident. However, in reality all layers have weaknesses (holes), which under certain circumstances can lead to an accident.

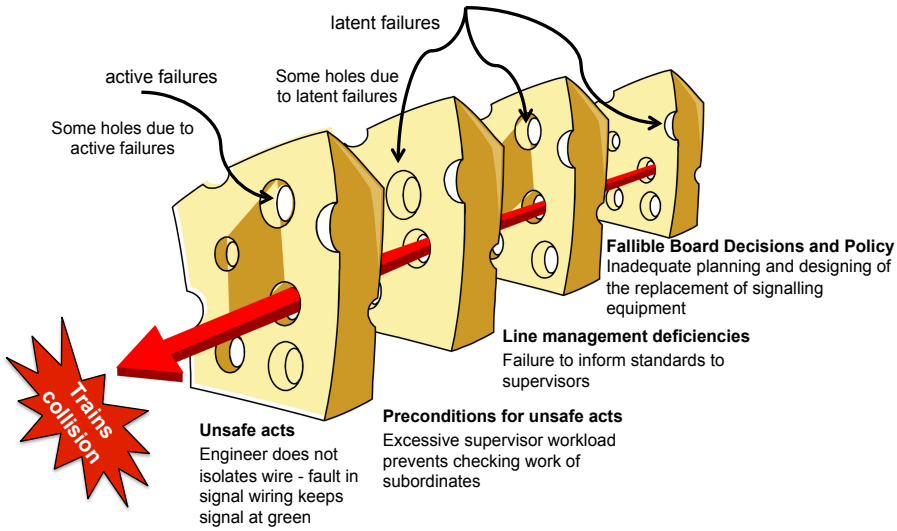



Figure 4-5 The Swiss Cheese model of human error, adapted from (Rail Safety and Standards Board, 2011)

With respect to human errors, the latent types are identified within the *fallible board decisions and policy*, *line management deficiencies* and *preconditions for unsafe acts*. They all refer to failures that may exist for years and never be related to any accident (or incident), until the time that a loss occurs, e.g. safety culture, lack of supervision. The active errors refer to *unsafe acts* and determine an operator's actions and decisions that take place exactly before the accident (or incident) and they are directly associated with the event (Reason, 1990, Hollnagel et al., 2006). According to the model, any accident is considered the result of errors that emerge from *unsafe acts* caused by frontline personnel in conjunction with the organisation's *latent failures*.

The Swiss Cheese model has been used to analyse accidents in several domains, including railways. An example was the Clapham Junction accident, which occurred in London in 1988, as illustrated in Figure 4-6. It is known that the accident was caused by a wiring fault during a re-signalling replacement process (Hidden, 1989). The faulty wire was left by the responsible signalling technician connected at the one end and loose and not insulated at the other. At first glance, it could be claimed that the accident happened due only to the technician's oversight. However, once the findings from the enquiry had been mapped to the Swiss Cheese model, several factors were identified that significantly contributed and finally resulted in the accident.

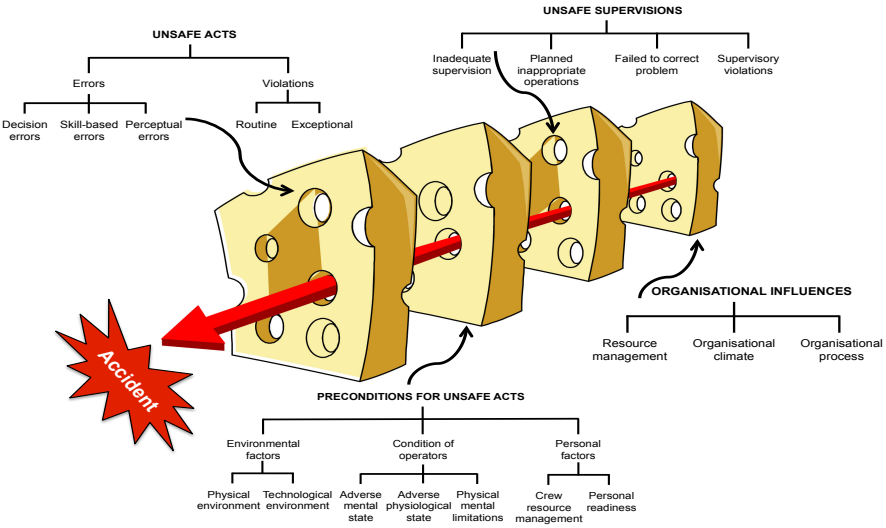



Figure 4-6 The Clapham Junction accident - the Swiss cheese model, adapted from (Rail Safety and Standards Board, 2011)

Starting with the *fallible decisions*, the board neither planned nor designed the replacement process adequately, since they did not ensure that sufficient staff would be available to perform the maintenance works. Consequently, supervisors were not properly informed about the required standards, leading to non-compliance with installation procedures. Furthermore, the understaffed supervising personnel were unable to monitor all frontline employees, including the technician. Finally, the technician who was asked to carry out the task was not sufficiently trained, which significantly contributed to his false decision not to secure the old wiring. The combination of all factors led to the fault in signalling, which resulted in the collision.

By introducing the *latent conditions* the Swiss Cheese model radically changed the way of considering human errors and accidents, which in turn expanded the scope of the investigation process beyond the proximate causes. However, the model has some considerable drawbacks. The main weakness is that the model does not provide analysts with a detailed guidance on its application in a real-world context (Shappell and Wiegmann, 2000). In particular, it does not indicate the exact notion of the *holes*, hence it can be difficult and ambiguous for analysts to exactly identify and define the weaknesses for a specific operation. As a result, the causal associations between the distant latent failures, i.e. fallible board decisions, and the active failures can be difficult to identify and tackle. However, new models, such as the Human Factors Analysis and Classification System (HFACS) (Shappell and Wiegmann, 2000) and the HFACS version for railways (HFACS-RR) (Reinach and Viale, 2006) (explained in Chapter 5) were introduced to address such limitations. Secondly, due to the linear nature of the model, although the cause of an error is no longer a single chain of events, the errors are still explained as the outcome of a fairly simple combination of

events. In addition, the failure of any defensive mechanisms is deemed as the failure of an individual system component, rather than the failure of the entire system (Hollnagel et al., 2006). Hence, the systemic models, which are characterised by, Rasmussen's (2002) Accimaps, Leveson's (2004) STAMP and Hollnagel's (2012) FRAM, were developed to overcome these limitations.

#### 4.4.2.3 The systemic models

As shown in the previous two sections, sequential and epidemiological models underpin the cause-effect concept, where accidents and human errors are expressed as the result of a combination of factors that occur in a linear fashion. More recently new approaches have been adopted, referred to as systemic models, which have two main features: (i) they focus on the performance of the system as a whole and (ii) they support the theory of non-linear causality.

Systemic models usually refer to the term of *concurrence* (Hollnagel et al., 2006). That is, an accident or error can be deemed as the outcome of an unexpected combination of conditions or events taking place at the same time, thereby affecting each other (Perrow, 1984, Hollnagel et al., 2006). Therefore, accidents and human errors are recognised as emergent phenomena, which arise from complex (nonlinear) interactions between the components of the system (Qureshi, 2007).

Leveson (2004) suggests that systems should be "*viewed as interrelated components that are kept in a state of dynamic equilibrium by feedback loops of information and control*". Within this context, human errors should be "*treated as part of an on-going process that is influenced by context, goals, motives, and mental models*" (Leveson, 2004, p.266). In particular, errors are seen as the product of faulty activities involving complex interactions amongst the frontline personnel, managerial and organisational structures as well as regulatory system components. Systemic models provide a detailed and robust framework to investigate in detail accidents and human errors. However, due to their complexity these models can be highly time-consuming, may require extensive resources and a large amount of information in order to provide conclusive results. Finally, in most of the cases it is questionable whether the systemic models add significant benefit compared to epidemiological models with respect to their generated remedial measures or countermeasures (Salmon et al., 2011).

### 4.4.3 Summary of human error modelling

Human error modelling is used to identify the origins of human errors through the prism of accident causation. The strengths and limitations of the dominant approaches are provided in Table 4-3.

Table 4-3 Error causation models

| Approach | Model | Concept | Pros | Cons |
|---|---|---|---|---|
| **Person** | **GEMS** | ▪ Combination of cognitive mechanisms<br>▪ Focus on cognitive errors<br>▪ Distinguishes errors and violations | ▪ Identifies the exact type of cognitive error just before an accident/incident | ▪ Focus only on "*unsafe acts*" without addressing any contextual factors<br>▪ Considers errors as random and independent events |
| **System** | **Sequential** | ▪ Identification of individual factors in a linear sequence | ▪ Identifies types of errors in the higher hierarchy levels of an organisation | ▪ Unclear what the single root cause is<br>▪ No insights concerning any underlying factors that contribute to an error |
| | **Epidemiological** | ▪ Errors seen as analogous to spreading of a disease<br>▪ Identification of active and latent errors<br>▪ Identification of defensive mechanisms | ▪ Identifies active and latent failures<br>▪ Causality no longer a single chain of events | ▪ Not always clear identification of latent failures<br>▪ Ambiguous definition of causal associations between active and distant latent failures<br>▪ Considers error as the failure of an individual component |
| | **Systemic** | ▪ Consideration of the whole system performance<br>▪ Consideration of complex interaction amongst systems components<br>▪ Consideration of errors in a nonlinear fashion | ▪ Errors are identified as emergent phenomena that arise from complex interactions amongst system's components | ▪ Time-consuming and intensive resources<br>▪ Remedial measures or countermeasures strongly related to analysts experience<br>▪ Require a lot of information for constructive findings |

The two approaches, the person and system, can be claimed to be complementary. This implies that in order to achieve the best understanding of the origin of human errors the two models should be integrated. Firstly, the system approach is used to identify the active and latent failures and address all the factors that may contribute to an error. In turn, the person approach is used to define the exact type of error that an individual made just before the accident/incident occurred. Depending on the complexity of the model, errors can be viewed as the product of linear or complex combinations of factors that happen coincidentally.

Based on the nature of the approach, analysts have the ability to investigate and identify common patterns in the occurrence of human errors with regard to or regardless of the characteristics of the individual in a broad variety of contexts. Subsequently, human error modelling supports analysts to examine and determine the contextual factors, apparent and hidden, that may affect the performance of the individual and contribute to the occurrence.

Thus, the next section discusses the context within which human error occurs.

## 4.5 Human Factors

The previous section presented the different approaches to classify human errors and identify their origins. It was determined that errors should be studied within the conceptual framework of a reference system, by taking into account the interactions between operators, organisations, technology, environment, procedures and regulations. Therefore, the discipline of Human Factors (HFs) was established and has been broadly applied, aiming at reducing human errors and the frequency of their unwanted consequences in the operation of several systems (Swain, 1990).

### 4.5.1 Definition of Human Factors

One of the first definitions for Human Factors (HFs), in the mid 1960's by Murrell (1965, p.xiii), is "*the scientific study of the relationship between man and his working environment".* Meister (1989, p.2) later expanded it to "*the study of how humans accomplish work-related tasks in the context of human-machine system operation and how behavioural and non-behavioural variables affect that accomplishment".* Further to this, the UK HSE (1999, p.5) introduced a more generic definition of HFs as "*the environmental, organisational and job factors, and human and individual characteristics which influence behaviour at work in a way which can affect health and safety"*.

With regard to railways the RSSB (2004e, p.181) defines HFs as "*the study of all of the 'people' issues that must be considered to assure the lifelong safety and effectiveness of a system or organisation"*, while Priestley and Lee (2008, p.1) define HFs as "*the application of knowledge of human capabilities and limitations to the design and management of the tasks that people do*".

Besides the above, many other definitions are listed in the literature, e.g. (Stanton et al., 2005a, Shaver, 2009). However, the *interpretation* of all definitions implies that a simple way

to consider HFs is to integrate three parameters: type of work, individual and organisation. Subsequently, both human error and human performance should be tackled as the consequence of the dependencies and potential failures between individual capabilities, tasks demands, organisational and environmental conditions.

### 4.5.2 Railway Human Factors

Railway HFs research has developed rapidly over the past two decades. Early work was conducted by a few institutions during the 1960s and 1970s, followed by a period of inertia in the 1980s and early 1990s. However, the evolution of the industry, the implementation of new technical systems and equipment, as well as the need to improve safety and reliability, have accelerated and supported the railway HFs research programs in several countries (Wilson and Norris, 2005).

The literature shows (Wilson et al., 2007a) that there are eight main areas of railway HFs research: (i) system ergonomics and staff attitudes, (ii) train driver behaviour and performance, (iii) performance and behaviour in rail traffic control, (iv) design of human machine interfaces, (v) automation, (vi) track engineering, (vii) human reliability, including reporting schemes, procedures and violations and finally (viii) design for neighbours, passengers and the public. It is apparent that research covers almost all aspects of the railway system such as: driving, signalling and control; maintenance; change of technical systems; passenger interests; design and planning. However, this research focuses on the performance of railway operators. Hence, the next subsections review only the studies that assess and evaluate human performance. A review of the remaining areas of research can be found in Wilson et al. (2005), Wilson and Norris (2005, 2006), Wilson et al. (2007a), Wilson et al. (2007b), Stanton (2012), and Dadashi et al. (2013).

## 4.6 Human performance and Human reliability

Human performance is broadly defined as "*the human capabilities and limitations that have an impact on the safety and efficiency of operations*" (Maurino, 1998, p.29). Within the concept of this thesis, human performance is expressed by the frequency of human error occurrences. Human performance is measured by Human Reliability (HR), which describes *"the probability that a person will accomplish a given task under given conditions in a given time interval within the acceptance limits"* (Bubb, 2005, p.358). Human reliability, in turn, is characterised by Human Error Probability (HEP), which is simply defined by Kirwan (1994, p.5) as:

$$HEP = \frac{Number\ of\ errors\ occured}{Number\ of\ opportunities\ for\ error} \quad (4\text{-}1)$$

Many techniques for the assessment of human performance have been created and implemented across several industries along the years. They are referred to as Human Reliability Analysis (HRA) techniques (Kirwan, 1994). However, by definition human reliability is strongly related to human performance; therefore the techniques are hereafter referred to as Human Performance Analysis (HPA). The majority of the techniques are developed on the concept of identifying and quantifying the factors that have an impact on human performance under different conditions, known as Performance Shaping Factors (PSFs). A number of traditional and new HPA techniques, which support the concept of PSFs are discussed in the next sections.

## 4.7 Overview of Human Performance Analysis

Human Performance Analysis (HRA) is generally described as "*the use of systems engineering and behavioural science methods in order to render a complete description of the human contribution to risk and to identify ways to reduce that risk"* (Swain and Guttmann, 1983, Boring, 2005). Its roots are associated with the study of human performance and its principle functions rely on the identification, quantification and reduction of human errors (Kirwan, 1994). HPA is a substantial part of the risk assessment process in complex systems, broadly known as Probabilistic Risk assessment (PRA) (Kirwan, 1994).

The main purpose of HPA is "*to estimate the likelihood of particular human actions (that may prevent hazardous events) not being undertaken when needed, or other human actions that may cause hazardous events (by themselves or in combination with other conditions) occurring*" (Wreathall et al., 2003, p.ix). In turn, as Hirschberg (2004) states, the main objectives of HPA are to (i) ensure that the key human interactions are systematically identified, analysed and incorporated into the safety analysis in a traceable manner, (ii) quantify the probabilities of human interactions success and failure and (iii) provide insights to improve human performance. With respect to the quantification of failure probabilities it should be noted that any attempt at estimating the likelihood of errors requires the consideration of the contextual framework within which the error occurred (Kyriakidis, 2009).

Kirwan (1994) introduces the traditional process of HPA, as shown in Figure 4-7. However, based on this description, Boring (2010a) indicates a simpler representation of the process, which can be seen as encompassing three main phases: (i) the modelling of the potential contributors to human error, (ii) the identification of the potential contributors to human error

and (iii) the quantification of human errors. The first phase, which comprises steps 1 to 3 of Figure 4-7, contains the definition of the problem and the required task analysis to break down an overall sequence of events into smaller, more suitable for the analysis units. The task analysis is used to determine what human actions could occur and the probability of errors for these actions, as shown in Chapter 2. In the second phase, step 4, the relevant PSFs are selected. The PSFs are used to determine which aspects of human behaviour have an impact on the outcome of human actions. Finally, within the third phase, steps 5-8, the quantification of errors is conducted with the calculation of human error probabilities, while the overall risk level is also determined. Many approaches are used for the quantification of errors including experts' estimation, the use of performance shaping factor multipliers, Bayesian approaches, and simulations.

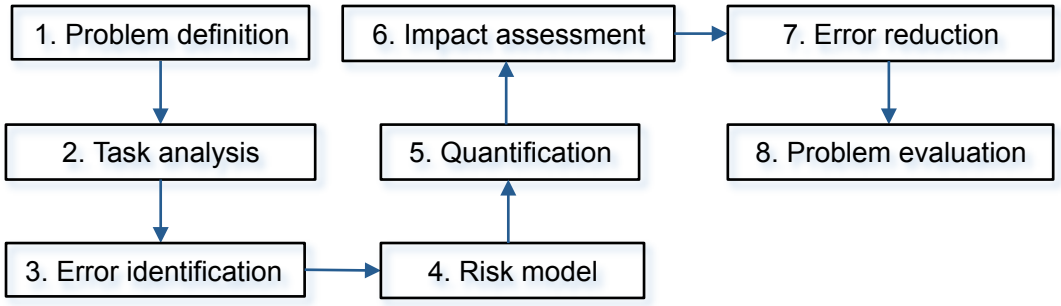



Figure 4-7 The Human Reliability Analysis process

## 4.8 Synopsis of HPA techniques

HPA techniques were initially developed to study the influence of human performance on the safety of nuclear power plants. However, over the years they have been widely applied to several domains, including transportation, health and oil. Due to their widespread use, many reviews have been conducted, including, the studies by Gertman and Blackman (1993), Kirwan (1994), Hollnagel (1998), Pyy (2000), Stanton et al., (2005a), Forester et al., (2006) the EUROCONTROL (2007), Gibson and Kirwan (2008a, 2008b) and Bell and Holroyd (2009).

Based on the findings from the literature, the HPA techniques can be classified into six categories according to: (i) generation, (ii) outcome, (iii) level of model details, (iv) contextual factors, (v) dependence on time, and (vi) type of data used. Figure 4-8 shows the classification of HPA techniques, and provides examples of techniques for each of the categories.

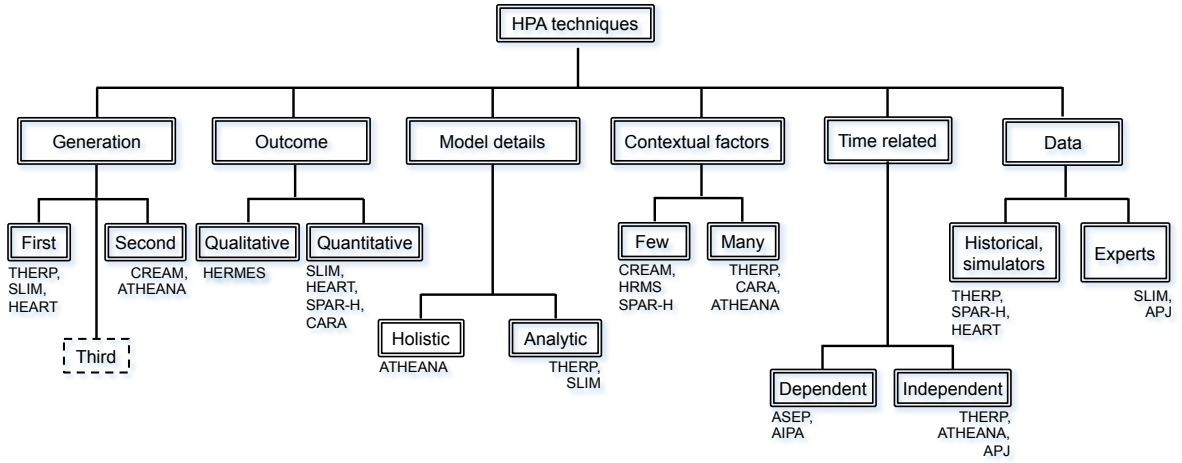



Figure 4-8 Classification of HPA techniques (Pyy, 2000, Boring, 2005)

### 4.8.1 First and second generation methods

Since the 1970s, HPA has become highly developed to improve the study of human influence and contribution to a system's reliability (Kyriakidis, 2009), developing techniques that evolved from the first to second generation HPA.

First generation approaches were developed to help risk assessors predict and quantify the likelihood of human error. They encourage researchers to break tasks into sub-tasks before considering the potential impact of modifying factors such as time pressure, equipment design and stress (Bell and Holroyd, 2009). Furthermore, they focus on the skill and rule base level of human action. Therefore, they are often criticised for failing to incorporate factors such as the impact of context, organisational factors and errors of commission (Bell and Holroyd, 2009).

The 1990s saw the development of second generation techniques, a process that it is still on-going. In contrast to the first generation methods, the main characteristics of these are (i) an emphasis on decision-making and cognitive performance, (ii) the use of a theory-based error taxonomy (often coincides with cognitive model of human behaviour) and (iii) the use of a complex match of error scenarios for the error identification and quantification (Bell and Holroyd, 2009). Such techniques attempt to consider context and errors of commission in human error prediction. However, despite their widespread use, their evolution has been slow, with their empirical validation still to be confirmed. Subsequently the benefits of the second generation over first generation approaches have not been established. (Bell and Holroyd, 2009, Kyriakidis, 2011).

### 4.8.2 Qualitative and quantitative methods

A further classification distinguishes the HPA techniques based on their outcomes, into qualitative and quantitative methods (Boring, 2005). The former focus on identifying and defining the human erroneous actions or the events are caused by them. In addition, they can be used to design a new system or redesign an existing one, based on the findings of the factors that affect the performance of operators. Qualitative methods also provide the necessary information for the quantitative part of the analysis (when a method supports both parts). Finally, they can be considered as the result of a task analysis or incident investigation.

On the other hand, quantitative methods focus on converting the identified errors into Human Error Probabilities (HEP). These techniques are dependent on existing data from previous studies on simulations. If such data are not available, then subject matter experts (SMEs) are asked to determine HEPs for a given context. The majority of the quantitative techniques portray human actions as a combination of tasks that operators should carry out in conjunction with the factors that affect their ability to perform these tasks. Quantification determines the likelihood that a particular action will fail, while the error likelihood ranges from 1/100,000 for high reliability to 1.0 for guaranteed failure, with the nominal values to be between 1/1000 to 1/100 (Gertman and Blackman, 1993).

### 4.8.3 Methods with performance shaping factors

Performance Shaping Factors (PSFs), describe the influence of the contextual factors on human performance and subsequently on the probability of human error. There are two main ways to classify HPA techniques based on the number of PSFs, such as techniques with (i) many and (ii) few PSFs (Boring, 2010b).

Techniques that include a few PSFs incorporate a higher level of factors that encompass large categories of contributing factors to human performance such as environmental, organisational or personal factors. In contrast, models that comprise many PSFs provide a more comprehensive list of factors, such as training, fatigue, stress or low visibility. Such models do not count all PSFs for every event or error, but only those that have been identified from the investigation scheme. Therefore, they also attempt to monitor and control the cases of double-counting of the related influences.

### 4.8.4 Holistic and analytic methods

Further to the distinction of many or few PSFs, HPA techniques could also be classified as holistic or analytic, primarily depending on how analysts perceive and tackle a particular problem. Holistic models account for human performance as indivisible part of a greater situation that cannot be broken into smaller parts. In contrast, analytic techniques consider human performance as a compound of individual elements of a whole situation, which can be decomposed and analysed individually (Pyy, 2000). Boring et al. (2005) state that analytic approaches estimate human error probabilities by judging the influence of individual PSFs on operator performance. On the other hand, in the holistic methodologies, the judgments do not account for the influence of each PSF, but they view the event instead as an irreducible whole.

### 4.8.5 Time dependence

Time dependence is another way to classify HPA methodologies. However, only a few techniques in this category are addressed in the literature (Gerdes, 1995, Pyy, 2000). They are used to assess time dependent error probabilities, which describe the relationship between the available time during which an operator is allowed to perform a task and the probability to fail to perform the task in time. Time dependent models are appropriate to be used for assessing error probabilities after an initiating event, especially when the reaction time is short. However, they are time consuming and cumbersome methods in terms of collecting data; therefore they are not broadly developed.

### 4.8.6 Data used

Finally, the HPA techniques can be distinguished based on the type of data they use to quantify human performance (Pyy, 2000). These can be data derived from the relevant literature, simulators, operating experience or expert judgment. Such classification is strongly related to the PSFs classification, since data are used primarily to assess and quantify the level of influence of PSFs. Within this classification, three approaches are identified. In the first, generic human reliability data, which are derived from the relevant literature are used to estimate human performance and then updated by using PSFs, e.g. the HEART methodology (Williams, 1988). In the second approach, specific operational data, when available, are directly implemented. Finally, in the third approach experts are asked to assess the status of PSFs and the level of PSFs influence on human performance (Pyy, 2000).

## 4.9 Performance Shaping Factors

As already stated, Performance Shaping Factors (PSFs) constitute a substantial part of the HPA process, widely used to assess human performance within a defined context. PSFs have been defined by Boring et al. (2007) as "*all these factors such as age, working conditions, team collaboration, mental and physical health, work experience or training which enhance or degrade human performance*". Hence, they express all various situational, organisational, systemic, personal and environmental factors that may influence the performance of an individual. PSFs account for both the positive and negative influence of the factors on human performance. However, generally, researchers are primarily interested only in the negative impact. Many HPA methods use the state (level of influence) of PSFs to assess human performance or to obtain qualitative insights about a specific operational scenario. The level of influence of PSFs is expressed in various ways. However, the qualitative approach of characterising the influence on a scale from low to high with increments of very important, medium or less important has been most broadly applied (Hollnagel, 1998, Groth, 2009).

PSFs have been numerously classified. Based on their characteristics they are primarily classified as internal and external, with the latter describing situational characteristics, job and task characteristics or environmental circumstances whereas internal PSFs refer to individual characteristics (Kumamoto and Henley, 1996). This classification, due to its simplicity, has been widely accepted and applied to several cases (Sasou and Reason, 1999). In addition to this classification, Boring et al., (2007) distinguish direct from indirect PSFs with the former describing those PSFs that "*can be measured directly, whereby there is a one-to-one relationship between the magnitude of the PSF and that which is measured*", e.g. training as expressed in hours of simulators. Conversely, indirect PSFs include those that "*cannot be measured directly, whereby the magnitude of the PSF can only be determined multivariately or subjectively*", e.g. team collaboration, there are several scales to measure team dynamics, but these are of the indirect variety. Although the distinction between direct and indirect PSFs is useful, without setting clear criteria to distinguish between them, researchers may form opinions that vary from case to case. Therefore, the use of direct and indirect PSFs may be inefficient and lead to inconsistent results amongst researchers.

Further to PSFs classification, several PSFs taxonomies are addressed in the literature, e.g. (Rasmussen, 1982, Swain and Guttmann, 1983, Williams, 1988, Hollnagel, 1998), including, amongst others, the THERP, SLIM, PHECA, HEART and CREAM taxonomies. The existing taxonomies are applied to a specific domain and for a pre-defined purpose, while they are

divided into those that comprise a detailed set of PSFs and those that include a more generic set. Despite this, detailed and generic taxonomies have considerable similarities and include similar PSFs regardless of their domain and application. Detailed PSFs taxonomies comprise factors such as supervision, temperature, training, or stress. In contrast, generic taxonomies include factors such as working conditions, organisational factors or personal factors. As shown in Table 4-4 there is an obvious overlap between the generic and detailed taxonomies, i.e. the generic taxonomies are expressed by categories, which contain the individual PSFs, e.g. the category organisational factors includes supervision.

Table 4-4: A sample of performance shaping factors

| Detailed PSFs | Generic PSFs |
|---|---|
| Training | Personal factors |
| Visibility | Environmental factors |
| Temperature | Working conditions |
| Supervision | Organisational factors |
| Fatigue | Adequacy of Procedures |
| Leadership | Available time |
| Time to execute a task | Team collaboration |
| Job satisfaction | |
| Roster planning | |

Although the role of PSFs in studying human performance is significant, several limitations of the existing taxonomies have been acknowledged. Boring et al. (2007) and Groth (2009) indicate that with the large number of current taxonomies, there is neither an agreed set to be used within the context of HPA methods nor a confirmed number of PSFs that should be included in each of the taxonomies. Forester et al., (2006) suggest that at least fifteen PSFs should be included in a taxonomy, however this suggestion cannot be applied to already established taxonomies.

A further limitation is related to the given definitions of current taxonomies. Definitions of current taxonomies are usually neither available nor explicitly described. For instance, working conditions can be described as the nature of the physical conditions such as ambient light, noise or temperature (Hollnagel, 1998). However, this definition is similar to what Kim and Jung (2003) define as environmental conditions. Therefore, due to the differences in definitions, researchers may reach controversial conclusions even when using the same factors.

Moreover, the impact of each PSF on human performance is another matter of concern, since not all PSFs affect performance equally. For instance, for a task A, lighting or training may influence human performance more than communication, while for another task B,

communication may be considered more significant. However, the quantification of the impact of PSFs on human performance is not straightforward. Thus, only a few HPA techniques provide analysts with a framework for such a quantification (Kirwan, 1994). Some techniques, such as the CREAM (Hollnagel, 1998) and SLIM (Forester et al., 2006) consider the different levels of the influence of PSFs on human performance and use simple equations to estimate this (Embrey et al., 1984, Hollnagel, 1998), yet there is no justified claim that they provide reliable results (Bell and Holroyd, 2009).

Furthermore, the majority of current taxonomies assume that the PSFs are independent, which in most of the cases is an arbitrary assumption. For example, when operators execute a complex task, their level of training is related to the level of stress. Therefore, the dependencies between PSFs should also be taken into account in order to derive more accurate estimations. Finally, the relationship between PSFs and time poses considerable problems since many of the PSFs are ”dynamic” variables, which implies that these they change while a task is executed, e.g. weather conditions or fatigue. Because it is too complex to estimate these changes, most of the techniques neither account for the relation between PSFs with time nor analyse PSFs separately for every condition (Kyriakidis et al., 2012).

## 4.10 Review of Human Performance Analysis Techniques

The need to include human factors research in the safety assessment of railways is widely recognised by most of the stakeholders of the railway system (Cacciabue, 2004b). Thus, to determine the appropriate HPA technique, it is necessary to start with a review of the existing approaches that have been developed and used in the railway domain. The characteristics of the approach should contain: (i) a detailed list of contextual factors, PSFs, that affect human performance in railway operations, including factors that may affect train drivers, signallers and controllers, (ii) the ability to quantify PSFs (considering their mutual dependencies) and subsequently to assess human performance based on PSFs quantification, (iii) the ability to compare human performance for certain actions or estimate the likelihood of errors and (iv) the ability to use the findings to direct resources more efficiently towards the development of sound solutions for improving operators performance. However, due to the limited number of the techniques customised for the railway industry, the review also considers approaches employed in other domains. These approaches have been developed primarily for the nuclear, aviation and oil industries. Figure 4-9 presents the methodology used in this thesis to determine whether an existing taxonomy can be used for the assessment of human performance on the influence of PSFs.

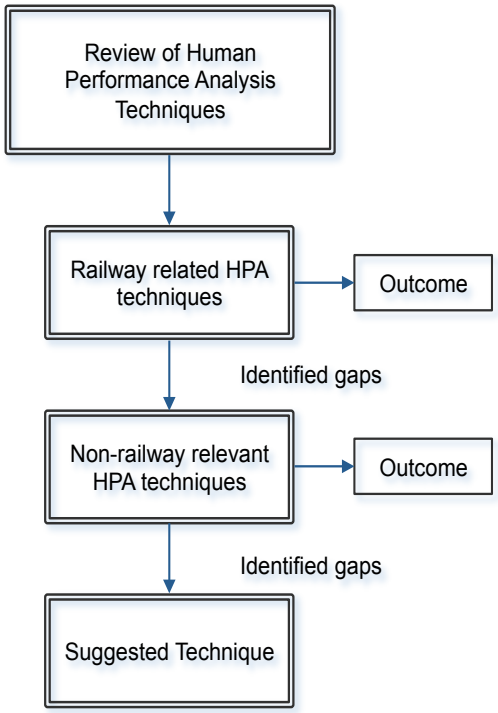



Figure 4-9 Review of HPA techniques

### 4.10.1 Railway Human Performance Analysis approaches

Throughout the continuous development of the research, different approaches have been developed to assess human reliability and performance. However, the relevant literature, (European Commission, 2000, Vanderhaegen, 2001, Wreathall et al., 2004, Wilson et al., 2005, Wilson and Norris, 2005, Wilson and Norris, 2006, Wilson et al., 2007a, Wilson et al., 2007b, Dadashi et al., 2013) shows that very few were established on the concept of PSFs, while only one can be argued to be a complete HPA methodology. Table 4-5 presents in some detail the approaches, their strengths and limitations.

.

Table 4-5 Railway Human Performance Analysis Techniques

| HPA technique | Overview | Advantages | Disadvantages |
|---|---|---|---|
| Analysing and modelling train driver performance (McLeod et al., 2005) | 1. An approach that aims to understand and assess risks of driver unreliability with extended uses of Automatic Warning Systems (AWS) on the UK rail network.<br>2. It provides a situational model to identify possible states of cognition at the time of interaction with any perceptual/cognitive task that could lead to the loss of driver's reliability. | 1. Analysis of driver's performance at a specific time "Now".<br>2. Illustrates factors for understanding human cognition and performance.<br>3. Takes into account characteristics of context and situation at the time behaviour of interest was occurred.<br>4. Proposes a detailed situational model for human performance. | 1. The method has been applied only for one set of H-M interaction.<br>2. It provides only qualitative analysis (level of risk of driver unreliability is given by ALARP).<br>3. Experts' judgment is involved in defining and modelling the scenarios.<br>4. Neither comprehensive nor rigorous approach. |
| Driver performance modelling to railway safety (Hamilton and Clarke, 2005) | 1. A model that is capable to predict the performance time, workload and error consequences of different operational conditions.<br>2. The method aims to support SPADs risk management strategies. | 1. Accounts for conflicts between time pressure, cognitive and behavioural demand.<br>2. Investigates interactions between train driver and equipment.<br>3. Objective and quantifiable approach.<br>4. It can be applied to several tasks. | 1. Time consuming.<br>2. Experienced analysts.<br>3. Time data are derived from literature.<br>4. It is centred only on the requirements of the train driver.<br>5. The quantification part is based on SHERPA, which has been developed for process industries. |
| Human factors in the railway system (Hammerl and Vanderhaegen, 2009) | 1. A model to structure the influence of PSFs on human performance.<br>2. Provides a model of work system core and shows the interactions of contextual factors within the concept of this model. | 1. New PSFs classification.<br>2. Accounts for dependencies between PSFs.<br>3. Introduction of human – barrier interaction.<br>4. Detailed qualitative analysis. | 1. It is not a complete HPA approach.<br>2. Quantitative approach is not introduced. |
| RARA (Gibson et al., 2013a) | 1. A methodology to quantify human error across railway industry developed on the HEART approach.<br>2. The first tailored, complete and customised HPA for railway industry. | 1. It is a quick and simplistic approach that needs little of training.<br>2. It gives error reduction suggestions.<br>3. It gives the analyst a quantitative output.<br>4. Accounts for overlaps amongst EPCs. | 1. It is a subjective method, thus, its reliability and consistency may be debatable.<br>2. Little guidance is offered as to the assignment of EPCs, or their exact number is included in a scenario.<br>3. Still on-going validation process.<br>4. Focusing primarily on drivers tasks. |

#### 4.10.1.1 Analysing and Modelling Train Driver Performance

McLEod et al., (2005) analysed and modelled the performance of train drivers in conjunction with a safety system, i.e. the Automatic Warning System, on the UK railway network. The research focused on understanding human performance within the AWS context at the time a signal is intended to influence the behaviour of driver. For this, the study introduced a situational framework. Such a framework highlights the significance of understanding the state of driver cognition at a specific time, i.e. "*Now*", in a specific situation and a specific context.

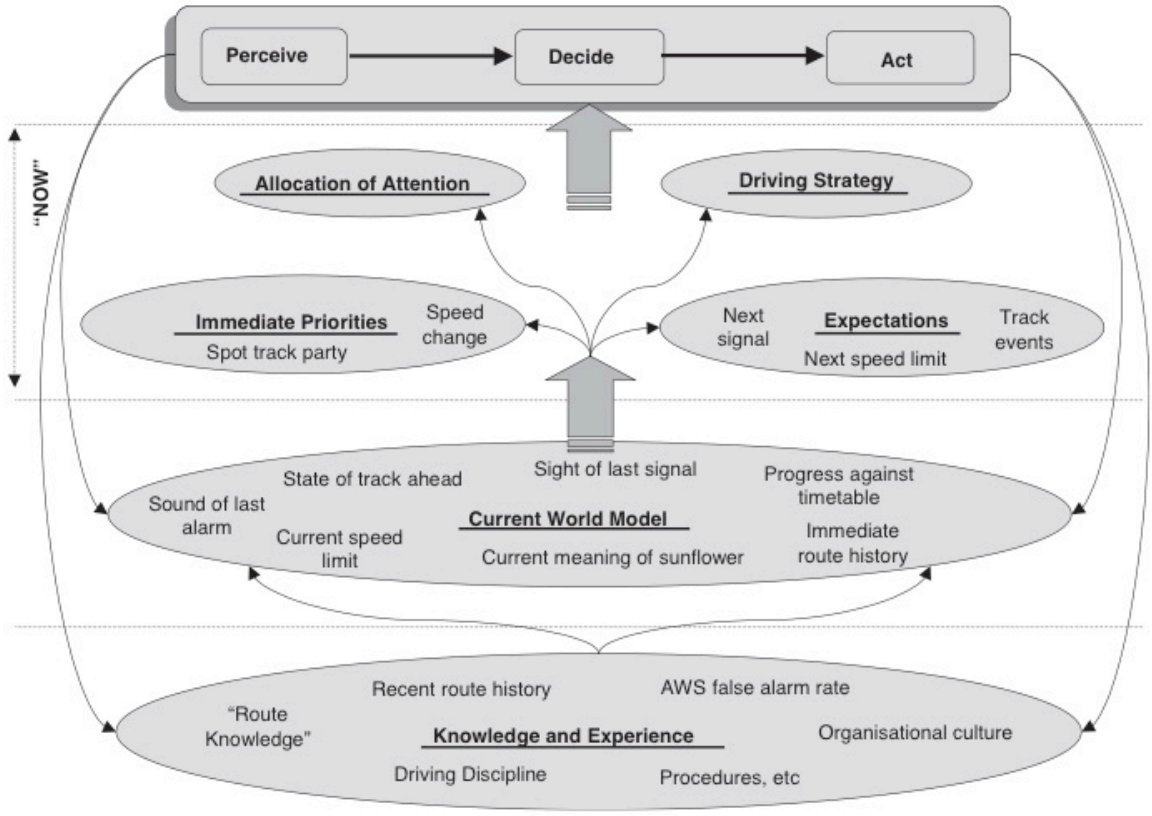



Figure 4-10 Situational model of driver performance in interacting with AWS (McLeod et al., 2005)

The model relies on a simple-information process, i.e. perceive, decide, act, alike the concept of railway operational errors (Section 4.1). According to the model, the performance of train drivers is based on their knowledge and experience, while the acknowledgement of the current "*state of the world*" has an impact on the their driving strategy. Subsequently, to define the cognitive status of train drivers at any specific moment two factors should be considered. The immediate priority and the expectation of what will be the situation within the coming few moments. Both immediate priorities and expectations change along the route, affecting a driver's allocation of attention and driving strategy.

The model aims at identifying the possible levels of cognition at the exact time of interaction with an AWS signal (or any other perceptual/cognitive task) that may compromise the reliability of a driver in a straightforward and simple way. However, neither the situational framework nor the methodology to assess train driver performance is comprehensive or rigorous. Further to this, the approach strongly depends on the experience and ability of researchers to bridge complex psychological constructs to real world situations. Therefore, although the main advantage of the method is focus on deeper understanding of complex psychological factors involved in real-time human performance, it does not provide detailed information of how inexperienced personnel could apply its concept. Furthermore, although the concept of the methodology is novel, the authors do not provide any remarks with respect to its transferability and necessary modifications if the framework changes, e.g. TPWS instead of AWS. Moreover, although it is supposed to support the assessment of human reliability, it does not provide any quantitative measurement. On the contrary, it does only qualitatively show the consequences of the interaction between drivers and AWS. Finally, although the generic concept of the model could be transferred and used for the assessment of the performance of signallers and controllers it requires substantial modifications. Thus, its current format is not suitable for such purposes.

#### 4.10.1.2 Driver Performance Modelling to support Railway Safety

The process of modelling train driver performance (Hamilton and Clarke, 2005) attempts to understand and manage the interactions between train drivers and infrastructure. The model deploys the features of cognitive theory and identifies driver performance in association with the infrastructure characteristics and operational conditions. Consequently, the model predicts the performance time, workload and consequences of driver error for different operational scenarios.

The model uses a defined driving line layout, where memory, perception and anticipation of drivers are measured. Since all the variables have different levels of capacity and reliability, their measurement shows where a driver's capacity is over-used, well-used or under-used. The model measures the timing of key tasks, e.g. stop at a signal, on the defined line, while it also takes into account parameters with respect to rolling stock, train speed, track and visibility conditions. The model assesses the driving capability of operators by calculating their workload, time to perform an action and error proneness. The time to execute an action is given by the Fitt's Law equation:

$$T_{Pa} = T_P + T_C + T_M + I_M log_2\left(\frac{D}{S} + 0.5\right) \quad (4\text{-}2)$$

where, $T_{Pa}$ is the time to perform the action, $T_P$ perceptual processing, $T_C$ cognitive processing, $T_M$ motor processing and $I_M$ movement iteration time.

However, for the calculations the model, instead of creating its own dataset, e.g. simulations, uses data from the literature for elementary behaviours on engineering psychology for three different reaction times: fast, median and slow. In addition, the model describes workload in three dimensions as: (i) time pressure workload, (ii) cognitive demand workload and (iii) behavioural conflict workload. However, although the authors refer to workload statistics to assess workload, this is not explicitly conveyed in the study. Finally, the method claims to support error prediction. Again, no clear instructions are given for this, apart from the fact that the SHERPA (as described in Stanton et al., 2005a) methodology is used to identify the generic error sources. However, since SHERPA is not a quantitate technique, errors predictions are purely qualitative.

The creators of the method claim that their work offers a powerful means for identifying human performance problems in an objective and quantifiable manner. In addition, they state that their approach could also be applied to any aspect of any task that imposes workload on the driver, e.g. the cab interfaces, the design and placement of furniture on the track, or the impact of speed profiles and timetabling on the way the route is driven (Hamilton and Clarke, 2005). However, the approach is time-consuming, cumbersome for inexperienced analysts, and requires extensive resources. In addition, although it accounts for external factors that affect driver reaction times, the approach does not consider any internal factors, e.g. distraction or stress that may affect the performance of the operators. Finally, the methodology focuses only on the train driver tasks and duties. Therefore, its transferability and applicability to other operators is tentative.

#### 4.10.1.3 Human Factors in the Railway System Safety Analysis Process

Hammerl and Vanderhaegen (2009) have addressed the issue of human factors in the railway system with the aim to resolve the limitations of existing HPA techniques. The authors reviewed: (i) the areas of human error in the quantitative risk analysis, (ii) the existing techniques for human reliability assessment and (iii) the great variability of human error. Consequently, they found that no agreed values were established for human error probabilities in the railway industry. In addition, they suggested that cross-industry HPA techniques have limitations (described in the next section) which restrict their direct transferability to the railway industry. Therefore, the authors propose that a new approach is

required for the railway industry. This approach should be "*less fine-grained*" in order to stay practicable and it should provide at least quantitative assessments of performance. Thus, the authors introduced an approach that relies on a model of working environments, which illustrates the engineering and psychology perspectives of the system, as shown in Figure 4-11.

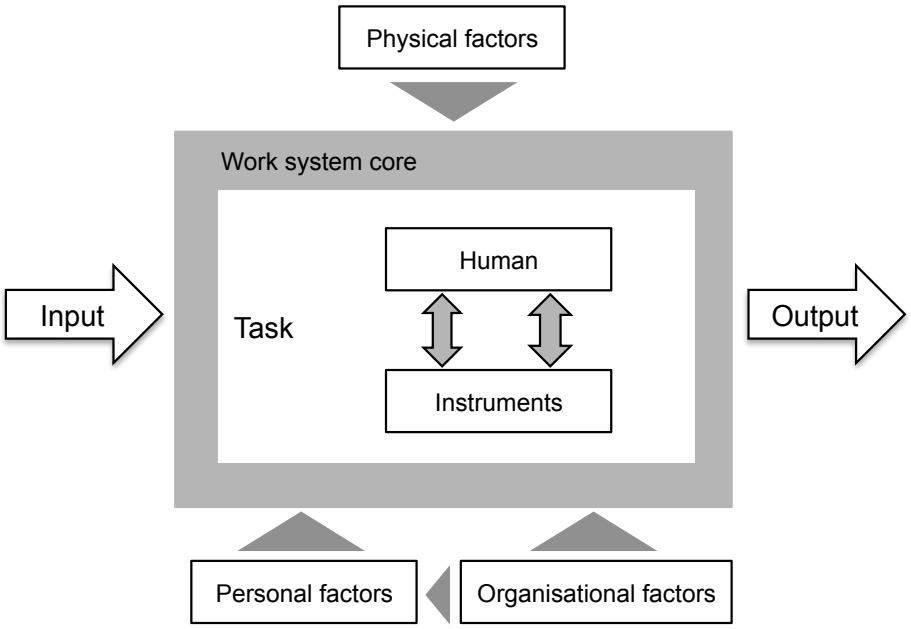



Figure 4-11 Model of working system and PSFs (Hammerl and Vanderhaegen, 2009)

In addition, this model defines the work system core as an interaction of individuals, their tasks and their instruments, while human performance is influenced by a number of PSFs. The PSFs are classified into three main categories, physical, personal and organisational factors. The three categories are further subdivided into seven groups and contain a list tailored to railway industry. Moreover, the model accounts for the dependencies between PSFs. For example, it indicates that some of the organisational factors affect personal factors, e.g. training methods have an influence on the competence skills of employees. Within the system two types of variables can be observed. The first is located at the top and bottom of the system core and continuously affect it. This type of variables, although continuously affecting the core system, does not change over a certain period of time, e.g. during a shift, and is therefore, defined as less dynamic. On the other hand, the second type, i.e. the variables on the horizontal axis, represent the exact moment of operation and therefore are defined as dynamic, i.e. information that drivers receive.

Hammerl and Vanderhaegen provide a simple, qualitative approach to structure the PSFs that affect the performance of railway operators. In addition, it is the only railway approach that concerns the dependencies between PSFs and the influence of such dependencies on

human performance. However, the dependencies are only qualitatively represented. The authors suggest that weighting of dependencies is required in order to firstly assess the degree of influence of particular PSFs on performance and secondly, reveal the most frequent and severe error favouring conditions. They propose that such weighting should be conducted either by both railway and human factors expert judgments or with the support of simulations and human factors studies in relevant environments. However, although they highlight the importance of the quantitative assessment, there are two main drawbacks. Firstly, such quantification is not supported in their study. Secondly, they do not address how the assessment would be conducted in the real world, e.g. methodology, experts characteristics, interpretation of weighting.

#### 4.10.1.4 Railway Action Reliability Assessment (RARA)

The Railway Action Reliability Assessment (Rail Safety and Standards Board, 2012b, Gibson et al., 2013a), introduced by the UK RSSB, is the most recent and comprehensive approach to human error quantification in the railway industry. It is the only approach that can be claimed to be a HPA methodology. The technique relies on the principles of Williams' HEART (1986) methodology (described thoroughly in Section 4.10.2.4) and was initially developed to capture human errors related to train driver tasks (Rail Safety and Standards Board, 2004b). The required steps for conducting the Railway Action Reliability Assessment (RARA) are:

1. Define the task under investigation.
2. Determine the nominal human reliability for the particular task, as expressed by the Generic Tasks, usually by consulting local experts.
3. Identify the Error Producing Conditions (EPCs), which are evident in the given situation and would have a negative effect on the outcome.
4. Assess the effect of each EPC on human performance on a scale 0-1, where 0 represents the minimum and 1 the total effect. This is again indicated as the Assessed Proportion of Affect (APOA).

The methodology is mathematically expressed by the Equation 4-3:

$$HEP = NEP \times \prod[(EPC - 1) \times APOA + 1] \quad (4\text{-}3)$$

where, NEP is the given nominal error probability for a selected generic task; EPC the given maximum predicted effect of the EPC on human performance; and APOA the assessed proportion of the value of the effect between 0.1 and 1, where 0.1 is a small effect and 1 is a full affect.

The methodology includes eight generic tasks tailored to the railway industry, primarily associated with train drivers' duties, e.g. "*fairly simple task performed rapidly or given insufficient or inadequate attention*". The tasks are grouped in three broad areas of human performance, based on the SRK framework (Section 4.4.1.1). In addition, the approach introduces twenty-seven EPCs that affect performance, classified in six main categories, e.g. task design, procedures and competence management. Contrary to the HEART, the RARA accounts for potential overlaps between EPCs or between generic tasks and EPCs. For instance, depending on the definition of EPCs, the terms 'unfamiliarity' and 'level of experience' may overlap. Within the context of the approach, both terms are related to the frequency that an operator is exposed to a specific task. However, while 'unfamiliarity' primarily indicates rare events, the 'level of experience' refers to routine tasks which may not often have been executed before. The overlapping of EPCs may lead to over conservative results in the cases that EPCs are twice considered in the assessment. Similarly, some generic tasks and EPCs combinations may imply overlapping and result in conservative risk estimations. For example, for the generic task "*performed rapidly or given insufficient or inadequate attention*" it is essential to define the underlying factors that cause inadequate attention, e.g. personal or task EPCs.

The RARA methodology thus provides a useful framework for human error quantification and human performance assessment in railways. Its main contribution is that it offers researchers a list of generic tasks which are prone to errors and, furthermore, provides the ability to estimate operator error probability for such tasks. Although developed based on the tasks of train drivers, evidence suggests that it could also be used to estimate human error of other railway employees, including signallers, electrical control room operators, station staff and track workers (Rail Safety and Standards Board, 2012b). A significant advantage of this technique, apart from is transferability, is that it accounts for the potential double-counting of EPCs and thereby avoids overly conservative risk assessments.

However, similar to the HEART, the RARA methodology has its limitations. First, it is a subjective method, and therefore, its reliability and consistency may be controversial since both depend strongly on researchers and their level of experience and understanding of the approach. In addition, inadequate instructions are offered to researchers concerning the assignment of EPCs, or the proportion of effect of such EPCs. For instance, for a particular task, a researcher may choose not only different generic tasks but also select EPCs and the proportions of effect of the tasks. Subsequently, completely different results may arise. Moreover, it is a very recent approach which is why its application and acceptability are yet to be established. Furthermore, although both the generic tasks and EPCs have been

tailored to the railway industry, this is not the case for their values, which primarily have been adopted from the HEART methodology. Moreover, although the technique provides a list of factors tailored to the railway industry, these factors have been derived based only on the UK railway operations, focusing only on the UK mail line train drivers. In addition, although the technique accounts for overlap amongst EPCs, it treats them as independent entities. Subsequently, it does not account for any potential influence of one EPC over another or their combined influence on human performance.

#### 4.10.1.5 Summary of railway HPA approaches limitations

The main drawbacks of the existing techniques can be summarised as follows:

1. *Adequacy of data*: The main problem of railway HPA is the lack of empirical data for model development and validation. Therefore, despite the extensive research on human performance / reliability in railway operations, only a few models have been developed in order for the quantitative evaluation of performance.

2. *Lack of data*: This has led to the exploration of alternative sources including subject matter experts (SMEs) and simulators. However, both may provide inaccurate or misleading results if not specified appropriately.

3. *Inadequate proof of the applicability of techniques*: The majority of the current techniques have not been applied widely in railway risk assessments. Since they have been developed for specific cases and scenarios, their applicability and transferability is limited.

4. *Dependencies of PSFs:* Only one of the techniques accounts for dependencies between PSFs. However, the dependencies are only represented qualitatively.

Due to these limitations, none of the current railway HPA approaches is suitable for this thesis research. However, some of the features of the methodologies, such as the lists of PSFs, are used here to inform the development of a new tool to quantify the impact of PSFs on human performance.

### 4.10.2 Cross - industry Human Performance Analysis techniques

Nine cross-industry, quantitative HRA techniques, developed on the concept of PSFs, are reviewed in this section. The techniques, which are referred to as Human Performance Analysis techniques, are the: THERP (Swain and Guttmann, 1983), SLIM (Embrey et al.,

1984), ATHEANA (Forester et al., 2007), CARA (Kirwan and Gibson, 2007), HEART (Williams, 1988), CREAM (Hollnagel, 1998), SHERPA (as described in Stanton et al., 2005a), HRMS (Kirwan, 1997), HEPI (Khan et al., 2006), and SPAR-H (Gertman et al., 2004). These represent some of the most well-known and broadly cited techniques in the literature ((Swain and Guttmann, 1983), (Kirwan, 1994), (Hollnagel, 1998), (Pyy, 2000), (Lyons et al., 2004), (Stanton et al., 2005a), (Forester et al., 2006), (Kim et al., 2006), (Bell and Holroyd, 2009), (Kyriakidis, 2011) and (Di Pasquale et al., 2013)). Table 4-6 illustrates a synopsis of the conducted literature review.

Table 4-6 Main features of human performance analysis techniques

| HPA technique | Overview | Advantages | Disadvantages |
|---|---|---|---|
| THERP (Swain and Guttmann, 1983) | 1. A first generation technique that provides both quantitative and qualitative results.<br>2. Initially designed for the nuclear industry, but has been applied to offshore, medical and transport domains.<br>3. Training is required on the method. | 1. It is a broadly used and acceptable methodology.<br>2. It offers a powerful methodology, which can be made auditable by the assessor.<br>3. It performs well in terms of accuracy.<br>4. It accounts for dependencies between human errors | 1. It is relatively unstructured.<br>2. It can be time consuming.<br>3. It is highly judgmental based on assessor's experience.<br>4. Its interaction between certain PSFs is unknown; therefore no guidelines can be given for possible combinations. |
| SLIM (Embrey et al., 1984) | 1. A first generation technique, initially designed for the nuclear industry, but has been applied to chemical industry.<br>2. Suitable for application in major hazard domains.<br>3. Requires an expert panel to conduct assessment. | 1. It is a flexible technique and good theoretical method.<br>2. It is able to deal with the total range of human error forms.<br>3. It does not need task decomposition (task analysis and error taxonomies). | 1. It is a complex method that needs intensive resource.<br>2. The choosing of PSFs is quite arbitrary.<br>3. It is a subjective method, which reduces its reliability and consistency.<br>4. Sometimes there are problems regarding experts' group synthesis.<br>5. There is a lack of valid calibration data. |
| ATHEANA (Forester et al., 2007) | 1. A second generation technique that provides both quantitative and qualitative results.<br>2. Initially designed for the nuclear industry.<br>3. Training is required on the method.<br>4. No evidence that it has been applied to other domains. | 1. It focuses on the prediction of a specific error and on the most influential factors which lead to this error.<br>2. It provides confidence that the major risk associated with the human failure has been captured.<br>3. It is able to estimate HEPs for all sorts of combinations of factors and various conditions.<br>4. It considers more PSFs than other techniques.<br>5. It provides a comprehensive approach to explore how action failures can occur. | 1. Too complex and difficult to be applied by inexperienced researchers.<br>2. Strongly dependent on subject matter experts and their level of experience.<br>3. The provided instructions are complex.<br>4. The quantification method is weak, and results are unjustified.<br>5. It requires extensive resources.<br>6. It is a subjective method, which reduces its reliability and consistency. |
| HEART (Williams, 1988) | 1. A first generation technique that quantifies human errors.<br>2. It has been applied in the nuclear, chemical, aviation, rail, health industries.<br>3. It can be applied to all major hazards sectors.<br>4. It is easy to use by engineers and human factors specialist. | 1. It is a quick and simplistic approach that needs little training.<br>2. It gives error reduction suggestions.<br>3. It gives the analyst a quantitative output.<br>4. A number of validation studies have produced encouraging results. | 1. It is a subjective method; hence its reliability and consistency may be questionable.<br>2. Neither dependence nor EPCs interaction are taken into account.<br>3. Little guidance is offered to the analysts in a number of the key HEART stages, such as the assignment of EPCs, or their exact number included in a scenario.<br>4. Sometimes there are problems regarding experts' group synthesis. |

| HPA technique | Overview | Advantages | Disadvantages |
|---|---|---|---|
| CARA (Kirwan and Gibson, 2007), | 1. A first generation technique that quantifies human reliability aspects as failure and success rates of mitigation actions in ATM.<br>2. Designed for the air traffic control domain, based on HEART approach | 1. It provides an initial indication that HRA can be used to deal with human factors arguments in a quantified ATM safety case context.<br>2. It is a quick and simplistic approach.<br>3. It gives a quantitative and qualitative output.<br>4. It uses customised to ATM GTTs descriptions. | 1. It has been applied to limited cases.<br>2. It can be time consuming when new tasks are introduced.<br>3. Neither dependence nor interactions between error production conditions are taken into account.<br>4. It is a subjective method, which reduces its reliability and consistency. |
| CREAM (Hollnagel, 1998) | 1. A second-generation technique that can be applied for both retrospective analysis and performance prediction.<br>2. Has been applied in the nuclear, rail and aviation industries.<br>3. It can be applied to other major hazards sectors.<br>4. No special skills or knowledge are required but human factors knowledge would be advantageous for its use. | 1. It can be extremely exhaustive.<br>2. It is a clear, well structured, schematic and systematic approach to error identification and quantification.<br>3. It can be used proactively and retrospectively as well as in several domains.<br>4. Its classification scheme is detailed, exhaustive and takes also into consideration environment and system causes of error. | 1. It appears complicated and daunting for a novice.<br>2. The exhaustiveness of the classification scheme (error taxonomy) makes the method more time and resource intensive than other techniques.<br>3. It has not been used extensively.<br>4. It does not offer any remedial measure.<br>5. It would presumably require analysts with knowledge of human factors and cognitive ergonomics, thus analysts with less experience might face difficulties on apply it. |
| HRMS (Kirwan, 1997) | 1. Originally designed for the nuclear industry.<br>2. Due to HRMS complexity a new, quicker technique, called JHEDI was developed, but it was more conservative.<br>3. In the cases that HEPs are found to be risk significant, the HRMS should then be applied.<br>4. Training should be provided for both techniques. | 1. It is useful for scenarios which require in depth analysis.<br>2. It covers the whole HRA process.<br>3. It was founded based on real-industrial data.<br>4. It provides a list of questions to assess PSFs.<br>5. Questions are factual rather than judgmental. | 1. It can be time consuming.<br>2. Highly resource-intensive.<br>3. Difficult to be applied by inexperienced researchers.<br>4. It has not been empirically validated.<br>5. It has only been applied to one assessment; no additional information could be found discuss its validity. |
| HEPI (Khan et al., 2006) | 1. Designed for offshore platform musters, based on SLIM methodology.<br>2. No evidence that it has been applied to other domains. | 1. It provides a quantitative approach for inclusion of HFs in risk assessments.<br>2. It provides insights about the mechanisms and modes under which human errors occur.<br>3. It offers a robust approach to overcome the lack of existing data. | 1. It has been applied to only three muster scenarios. Therefore its validity needs to be examined.<br>2. It can be time consuming and requires intensives resources.<br>3. It does not consider the dependencies and interaction amongst PSFs. |
| SPAR-H (Gertman et al., 2004) | 1. A second generation approach, originally designed for the nuclear industry.<br>2. No evidence that the methodology has been | 1. It is a simple and straight forward to be applied method.<br>2. It provides a list of eight PSFs with precise definitions. | 1. It can be difficult to address the influence of one PSF on another.<br>2. There are no clear instructions on how to |

| HPA technique | Overview | Advantages | Disadvantages |
|---|---|---|---|
| | applied to other domains. | 3. It considers dependencies and interactions amongst PSFs.<br>4. It decomposes probabilities in contribution from diagnosis and action errors.<br>5. It provides a worksheet to simplify estimation process | incorporate a wider range of PSFs when needed.<br>3. The validation of the nominal HEPs was not clearly justified. |

The nine techniques were initially evaluated according to their level of accuracy, validity, usefulness, effective use of resources, acceptability and maturity, as shown in Table 4-7. Kirwan (1994) describes the set of criteria as:

1. *Accuracy* represents the level of agreement between the numerical outcomes of the methodology with already known HEPs values, as well as the level of consistency between the experts and the assessors.
2. *Validity* considers whether the technique uses a set of PSFs that supports the quantification, empirical validity and degree of validity as perceived by the experts and assessors.
3. *Usefulness* refers to the level of qualitative practicality in identifying the error mechanisms and generating reduction measurements and whether such measurements can be assessed.
4. *Effectiveness of resources* concerns the appropriateness of equipment, personal effort and data to conduct the analysis.
5. *Acceptability* refers to the level of the techniques' acceptance by the assessors, experts and relevant scientific community. Finally,
6. *Maturity*, describes the current status of techniques' maturity and the potentials for future development.

Table 4-7 Comparison of Human Performance Analysis techniques

| | Accuracy | Validity | Usefulness | Effective use of resources | Acceptability | Maturity |
|---|---|---|---|---|---|---|
| THERP | H | M | M | L/M | H | H |
| SLIM | M | M/H | H | M | H | H |
| ATHEANA | L/M | M | M | L | M | L/M |
| CARA | M | M | H | H | M/H | M |
| HEART | M/H | M/H | H | H | H | H |
| CREAM | M | L/M | H | M/H | M | L/M |
| HRMS | L/M | M | M/H | L/M | M | L/M |
| HEPI | L/M | M/H | M/H | M | L/M | L/M |
| SPAR-H | M | M/H | H | M | M/H | M/H |

Note: L = Low, M = Medium, H = High

This assessment was conducted to explore whether the techniques show sufficiently significant performance, to be further considered for application to this thesis.

#### 4.10.2.1 THERP (Technique for Human Error Rate and Prediction)

THERP, as described by Swain and Guttmann (1983), is the best-known, first generation HRA method (as cited in EUROCONTROL, 2007). It aims to calculate the probability of the successful performance of the activities necessary for the accomplishment of a task. The technique utilises PSFs and models human errors using probability trees and models of dependence. THERP involves a task analysis to describe the performance characteristics of the human tasks being analysed. The results of the task analysis are represented graphically as event trees that are a formal representation of the required sequence of actions. Hollnagel (1998) describes the six steps of THERP as:

1. Define the system failures that may be influenced by human errors and for which probabilities are to be estimated.
2. Identify, list and analyse the human operations performed and their relationships to system tasks and function of interest, i.e. undertake a task analysis.
3. Estimate the relevant Human Error Probabilities, i.e. predicted error rates.
4. Determine the effects of human errors on the system failure events of interest.
5. Recommend changes to the system in order to reduce system failure rates to an acceptable level.
6. Review the consequences of proposed changes with respect to availability, reliability and cost benefit.

THERP has been demonstrated to achieve high accuracy. Initially developed for estimating risks in the nuclear power industry, THERP has subsequently been applied to other sectors such as offshore oil and gas, and healthcare (Bell and Holroyd, 2009). According to Kirwan (1994) THERP is one of the few complete first generation techniques, in the sense that it describes both how events should be modelled (event trees) and quantified. The dominance of the event tree, however, means that the classification scheme and model are necessarily limited, as the event tree can only account for binary choices (success-failure). Therefore, it is difficult to introduce more complex error modes in the context of THERP. In addition, THERP neither accounts for dependencies amongst the PSFs nor supports the incorporation of such a feature. Finally, data used for the quantification of the HEPs are derived only from scenarios tailored to process industries.

#### 4.10.2.2 SLIM (Success Likelihood Index Method)

SLIM was developed by Embrey (1983) and provides a structured judgement about error probabilities in both procedural and cognitive tasks (EUROCONTROL, 2007). SLIM is a

decision-analytic approach that uses expert judgement to quantify the PSFs that affect human performance. Therefore, a panel of experts is asked to identify the PSFs that are considered to affect human performance for a specific scenario and task. These PSFs are then used to obtain a Success Likelihood Index (SLI), which represents the overall belief of experts concerning the positive or negative effect of each PSF on the likelihood of success for the task under consideration. The SLI can also be calibrated against existing data to derive the final HEPs. The SLIM methodology comprises the following nine steps:

1. Selection of the expert panel.
2. Definition of situations and subsets.
3. Elicitation of PSFs.
4. Rating of the tasks on the PSF scale.
5. Ideal point elicitation and scaling calculations.
6. Independence checks.
7. Weighting procedure.
8. Calculation of SLIs.
9. Conversion of SLIs into probabilities.

Typical PSFs used in SLIM include time pressure or stress levels, quality of information or interfaces, quality of procedures, task complexity, and adequacy of training or level of competence. The SLI for n PSFs is deduced using the Equation 4-4:

$$SLI = \sum_{i=1}^{n} w_i \times r_j \quad (4\text{-}4)$$

where, $w_i$ and $r_i$ are the weighting and rating factors respectively. In turn, SLIs are converted into probabilities by using Equation 4-5:

$$log(HEP) = a \cdot SLI + b \quad (4\text{-}5)$$

where, HEP is the success probability, a (slope of the line) and b (the intercept of the line with the vertical axis) are constants which are derived by processing simultaneous equations, as long as at least two calibration probabilities have been assessed within each task subset or from already estimated HEPs.

The main feature of SLIM is that it defines preferences representing the relative likelihoods of different human errors as a function of various PSFs (Kirwan, 1994). However, as the number of PSFs increases, the method becomes cumbersome and time-consuming, and the panel of experts must be carefully selected to meet the criteria of the study. Therefore, for complex SLIM models empirical data are required to validate the results from experts judgments (Hollnagel, 1998).

The level of accuracy associated with SLIM is indeterminate due to lack of data. However, its theoretical validity is at a reasonably high level. In addition, the list of PSFs, which are included in SLIM is not predefined but tailored every time to the scenario under examination. Subsequently, the methodology is regarded as flexible and applicable to many domains. Furthermore, although SLIM does not account for dependencies amongst PSFs, its concept supports and allows for such consideration. Finally, SLIM allows analysts not only to estimate error probabilities, but also to compare several actions in terms of their proneness to errors by estimating the success likelihood index. Therefore, the SLIM concept is considered suitable for the development of an approach to estimate human performance in railway operations.

#### 4.10.2.3 ATHEANA (A Technique for Human Event Analysis)

ATHEANA was developed and designed for the US Nuclear Regulatory Commission (Cooper et al., 1996) to support the understanding and quantification of human failures in nuclear power plants. The approach is also suitable for application in other industries (Subotic, 2007, Bell and Holroyd, 2009). The basic concept of ATHEANA is that significant human errors are considered as result of “error-forcing contexts” (EFCs), which are defined as the combination of the system conditions and all other influences that contribute and lead to errors caused by operators (Bell and Holroyd, 2009). Forester et al. (2007), summarise the nine steps of ATHEANA approach as follows:

1. Define and interpret the issue (in this step analysts define the objective that is to be achieved by performing the HRA).
2. Define the scope of the analysis.
3. Describe the accident scenario and its nominal context.
4. Define the corresponding Unsafe Actions, which may affect the task in question.
5. Assess information relevant to human performance and characterise the factors (PSFs) that could lead to potential vulnerabilities.
6. Search for plausible deviations from the assessment scenario.
7. Evaluate the potential for recovery.
8. Estimate the HEPs for the unsafe actions.

For a given scenario S the final quantification step of ATHEANA is expressed by:

$$P\left(\frac{E}{S}\right) = \sum_{\substack{unsafe \\ action\ (i)}} \sum_{\substack{error\ forcing \\ context\ (j)}} P_{ij}(S) \qquad (4\text{-}6)$$

where P(E/S) is the probability of the Human Failure Event (HFE) in scenario "S" , Pij(S) is the probability of unsafe action "i" resulting from EFC "j" in scenario "S".

The literature, e.g. (Cooper et al., 1996, Forester et al., 1998, Reer, 2008, Bell and Holroyd, 2009) indicates that ATHEANA is considered as useful methodology for deeper and more comprehensive understanding of the contexts of human performance as the cause of occurrences. However, it is also highlighted that its quantification capabilities are weak with questionable levels of accuracy. In addition, within its concept, the quantification of PSFs is not supported. Results are excessively dependent on assessors, making its findings difficult to verify independently. Furthermore, ATHEANA can be extremely complex resource intensive to operate (Bell and Holroyd, 2009).

#### 4.10.2.4 HEART (Human Error Assessment and Reduction Technique)

HEART was introduced by Williams (1986) as a quick, easy to use and understand approach for the assessment and quantification of the probability of human failure. HEART defines error probabilities for a set of predefined generic task types, and introduces a list of error producing conditions (EPCs) that address only negative influences on human performance. One of the main features of the approach is that in order to reduce resource usage, it only deals with those errors that have a significant impact on the system under consideration. HEART is built upon three main assumptions: (i) basic human reliability is dependent upon the generic nature of the task to be carried out; (ii) under perfect conditions, reliability tends to be achieved consistently with a given nominal likelihood within probabilistic limits; and (iii) given that these perfect conditions do not exist in all circumstances, the predicted human reliability may degrade as a function of the extent to which the identified EPCs may apply. Stanton et al., (2005b) define the six steps of HEART as follows:

1. Refine the task in terms of its generic, proposed and nominal level of human reliability.
2. Identify the full range of sub-tasks that a system operator would be required to complete within a given task.
3. Determine a nominal human reliability for the particular task, usually by consulting local experts. Establish a $5^{th}$ – $95^{th}$ percentile confidence range based around this calculated score.
4. Identify EPCs, which are evident in the given situation and would have a negative effect on the outcome.

5. Assess the effect of each EPC on human performance in a scale 0-1, whereas 0 shows the minimum and 1 the total effect.
6. Calculate a final estimate of the human error probability (HEP).

HEART is mathematically expressed by the Equation 4-7:

$$HEP = NEP \times \prod[(EPC - 1) \times APOA + 1] \quad (4\text{-}7)$$

where, NEP is the given nominal error probability for a selected generic task type; EPC the given maximum predicted effect of the EPC on human performance; and APOA is the assessed proportion of effect value between 0.05 and 1, with 0.05 being very weak effect and 1 representing full effect.

HEART is one of few HRA techniques that has been independently and empirically validated, where the results show a strong correlation between the assessed values and the true values (Bell and Holroyd, 2009). Therefore, the HEART framework has been used to develop new techniques for specific domains, such as the CARA (Kirwan and Gibson, 2007) for aviation and RARA (Rail Safety and Standards Board, 2012b) for railways. In addition, HEART is considered to be a quick and straightforward technique that provides recommendations for the mitigation of errors, does not require a lot of training, while its validity and acceptability are quite high (Kirwan, 1994). However, HEART has certain limitations. Firstly, it does not provide clear instructions on how the generic tasks should be selected, thus making its consistency questionable. Furthermore, there is little guidance and instructions on how the analysts should rate the EPCs. As a result, often significantly different weightings are derived for the same scenario and conditions. Finally, HEART does not consider either dependencies or interactions amongst EPCs.

#### 4.10.2.5 CARA (Controller Action Reliability Assessment)

Building upon the framework of HEART, CARA was introduced by Kirwan and Gibson (2007) to assess and quantify human performance in the context of Air Traffic Management (ATM) safety assessments. CARA introduces a set of generic task types (GTTs), which are specific to the ATM environment, and have been quantified using existing human performance data. In addition, new EPCs tailored for ATM operations are used to assess the negative influence of the contextual factors on human performance. Finally, for the HEP estimation CARA uses the calculation method of HEART, as shown in Equation 4-8:

$$HEP = GTT \times [(EPC_1 - 1) \times APOA_1 + 1] \ldots . \times [(EPC_n - 1) \times APOA_n + 1] \quad (4\text{-}8)$$

where GTT is the human error probability associated with a GTT; EPC is the maximum affect associated with an EPC; and, again, APOA is the assessed proportion of the effect value between 0.05 and 1, where 0.05 is a very weak effect and 1 the full effect.

CARA has been applied to three safety ATM scenarios with initial results to indicate that it can be successfully used to deal with human factors arguments in a quantified ATM safety case (Gibson and Kirwan, 2008a). However, CARA does not address any of the HEART's limitations; subsequently it shares the same limitations and drawbacks, without providing any further improvements.

#### 4.10.2.6 CREAM (Cognitive Reliability and Error Analysis Method)

CREAM (Hollnagel, 1998), was developed to address the limitations of the existing first generation techniques. It was designed to predict, analyse and quantify potential human errors. CREAM enables researchers to (Stanton et al., 2005a): (i) identify the tasks or actions that require or depend on human cognition and subsequently can be affected by variations in cognitive reliability; (ii) define the circumstances under which the cognition reliability can be reduced and in turn where tasks or actions may constitute a source of risk; (iii) provide an evaluation of the consequences of human performance on the safety of the system; and (iv) develop and specify countermeasures to improve the conditions that affect human performance, increase human reliability and subsequently reduce risk.

CREAM uses a model of cognition called Contextual Control Mode (COCOM), which analyses how operators choose their actions. It assumes that the degree of operators' control on their choices varies, and thus their performance reliability varies. The model relies on the fundamental distinction between competence and control, where competence can be defined in terms of individuals' skills and knowledge. On the other hand, control can be described by "*a continuum, going from a situation where a person has little or no control over events to conditions where events are under complete control*" (Hollnagel, 1998, p.154). The performance reliability of operators is expressed by four control modes, including scrambled, opportunistic, tactical and strategic. The higher level of control indicates higher operators' performance reliability, as shown in Figure 4-12.

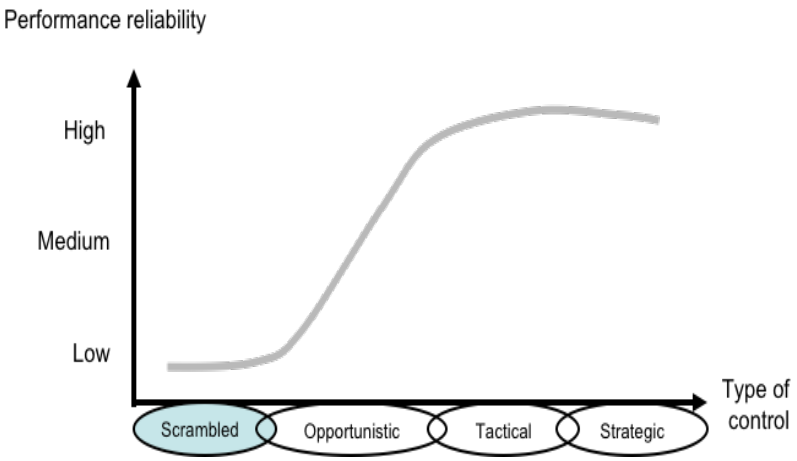



Figure 4-12 Proposed relation between control mode and reliability (Hollnagel, 1998)

Within the CREAM context, human errors are the result of degraded human performance which is affected by a set of factors, referred to as Common Performance Conditions (CPCs). CREAM introduces nine CPCs with specifically assigned levels of effect on human performance. The sum of the performance reliability for each CPC delivers a total CPC score and subsequently the reliability range as shown in Figure 4-13.

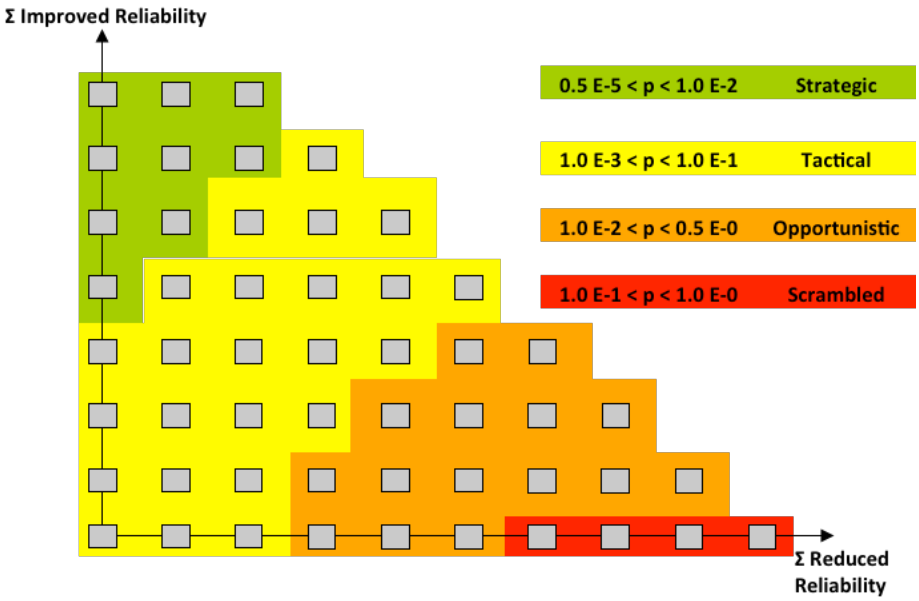



Figure 4-13 Relations between CPC score and control modes (Hollnagel, 1998)

For the purpose of HRA, Stanton et al. (2005a) summarise CREAM's main steps as follows:

1. Conduct the task analysis.
2. Describe the context of the scenario under analysis, define and assess the relevant CPCs.

3. Specify the initiating events that will be subject to the error predictions.
4. Predict errors by defining how an initiating event could potentially lead to an error.
5. Select the task steps to be quantified.
6. Quantitative performance prediction.

CREAM is a well-structured and -documented, schematic and systematic technique to identify and quantify human errors. It can be implemented proactively and retrospectively in several domains. It is one of the very few techniques that accounts for dependencies between PSFs. While it can be extensive, it provides detailed insights of how human actions affect the safety of a system. In addition, its detailed classification scheme enhances the error identification by taking into account the environmental and systemic causes of an error. Although a generic technique, CREAM has not been used extensively, in part because of questions as to its credibility and reliability, especially with respect to how assessors perceive the cognition mechanisms of tasks. Moreover, depending on its level of comprehensiveness, CREAM can be extremely complex and resource-intensive especially for novice researchers. In addition, it does not introduce any remedial measures for the recovery from the human errors, and even for very simple scenarios the required time to apply CREAM is high. Finally, due to its complex nature, its analytical application is presumably limited to researchers with knowledge of human factors and cognitive ergonomics.

#### 4.10.2.7 HRMS (Human Reliability Management System)

The HRMS technique was developed by Kirwan (1997) to provide a comprehensive and accurate approach to assess human contribution to risk in the nuclear industry. The technique comprises three main parts: (i) a detailed task and error analysis, (ii) the quantification of error probabilities and (iii) the error reduction scheme. Kirwan (1997) acknowledged that the HRMS can be resource-intensive, therefore he subsequently introduced a simplified version, referred to as Justification of Human Error Data Information (JHEDI).

HRMS uses a set of six PSFs to describe the factors that affect human performance. For its implementation, an audit document, which includes fifty questions, is used. For each question, experts provide their input (‘yes, ‘no’, ‘not applicable’) to represent the impact of each PSF on human performance. Subsequently, a profile of each PSF is generated for every task and further related to the known value of human error probability for that particular task, which is derived from existing databases. However, Everdij and Blom (as

cited in Bell and Holroyd, 2009) note that HRMS is currently not in use or it is used rarely. Based on its limited use, its complexity and the required extensive resources, HRMS is not a practical model.

#### 4.10.2.8 HEPI (Human Error Probability Index)

HEPI was introduced by Khan et al. (2006), to provide a methodology for the identification, assessment and mitigation of risk caused by human errors during offshore platform emergency musters. HEPI is built upon the concept of SLIM methodology, and therefore, comprises similar steps, including:

1. Establish criteria to choose the assessors.
2. Define the muster scenario parameters.
3. Elicit the data.
4. Calculate HEPs through SLIM procedure.
5. Develop the PSF weight and rating reference graphs.
6. Develop the conversion curves between SLI values and the probability of success (POS).
7. Develop the consequence and risk tables.
8. Develop risk mitigation measures (RMMs) for each muster action based on training, procedures, management systems and equipment.

while it also uses Equations 4-4 and 4-5 as described in Section 4.10.2.2.

HEPI enhances assessors' understanding of human errors, the mechanisms that cause such errors, the modes under which human failures occur, and the factors that affect human performance. However, the methodology fails to tackle any of the disadvantages of the SLIM methodology with its main weakness being that it does not consider dependencies and interactions amongst PSFs. In addition, HEPI has been applied to only three muster scenarios, and although the findings were promising, its limited use reduces its credibility and validity. Finally, the creators of HEPI acknowledge that the approach to ranking the risk level is subjective and depends strongly on the experience and perception of the assessors.

#### 4.10.2.9 SPAR-H (Standardized Plant Analysis Risk Human Reliability Analysis)

SPAR-H was introduced by Gertman et al. (2004) as a simplified method to estimate human error probabilities associated with operators' actions and decisions in nuclear power plants. According to its authors the SPAR-H approach contains five steps:

1. Define the model of human behaviour.
2. Describe the details of the model in behavioural terms, into action and diagnosis tasks.
3. Identify the PSFs that characterise behaviour with the operational context
4. Refine and assess the list of PSFs an.
5. Develop the quantification scheme.

The methodology incorporates PSFs capable of influencing human performance. All PSFs are considered to be significant contributors to human error and dependencies are taken into account. The final human error probability is calculated by multiplying the influence of PSFs, with the nominal error probability for the two different types of tasks, action and diagnosis. The mathematical expression is given by Equation 4-9:

$$HEP = \frac{NHEP \cdot PSF_{composite}}{NHEP \cdot (PSF_{composite} - 1) + 1} \quad (4\text{-}9)$$

where, the NHEP is the nominal error probabilities for the two types of tasks, 0.01 for diagnosis and 0.001 for action and $PSF_{composite}$ the multiplicative calculation of PSFs that are involved in the particular tasks.

SPAR-H is a flexible technique, which gives analysts the ability to implement the methodology beyond specific tasks and to assess human error probability based on the context of specific PSFs. On the other hand, this great dependency of the methodology on PSFs is probably its main weakness. SPAR-H does not provide explicit guidance to address a wider range of PSFs when needed. Therefore, SPAR-H encourages analysts to use more recent context in developing methods when more detail about the application of PSFs is required. Finally, although it addresses the issue of PSFs dependency and despite the detailed discussion of their potential interaction effects, it treats PSFs as independent entities.

#### 4.10.2.10 Summary of cross – industry HPA techniques limitations

The previous section described in detail nine quantitative performance analysis techniques which have been broadly applied to several domains. All the techniques have several limitations, which reduce their accuracy and degrade their effectiveness and contribution to overall risk identification. In addition, in many cases such limitations render the techniques cumbersome and difficult to apply. The main collective drawbacks of the existing techniques can be summarised as:

1. *Adequacy of data*: It is acknowledged in the literature that one of the main problems of all human performance analysis techniques is the lack of empirical data for model development and validation. Furthermore, even in the cases where empirical data exists, e.g. THERP methodology, the data have been collected for a particular industry and operation and are therefore not universally applicable. Finally, even when data allow the estimation of error probabilities, there is no agreed methodology of how to interpret the probabilities in terms of safety.

2. *Lack of data*: This, again, has led to the exploration of alternative including subject matter experts (SMEs) and simulators. However, as mentioned in Section 4.10.1.5, both may provide misleading results if not specified appropriately.

   *2a*. With respect to SMEs, there is currently no agreement either on the criteria or the characteristics that SMEs should fulfil (Salmon et al., 2011). Although a generic common practice indicates that the more experienced the SMEs the better (Kirwan, 1994), this does not provide any insights with regard to their characteristics. The differences on the selected panel of experts, within the context of the scenario under consideration, may lead to inconsistent and misleading outcomes due to the biased judgements of SMEs, e.g. frontline vs. managerial personnel. Therefore the SMEs should be selected based on the defined problem, the scope of the analysis and the domain of applicability. For instance, if the scenario assesses train drivers performance, the panel of experts should at least consist of professionals familiar with the tasks of train drivers. In addition the panel should also include professionals familiar with the terms of PSFs and the analysis of train accidents, i.e. human factors experts.

   *2b*. On the other hand, concerning data derived from simulations, there are two main areas of concern. Firstly, it is debatable whether simulation data can replicate empirical data sufficiently. Secondly, simulator observations have limited capabilities concerning the measurement of the factors that affect human performance. For example, the levels of fatigue or workload can be measured with the evolution of technology (Hammerl et al., 2009). However, factors such as motivation of employees or the incentives that organisations provide for their personnel are difficult to be covered. Subsequently, simulators can only assess human performance based on a limited number of measurable factors.

3. *Inadequate proof of the accuracy of techniques*: Due to the lack of empirical data confirmation regarding the validation and accuracy of the techniques is limited (Bell and Holroyd, 2009). Additionally, most of the recent techniques are still under further

development and they are often tailored to a very specific context. Therefore, it is difficult to apply them to a large matrix of scenarios and consequently the process of their validation delays (Kyriakidis, 2009).

4. *Dependencies of PSFs:* Only few of the techniques account for dependencies between PSFs. Therefore, in most of the cases the assessment of performance is based on the direct influence of PSFs, e.g. how fatigue affects performance, without taking into account any indirect contribution of one PSF on another PSF, e.g. how training may influence workload.

5. *Repetitive errors are not taken into account:* The techniques assume that the same human error can happen only once while a task is executed. However, this is not always the case. Very often the same errors are repeated until the occurrence. For instance, a train driver may repeatedly exceed speed limits and manage to avoid SPADs, until the SPAD eventually occurs.

6. *Application of techniques*: Several methods or models can be found in the literature, however only few of them have been cross-industry applied including SLIM, THERP and HEART. Subsequently, the transferability of the techniques to other domains is not widely assessed.

Due to the limitations of the existing cross-industry HPA techniques no single approach will be used per se for the purposes of this thesis. However, the features of the SLIM methodology will be employed as they meet most of the criteria of the suggested methodology, as described in Section 4.10. However, for the implementation of SLIM, the relevant set of PSFs should be first identified. This is thoroughly described in the following chapter.

## 4.11 Summary

This chapter presented the fundamental attributes of human error and human performance. First, the concept of human error in railway operations was determined. Based on this, different approaches used to indicate the origins of human error were reviewed. The results of this review showed that errors should not be treated as isolated events, but rather as results of a sequence of actions and interactions between human and systemic factors. Within this context, it was also shown that human error is associated with degraded human performance, which in turn is affected by a large number of contextual factors, broadly

known as Performance Shaping Factors. Therefore, the human performance should be considered as function of various factors that can affect it.

In view of that, the relevant railway and cross-industry Human Performance Analysis techniques were examined to indicate whether any of the existing approaches could be per se used to study the influence of the PSFs on railway operators' performance. Despite the continuous improvement of these techniques, significant limitations with respect to their accuracy, validity, and included PSFs, were observed, and therefore none of the existing approaches was deemed appropriate to be applied to this thesis.

Therefore, a new approach is proposed, which will allow researchers to overcome the identified limitations. This approach is based on a set of PSFs tailored specifically to railway operations, and accounts for factors that not only affect train drivers' but also signallers' and controllers' performance. This set of factors, referred to as Railway-Performance Shaping Factors (R-PSFs) taxonomy, is introduced in the next chapter. The R-PSFs taxonomy will be then used in Chapter 7 for the development of an approach to study and analyse the impact of the relevant R-PSFs on the performance of railway operators.

# Chapter 5 THE RAILWAY PERFORMANCE SHAPING FACTORS TAXONOMY – LITERATURE REVIEW AND INITIAL DEVELOPMENT

The thesis has so far discussed and justified the significant role humans play in railway operations, and the need for a new method for the assessment of human performance. The remainder of this thesis develops this new method using a new list (or taxonomy) of contextual factors established for the railway domain, referred to as the Railway Performance Shaping Factors (R-PSFs) taxonomy. This chapter initiates this process by presenting the process for the development of this taxonomy.

## 5.1 The characteristics of R-PSFs taxonomy

In the previous chapter (Section 4.9) the limitations of the existing PSFs taxonomies were identified, including: (i) a lack of clear and precise definitions, (ii) poor assessment of the impact of PSFs on human performance, and (iii) assumption of PSFs independency. On account of these limitations, and considering the lack of an agreed taxonomy within the context of the HPA techniques, this thesis develops a new, simple and detailed PSFs taxonomy that focuses on the railway domain. This taxonomy, referred to as the Railway Performance Shaping Factors (R-PSFs), has the aim to enable researchers, irrespective of experience in the field of human factors, to better examine and study human performance in railway operations. Therefore, the main objectives of the R-PSFs taxonomy are to:

- identify the PSFs that influence the performance of railway operators;
- classify PSFs based on their common characteristics and provide precise definitions;
- investigate and account for the dependencies amongst PSFs;
- account for the distinction between dynamic and static PSFs;
- quantify the contribution of each PSF on the performance of operators;
- direct resources more efficiently to develop of effective ways to improve operator performance;

- provide the means to develop a tool for the assessment of human performance in railway operations; and
- account for transferability.

## 5.2 Methodology to develop the R-PSFs taxonomy

The holistic approach captured in Figure 5-1 has been used to determine the set of R-PSFs relevant to railway operations and develop the corresponding taxonomy. In addition, Figure 5-2 illustrates the detailed framework for the development of this taxonomy.

The approach starts with a review of the contextual factors from the most relevant current railway PSFs taxonomies. Due to the limited number of taxonomies for the railway industry, they are augmented with the relevant ones from the other domains to realize a full set of PSFs. Many of these taxonomies are included in the HPA techniques presented in the Chapter 4.

Further to the literature review, 479 complete and detailed worldwide accident and incident investigation reports were analysed. The reports were obtained from 14 different National Transport Safety Boards (NTSBs), National Investigations Bodies (NIBs), Railway Undertakings (RUs) and Infrastructure Managers (IMs). The aim of the analysis was twofold. Firstly, to consolidate the results from the literature in order to identify if any factors should be added, replaced or removed. 192 reports were reviewed to consolidate the results from the literature and the findings were verified by Subject Expert Matters (SMEs). The SMEs were also employed to, indicate the dependencies between R-PSFs. Secondly, the analysis had the aim to introduce a transferable list of R-PSFs that could be applied to more than one country or type of network. Therefore, 287 additional reports were gathered and combined with the first set of reports, and analysed to generate the transferable list of R-PSFs.

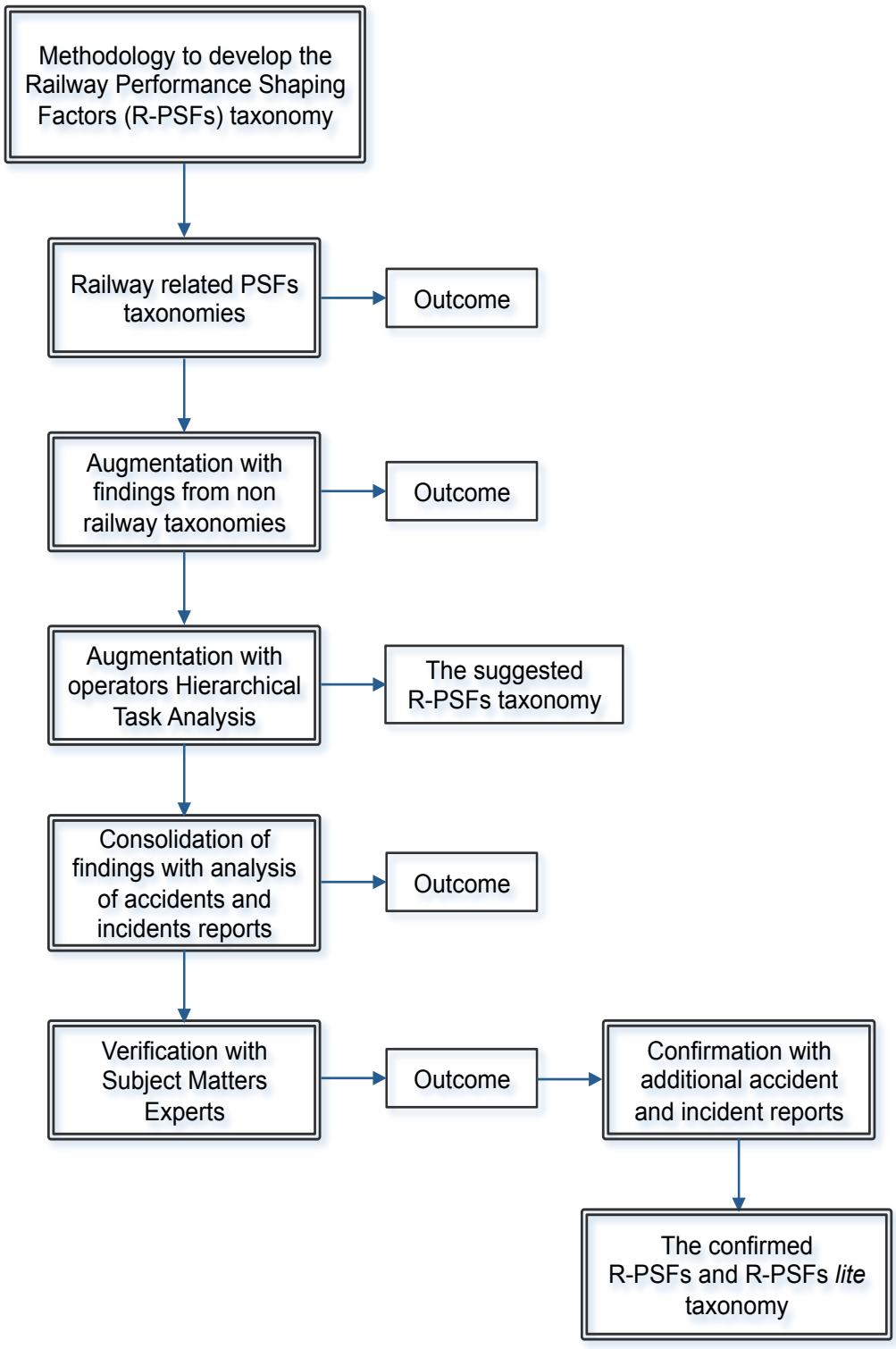



Figure 5-1 Methodology to develop the R-PSFs Taxonomy

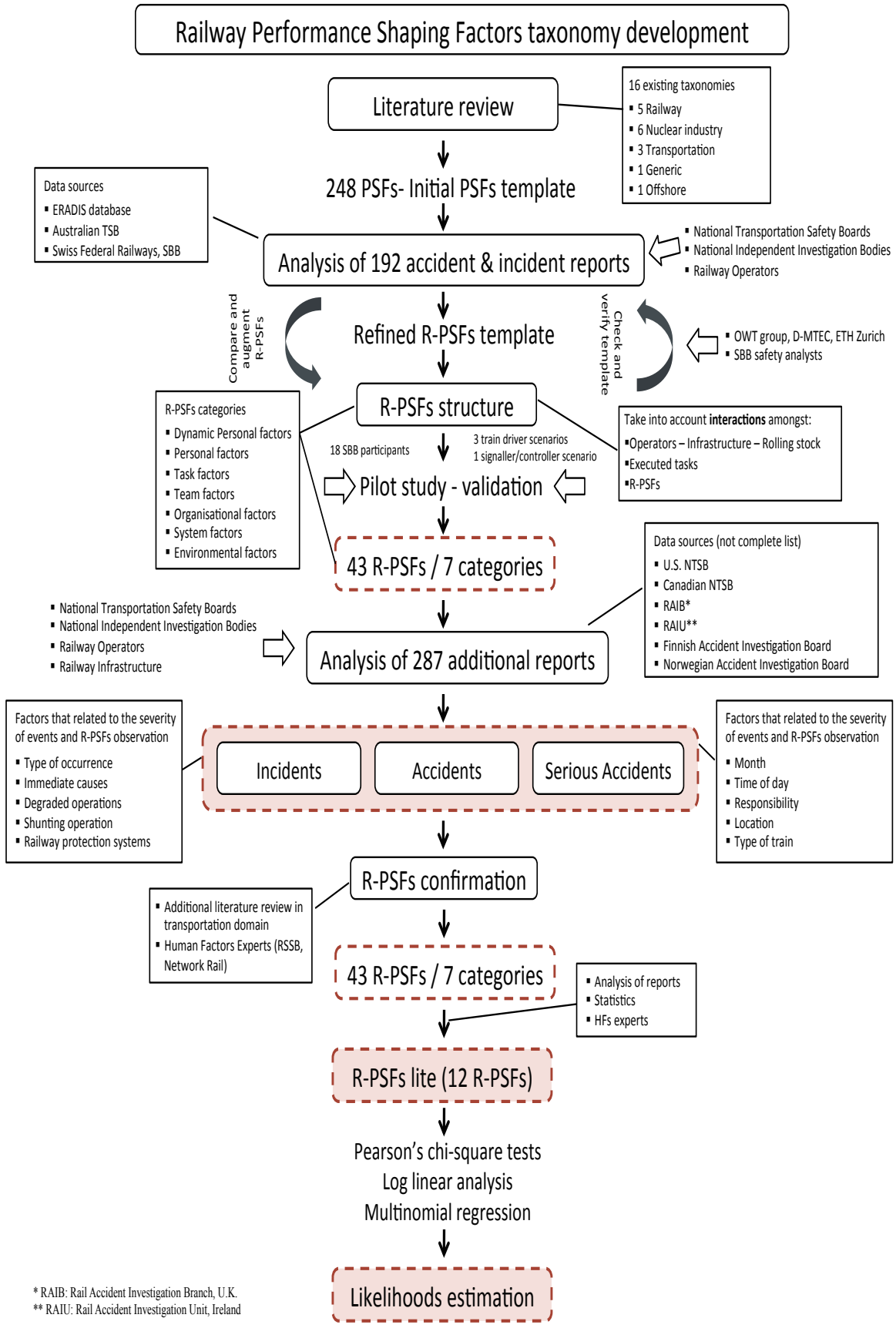



Figure 5-2 Detailed framework of the R-PSFs taxonomy development

## 5.3 Review of existing PSFs taxonomies

The methodology for the selection of the relevant PSFs to railway operators, starts with a review of the factors as identified in the most relevant taxonomies and HPA techniques.

### 5.3.1 Human Factors in the Railway System (TRACEr-Rail)

The TRACEr-Rail (Rail Safety and Standards Board, 2004b), was introduced by the U.K. RSSB as a part of research aimed at developing a rail-specific technique to assess train drivers performance. This technique was not introduced in the previous chapter, as it is a part of the process for the development of the RARA technique. The TRACER-Rail is developed based on the list of factors included in the Technique for the Retrospective Analysis of Cognitive Errors (TRACEr) (Shorrock and Kirwan, 2002), which was established to gain a better understanding of the errors caused by air traffic controllers. Therefore, the starting point for the creation of the TRACEr-Rail was to review the factors included in the TRACEr and exclude those irrelevant to railway operations.

The second step was to re-categorise and rename, when necessary, the remainder of the factors. The original version of TRACEr contains a list of nine main PSF categories: traffic and airspace (e.g. traffic complexity), pilot/controller communications (e.g. workload), procedures (e.g. accuracy), training and experience (e.g. task familiarity), workplace design, Human Machine Interface (HMI) and equipment factors (e.g. radar display), ambient environment (e.g. noise), personal factors (e.g. alertness/fatigue), social and team factors (e.g. handover/takeover) and organisational factors (e.g. conditions of work). After the assessment of the original factors, the TRACEr-Rail taxonomy was derived, and comprises the following nine categories of PSFs: infrastructure and traffic factors; communication factors; procedures and documentation; information; training, knowledge and experience; workplace design, HMI and quality of equipment; in-cab environment; personal factors; and social and team factors. Finally, each of the categories was sub-divided resulting in 29 individual PSFs. Examples of these are distraction, job clarity, procedures clarity, risk perception, unfamiliarity and time availability.

### 5.3.2 The Railway Action Reliability Assessment (RARA)

The RARA (Rail Safety and Standards Board, 2012b, Gibson et al., 2013a) taxonomy was developed primarily to identify the factors that affect the performance of train drivers. It can also be used for other railway personnel as well, e.g. signallers. It introduces a group of six

key categories of PSFs (EPCs[6]): (i) task design, i.e. how a task occurs in practice and the types of demands it places on personnel; (ii) interface design, i.e. the human-machine interface used for completing a task including tools, controls and displays; (iii) competence management, i.e. the quality of training and selection processes which support the task; (iv) procedures, i.e. the quality of procedures and documentation relevant to a task; (v) person, i.e. population of users for the task, and how their perceptions of the task or their mental/physical characteristics might influence their reliability and (vi) environment, i.e. the physical environment (heating, lighting, noise) which may influence human performance. Each of the six categories is broken down further, resulting in 27 individual elements including: unfamiliarity with situation, time availability, unclear allocation of responsibility, operators inexperience, ambiguity of given procedures, fatigue, stress and physical environment. Finally, besides the six categories, the RARA also acknowledges four additional factors, workload, distraction, task complexity and unavailability of staff, not been included in the initial list because they are associated with more than one of the existing EPCs.

### 5.3.3 Human Factors Analysis Classification System Railways (HFACS-RR)

The HFACS-RR (Reinach and Viale, 2006) was developed on the principles of HFACS (Shappell and Wiegmann, 2000) to systematically examine underlying human causal factors and to improve railway accident investigations. The approach is similar to the Swiss Cheese model, hence it identifies human error as a sequence of events which is expressed by the unsafe actions of the frontline personnel. However, in contrast to the Swiss Cheese model, the HFACS-RR introduces a set of contributing factors that affect the performance of operators and subsequently increase the likelihood of the occurrence of errors. This set of factors consists of five main categories: conditions of operators (physical and psychological limitations); environmental factors (physical and technological environment); personal factors (training, staffing), organisational factors (supervision, safety culture, procedures, rules and regulations), and outside factors (regulatory oversight and political/social environment). However, by definition *outside factors* are considered exceptional and rare. Therefore, they are excluded from the set of candidate R-PSFs.

[6] As stated in Chapter 4, RARA approach uses the term Error Performance Conditions (EPCs) when referring to PSFs.

### 5.3.4 Human Error Risk Management for Engineering Systems (HERMES)

The HERMES (Cacciabue, 2004a) is a qualitative generic approach to enhance the identification of potential human errors, and is strongly influenced by the THERP methodology. Within its concept the HERMES provides a set of Performance Influencing Factors (PIFs), which have been successfully applied to the railway domain. The PIFs are: communication within organisation, communication means, technological interfaces, maintenance of trains/railway, comfort of working contexts, roster and shifts planning, regulations/rules, training methods and simulators. These are further sub-divided into indicators of safety, including amongst workload, stress, training, quality of equipment and supervision.

Due to the fact that this thesis does not account for occurrences related to maintenance, only the factor *maintenance of trains/railway* is excluded from the candidate set of PSFs.

### 5.3.5 Human Factors in the Railway System Safety Analysis Process

The model of Hammerl and Vanderhaegen (2009) provides a list of PSFs that was derived and restructured on the basis of the PSFs included in the THERP and TRACE-Rail taxonomies. As shown in the Chapter 4, the model addresses three main categories of factors: physical, personal and organisational. The physical factors are sub-divided into: (i) anthropometry, i.e. basic layout of working environment; (ii) working conditions, i.e. physical conditions as temperature, humidity, light, noise; and (iii) design of HMI, e.g. layout of HMI, quality of feedback. The personal factors comprise individual and dependent factors, with the former including stress, age, gender and health, and the latter including experience, motivation, safety awareness and fatigue. Finally, the organisational factors are classified into: (i) employee related factors, e.g. roster planning, training, safety culture and (ii) standard factors, e.g. rules and guidelines. Although, this list is not exhaustive, the model allows researchers to enhance the taxonomy by adding any factors relevant to a given task or scenario.

The five contextual factors taxonomies presented above (TRACEr-Rail, RARA, HFACS-RR, HERMES, HFs in Railways) were either modified or developed for the railway environment. In general, they consider PSFs as a set of factors that can either enhance or degrade human performance and subsequently affect the likelihood of errors occurrence. The assessment of the five models results in a total of twelve PSFs categories, as shown in the Table 5-1, and more than 75 individual elements (exhaustive list in Appendix III).

The PSFs categories are: personal factors; team factors; task factors; organisational factors; environmental factors; infrastructure and traffic; procedures and documentation; training and experience; communication and information; roster and shift patterns; interfaces and finally working environment. On the other hand, the individual contextual factors include amongst others stress, training, safety culture, design of workplace, supervision, familiarity, experience, quality of equipment, fitness to duty and fatigue.

Table 5-1 Review of railway contextual factors taxonomies

| Taxonomy | Terminology of contextual factors | High-level contextual factors | Elements of contextual factors (not exhaustive list) |
|---|---|---|---|
| TRACEr-Rail | ▪ Performance Shaping Factors (PSFs) | ▪ Infrastructure and traffic<br>▪ Communication<br>▪ Procedures and documentation<br>▪ Information<br>▪ Training, knowledge and experience<br>▪ Workplace design, HMI<br>▪ In-cab environment<br>▪ Personal<br>▪ Social and team | ▪ Distraction<br>▪ Time availability<br>▪ Workload<br>▪ Risk perception<br>▪ Experience<br>▪ Procedures clarity<br>▪ Quality of information<br>▪ Stress<br>▪ Health<br>▪ Fatigue<br>▪ Team relations |
| RARA | ▪ Error Producing Conditions (EPCs) | ▪ Task design<br>▪ Interface design<br>▪ Competence management<br>▪ Procedures<br>▪ Person<br>▪ Environment | ▪ Unfamiliarity<br>▪ Time availability<br>▪ Quality of information<br>▪ Poor system feedback<br>▪ Experience<br>▪ Quality of information<br>▪ Fatigue due to shift pattern<br>▪ Workforce morale |
| HFACS-RR | ▪ No terminology | ▪ Conditions of operators<br>▪ Personal factors<br>▪ Organisational factors<br>▪ Environmental factors | ▪ Mental state<br>▪ Physiological state<br>▪ Crew resource management<br>▪ Technological environment<br>▪ Supervision<br>▪ Planning of operations<br>▪ Organisational climate |
| HERMES | ▪ Performance Influencing Factors (PIFs) | ▪ Communication within organisation<br>▪ Communication means<br>▪ Technological interfaces<br>▪ Comfort of work context<br>▪ Roster/shifts planning<br>▪ Regulations/rules<br>▪ Training methods and simulators | ▪ Unwritten rules<br>▪ Mental conditions<br>▪ Quality of tools/actuators<br>▪ Workplace design<br>▪ Stress<br>▪ Physical conditions<br>▪ Training<br>▪ Supervision<br>▪ Procedures |
| HFs in Railways | ▪ Performance Shaping Factors (PSFs) | ▪ Physical<br>▪ Personal<br>▪ Organisational | ▪ Emotional tension<br>▪ Training<br>▪ Safety culture<br>▪ Task design<br>▪ Experience<br>▪ Physical conditions |

However, as can be seen in the Table 5-1, the reviewed taxonomies do not include categories of factors at the same level of detail. For instance, the *HFACS-RR* and *HFs in Railways* taxonomies include categories that in general refer to individuals, organisations and working conditions, while the remaining are more specific, containing categories such as procedures, communication and technological interfaces. Subsequently, the categories of the reviewed taxonomies have a certain degree of overlap. The factors category *procedures and documentations* for example, is related and could be included within *organisational factors*, while *conditions of operators* can be accommodated within *personal factors*.

Therefore, this thesis introduces an alternative classification of the PSFs categories to bridge the gap due to the different perspectives amongst existing taxonomies. This list includes six categories of factors, which embrace all the identified categories according to their common characteristics. The suggested classification consists of:

(i) *personal* factors, which includes the factors that characterise and affect the performance of individuals, e.g. individual characteristics, conditions of operators, training and experience;

(ii) *task* factors, that characterise the features of an executed task e.g. task design;

(iii) *team* factors, e.g. communication between team and social climate;

(iv) *organisational* factors, which include the factors defined and controlled by a railway organisation. Such factors embrace the characteristics and attitudes of an organisation, as well as certain organisational behaviours that influence the performance of employees (Groth, 2009); e.g. procedures and documentation, training methods, roster and staffing;

(v) *system* factors, which characterise the attributes of the railway system and how the interactions between operators and system characteristics may affect operator performance at, e.g. interfaces, infrastructure layout and working environment;

(vi) *environmental* factors, that refer to weather conditions at the time the task is executed.

The review of railway related PSFs taxonomies provides a large set of high-level and detailed contextual factors relevant to railway operations. Nevertheless, it is worth exploring whether other factors exist within relevant approaches and methods developed for non-railways operations. The following sections present an overview of the relevant findings.

### 5.3.6 Technique for Human Error Rate Prediction (THERP)

The THERP taxonomy was developed to identify the factors that influence the performance of operators in the control rooms of nuclear power plants (Swain and Guttmann, 1983). It is acknowledged as the first and most detailed taxonomy including more than 60 PSFs (Kirwan, 1994). The taxonomy classifies all factors into external, internal, and those that act as psychological and physiological stressors. Further to the initial classification, the THERP sub-divides external PSFs into situational characteristics, task and equipment characteristics, and work and task instructions. The situational characteristics describe generic PSFs that can be found in more than one job in a work situation. The task and equipment characteristics describe all PSFs that are specific to tasks in a given work, while the task instructions involve the PSFs, which are important tools for most of the tasks. The internal factors are defined as those related to the individual, i.e. characteristics of people resulting from internal and external influences. Finally the two types of stressors refer to those PSFs that directly affect mental and physical stress. The PSFs included in the THERP taxonomy are presented in Table 5-2.

Table 5-2 Factors influencing human actions in THERP (Kirwan, 1994)

| Category of PSFs | Performance Shaping Factors |
|---|---|
| **External PSFs** | |
| Situational characteristics | Architectural features; Quality of environment (temperature, air humidity, air quality, radiation exposure, lighting, noise, vibration, cleanliness); Working hours; Work breaks; Availability of special work resources; Manning; Organisational structure (authority, responsibility, communication channels); Actions by supervisors, co-workers, union representatives; Remuneration incentives (recognition, rewards, benefits) |
| Work and task instructions | Required procedures (written, non-written); Written and verbal communication; Cautions and warnings; Work-methods; Plant policy |
| Task and equipment characteristics | Perceptual requirements; Requirements for motor system (speed, power expenditure, accuracy); Relationship between operators and display; Anticipatory requirements; Interpretation; Decision making; Complexity (information loading); Narrow nature of task; Short term and long term memory; Frequency and repetitiveness; Task criticality; Calculations; Feedback (knowledge regarding results of an action); Dynamic of gradual actions; Team structure and communications; Man-machine interface factors (design of work resources, test instruments, maintenance equipment, work aids, tools, accessories) |
| **Internal PSFs** | |
| Organismic factors | Previous training / experience; State of current practice or abilities; Personality and intelligence variables; Motivation and attitudes; Emotional states Stress (mental or physical); Knowledge of required performance standards; Gender differences; Physical conditions; Attitudes deriving from family, social environment; Group identifications |
| **Stressors** | |
| Psychological | Suddenness of occurrence; Duration of stress; Task speed; Task load; High hazard |

| Category of PSFs | Performance Shaping Factors |
|---|---|
| stressors | risks; Threats (fear of failure, loss of job); Monotony, degrading or meaningless work; Duration of uneventful vigilance periods; conflicts or motives about work performance; Reinforcement absent; Distractions (noise, glare, movement, flicker, coloration); Inconsistent cues |
| Physiological stressors | Duration of stress; Fatigue; Pain or discomfort; Hunger or thirst; Extreme temperatures; Radiation; Extreme gravitational forces; Extreme atmospheric pressure conditions; Oxygen insufficiency; Vibration; Restricted movements; Absence of physical exercise; Disruption of circadian rhythm |

A review of the PSFs within the THERP taxonomy reveals that most can be allocated to the factors identified by the first five railway related approaches. Several other factors, such as *decision-making* may be classified as personal, while *monotony* could be characterised as related to the executed task. However, factors such as extreme G-forces, extreme atmospheric conditions, oxygen insufficiency and extreme gravitational forces, by nature, cannot be identified within the factors that affect railway operators performance. Therefore, they are excluded from the set of relevant to railways PSFs.

### 5.3.7 Human Error Assessment and Reduction Technique (HEART)

The HEART taxonomy (Williams, 1986) is considered generic and has been extensively used in the nuclear, chemical and petrochemical industries. The approach defines a set of Error Producing Conditions (EPC) including ergonomic, task and environmental factors that could negatively affect human performance. It identifies 38 different EPCs (Rail Safety and Standards Board, 2012b), which can be classified into two main groups, directly transferable to railways and non-transferable. The EPCs relevant to the railway industry are those already introduced in the RARA taxonomy. The EPCs not relevant to railways include several factors. For example, the factor "*inconsistency of meaning of displays procedures*" is not applicable in railway operations. Furthermore, the EPC "*incentive to use other more dangerous procedures*" is also not applicable to railways as "*dangerous procedures*" imply direct violations of the rules. Such factors are excluded and not taken into account in the candidate set of R-PSFs.

### 5.3.8 Human Event Repository and Analysis (HERA)

The HERA (Hallbert et al., 2007) taxonomy was developed within the HERA project to analyse and gather information about human performance from commercial nuclear power plants. The taxonomy introduces eleven PSFs, which were obtained by merging the PSFs included in the SPAR-H taxonomy and in the study "*Good practices for Implementing*

*Human Reliability Analysis*" (Kolaczkowski et al., 2005). The eleven PSFs are: available time, stress and stressors, experience and training, complexity, procedures and reference documents, ergonomics and HMI, work processes, fitness for duty/fatigue, environment, and team dynamics/characteristics, and communication. These categories are transferable to railways, and therefore, included in the list of candidate R-PSFs.

### 5.3.9 A Technique for Human Error Analysis (ATHEANA)

The ATHEANA (Cooper et al., 1996) provides a list of factors that have an impact on the performance of operators in nuclear power plants. Similar to other approaches, ATHEANA introduces a generic set of factors, such as: training and experience; procedures; availability and clarity of instrumentation; time pressure; workload and stress; complexity of required actions and familiarity with the situation; dynamics and characteristics of team; staffing; ergonomics, quality and human-machine interface; communications; environmental and working conditions; and personal fitness level. All the PSFs within ATHEANA are relevant to railway operators and operations, and therefore, included in the list of the R-PSFs taxonomy.

### 5.3.10 Controller Action Reliability Assessment (CARA)

The CARA taxonomy (Gibson and Kirwan, 2008a) was developed for the assessment of human performance in the aviation industry and in particular in Air Traffic Management (ATM), based on the principles of HEART. It introduces 22 EPCs obtained from the review of several existing taxonomies, including amongst others the HEART, the SPAR-H and the CREAM. Similar to the EPCs in the HEART taxonomy, the EPCs included in CARA are again classified into the two main groups of transferable and non-transferable. The EPCs relevant to railways are the same as in the HEART and RARA taxonomies, and are included in the candidate R-PSFs.

### 5.3.11 Human Error in Air Traffic Management (HERA-JANUS)

HERA-JANUS (EUROCONTROL, 2003) represents another approach for the analysis of human error in the ATM. It was developed from an extensive literature review and operational engagement of air traffic controllers, incident investigators, and safety managers (Subotic, 2007). It has a set of eleven categories of factors, referred to as Contextual Conditions (CCs). These are: pilot-controller communications, pilot actions, traffic & airspace, weather, documentation & procedures, training & experience, workplace design &

HMI, environment, personal factors, team factors, and organisational factors. Each of the CC groups is sub-divided, resulting in more than 200 individual contextual factors.

All, but one of the CCs defined in the taxonomy are relevant to railway operators and operations. Most are directly transferable to railways, while others, i.e. *pilot/controller communication* and *traffic & airspace* require renaming and redefinition. The category *pilot actions* is removed as non-relevant. All the remaining factors (including where modification is required) are included in the set of candidate R-PSFs.

### 5.3.12 Cognitive Reliability and Error Analysis Method (CREAM)

The CREAM (Hollnagel, 1998) taxonomy provides a generic list of factors applicable to any industry. Based on prior research (i.e. THERP), it has the goal of delivering a more structured taxonomy where related categories of contextual factors are assembled together. It has a small set of factors, referred to as Common Performance Conditions (CPCs), that include the general determinants of performance, such as (Hollnagel, 1998):

- adequacy of organisation, e.g. the quality of the roles and responsibilities of team members, additional support, communication systems, Safety management Systems;
- working conditions, e.g. the nature of physical working conditions such as ambient lighting, glare on screens, noise from alarms;
- adequacy of Man-Machine Interface (MMI) and operational support, e.g. the MMI in general, available information on control panels, computerised workstations;
- availability of procedures/plans, e.g. operating and emergency procedures, familiar patterns of response, routines;
- number of simultaneous goals, e.g. the number of tasks a person is required to pursue or attend to at the same time;
- available time, e.g. the time available to carry out a task;
- time of day (circadian rhythm), e.g. the time at which the task is carried out;
- adequacy of training and preparation, e.g. level and quality of training provided to operators and operational experience;
- crew collaboration, e.g. teamwork, level of trust, social climate.

All the CPCs are transferable to railways, and are therefore, included in the development of the R-PSFs taxonomy.

### 5.3.13 Human Reliability Management System (HRMS)

The taxonomy included in the HRMS approach (Kirwan, 1997) was developed to study the factors that affect human performance in the nuclear industry. Six PSFs, derived based on the assessment of existing taxonomies (Kirwan, 1997), are provided. These are: procedures, time, quality of information and interface, training/expertise/experience/competence, task organisation and task complexity.

The PSFs are all relevant to railways, and are included to the candidate set of R-PSFs for the railway taxonomy.

### 5.3.14 Simplified Plant Analysis Risk Human Reliability Assessment (SPAR-H)

The SPAR-H (Gertman et al., 2004), introduces a list of contextual factors for the estimation of human performance in nuclear power plants. Nine PSFs are identified, derived from an extensive literature review in the field of behavioural sciences (Gertman et al., 2004). These PSFs are: available time; stress and stressors; training and experience; task complexity; ergonomics - Human Machine Interface; procedures; fitness for duty; and work processes, i.e. organisational factors, safety culture, work planning, communication.

All the factors are relevant to railway operators and operations and thus are included in to the candidate set of R-PSFs for the railway taxonomy.

### 5.3.15 Success Likelihood Index Methodology (SLIM)

The SLIM (Embrey, 1983) approach was developed to quantify the PSFs that affect human performance in nuclear power plants. In contrast to the other taxonomies, SLIM does not provide a fixed set of PSFs. It allows analysts to identify the factors related to individuals, environment and task, which are considered most important to the scenario under investigation. Nevertheless, the literature (Forester et al., 2006) indicates that there are at least 14 PSFs within the SLIM approach that could be used as template by the researchers. These are: training and experience; procedures and administrative controls; availability and clarity of instrumentation; time available and time required to complete the task; complexity of the required diagnosis and response; workload, time pressure and stress; team dynamics and characteristics; staffing; quality of the human-system interface; environmental factors;

accessibility and operability of the equipment to be manipulated; need for special tools; communication within organisation and special fitness needs.

All but three of the PSFs are directly transferable to railways. From the remaining three, two, i.e. the "*need for special tools*" and "*accessibility and operability of the equipment to be manipulated*" are not related to railway operations. The third, i.e. "*complexity of the required diagnosis and response*", is primarily related to the cases of emergency, hence it is incorporated within existing factors, e.g. workload and time to response to a task.

### 5.3.16 Human Error Probability Index (HEPI)

The HEPI taxonomy provides a set of PSFs for the petrochemical industry. It is built upon the principles of SLIM, and allows researchers to define the factors related to individuals, environment and task, which are considered to the most important for the scenario under investigation. HEPI has a list of nine PSFs: procedures, stress, complexity, time pressure, teamwork, training, experience, distraction and environmental factors. All the HEPI factors are directly transferable to railway operations.

### 5.3.17 Summary of findings

The candidate set of Railway Performance Shaping Factors (R-PSFs) has been selected on the basis of several sources of information. In general, the contextual factors throughout the assessed taxonomies show strong similarities, where contextual factors are considered as conditions or factors that influence human performance. The process followed to select the relevant R-PSFs started with an initial selection based on the review of contextual factors identified in five railway related taxonomies (Table 5-1). As a result, six PSFs categories were determined as relevant: personal, task, team, organisational, system and environmental factors, which embrace all the generic categories of factors as identified by the existing taxonomies. In addition, more than 75 individual PSFs relevant to railway operations were identified.

The initial findings were augmented with a review of non-railway related taxonomies. This phase involved the review of eleven well-documented and widely implemented taxonomies designed primarily to analyse human errors in the nuclear, aviation and process industries (Kirwan, 1994, Forester et al., 2006, Bell and Holroyd, 2009). The literature shows that taxonomies have a tendency to be repetitive, i.e. include the same PSFs (Wilson et al., 2007a). Therefore, although the selected taxonomies may not be exhaustive, it can

reasonably be argued that they contain the majority of the factors that influence human performance in every working environment.

The assessed taxonomies and their related factors are presented in tabular form in Appendices II and III. Due to the fact that there is no agreed taxonomy for the railway industry, taxonomies from other domains were analysed and compared with those within the railway domain. The comparisons were conducted on two levels, firstly between the categories of PSFs and secondly, amongst the individual PSFs. The comparisons were mostly straightforward as particular factors were identified in almost all taxonomies, e.g. the factors "*procedures*" and "*training*". The comparison on the first level did not reveal any categories that should be added to the six categories as obtained from the railway taxonomies. However, this was not the case for the comparison on the second level, as six R-PSFs elements could not be directly identified as belonging to any of the railway related taxonomies. These factors are: *routine*, *expectation*, *vigilance*, *interpretation*, *decision-making skills* and *Safety Management Systems* (SMS). However, routine is indicated as an underlying factor within the RARA taxonomy. In addition, the factors vigilance, interpretation, decision-making skills and expectation are defined within the TRACEr-Rail approach as operational error types. Finally, although not identified in any railway relevant taxonomy the factor SMS is clearly included within the CREAM taxonomy. Therefore, they are all retained within the set of candidate R-PSFs due to their relevance. The final list of candidate R-PSFs contains more than 250 elements, as shown in the Appendix III.

## 5.4 Identification of dynamic R-PSFs

In Chapter 4 it was stated that the relationship between PSFs and time has not been addressed adequately. This is confirmed from the results in Section 5.3 where no taxonomy refers directly to the *dynamic* perspective of PSFs. However, many factors may gradually or rapidly change while a task is executed, e.g. distraction, stress or fatigue. These factors can be claimed to have a direct impact on human performance just before an error is made and subsequently an accident or incident occurs. To address this issue, this thesis introduces an additional R-PSFs category, referred to as *dynamic personal factors*. It describes all the factors that characterise and affect the individuals and are *strongly* related to the precise moment of operation/occurrence, e.g. distraction. In addition, environmental R-PSFs are considered as dynamic factors.

The remaining categories are considered as *static* factors, i.e. the PSFs that have less association with the time of the occurrence. However, under certain conditions, e.g. an

accident on the network, factors such as the workload and time pressure may be affected significantly by the precise moment of the operation/occurrence. For these conditions the factors are considered dynamic.

## 5.5 Augmentation with operators hierarchical task analysis

After the identification of the R-PSFs, as derived from the relevant literature, the task analysis of operators was again reviewed. The purpose of this review is twofold. First, to map each of the factors with the duties of operators so as to confirm that the extracted R-PSFs may indeed potentially affect operators' performance. Second, to identify whether any factors have been either ignored, neglected or do not match to the operational concept of railways. The review focused particularly on the factors that affect the signallers and controllers performance as the majority of the railway approaches are built on train drivers' performance.

The mapping showed that all identified R-PSFs are relevant to the railway operational concept. Furthermore, no factors were conspicuously identified as missing or ignored. For train drivers this outcome was rather expected since the majority of relevant taxonomies were developed on their tasks. On the other hand, with respect to signallers, two are the most likely reasons for such observation. The first is that the number of PSFs is limited regardless of the task or operational procedures (Kirwan, 1994). The second is that within a generic operational concept it can be claimed that the duties of operators in control rooms of industries other than railways, e.g. aviation, have many similarities to the duties of train signallers and controllers. For instance, both air traffic controllers and railway signallers are responsible to: maintain the scheduling of services, contact with either pilots or train drivers when necessary, and prevent any type of conflicts. Subsequently, it is assumed that the majority of the identified factors from the review of relevant non-railway taxonomies cover sufficiently the factors that affect the performance of railway signallers and controllers.

## 5.6 Presentation of the R-PSFs taxonomy

Having reviewed sixteen PSFs taxonomies and the tasks of railway operators, the R-PSFs taxonomy was finally created. The review revealed seven R-PSFs categories and more than 250 R-PSFs elements. However, many of the individual elements are either identical or overlap. Therefore, they have been combined based on their common characteristics and

when necessary renamed and redefined, e.g. the R-PSFs "*Railway communication means*", instead of "*Communication equipment*" as indicated in the HERA taxonomy.

Finally, 43 R-PSFs were identified and allocated to each one of the seven R-PSFs categories as displayed in the Table 5-3. Furthermore, detailed definitions (Appendix IV) to both categories of factors and the elements are provided in order not only for the taxonomy to be implemented by researchers with different backgrounds and perspectives, but also for the derived results to be comparable. Definitions aim to be short and transparent, so as to be easily applied from both academic researchers and practitioners. Unless otherwise stated, definitions were derived from the Collins English Dictionary (2009). Finally, constructive examples are also given (Appendix IV), which aim at supporting researchers to better understand the context within each of the R-PSFs should be used.

Table 5-3 The complete R-PSFs taxonomy

| Category of R- PSFs | Railway Performance Shaping Factors |
|---|---|
| Personal Factors | ▪ Experience<br>▪ Familiarity<br>▪ Fit to work (health)<br>▪ Individual characteristics<br>▪ Motivation<br>▪ Training - competence |
| Dynamic Personal Factors | ▪ Decision making skills<br>▪ Distraction-loss of concentration<br>▪ Expectation<br>▪ Fatigue<br>▪ Interpretation<br>▪ Perception<br>▪ Situational awareness<br>▪ Stress<br>▪ Vigilance |
| Task Factors | ▪ Monotony<br>▪ Routine<br>▪ Task complexity<br>▪ Task instructions<br>▪ Time pressure - time to respond<br>▪ Workload |
| Team Factors | ▪ Communication between employees<br>▪ Relations within team<br>▪ Teamwork<br>▪ Trust in information |

| Category of R- PSFs | Railway Performance Shaping Factors |
|---|---|
| Organisational Factors | ▪ Communication within organisation<br>▪ Fit to work aspect<br>▪ Incentives for employees<br>▪ Leadership<br>▪ Quality of procedures, standards and regulations<br>▪ Relations within organisation<br>▪ Safety culture (disregard procedures)<br>▪ Safety Management Systems<br>▪ Shift pattern (working hours, breaks, manning)<br>▪ Supervision<br>▪ Training / training methods |
| System Factors | ▪ Human Machine Interface (HMI)<br>▪ Railway Communication Means (RCMs)<br>▪ System design<br>▪ Trust in equipment<br>▪ Working environment |
| Environmental Factors | ▪ Visibility<br>▪ Weather conditions |

In addition, Figure 5-3 illustrates the structure of the taxonomy by showing the interactions between operators and R-PSFs.

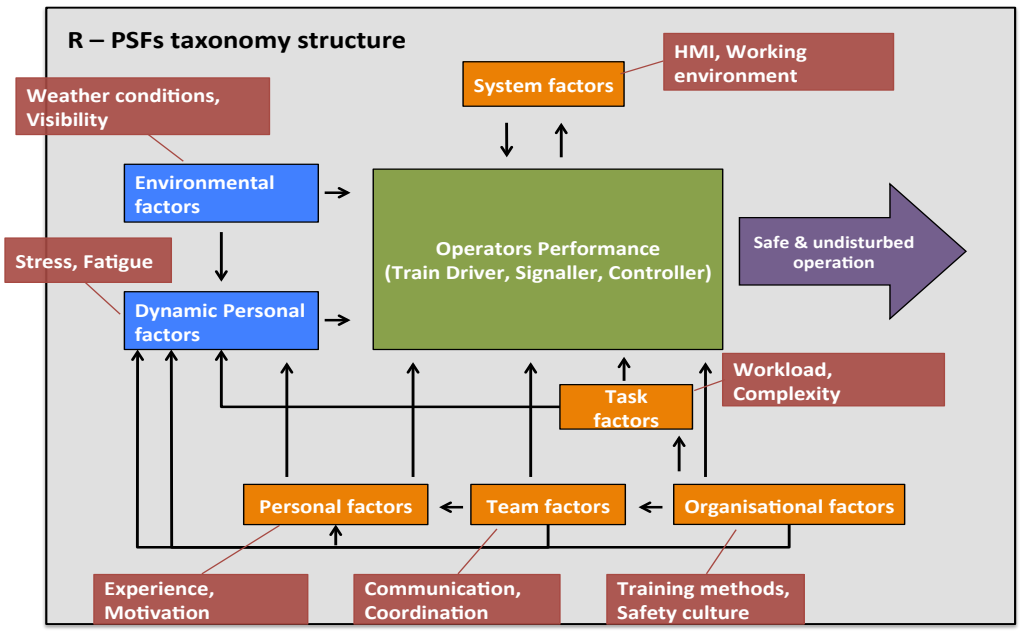



Figure 5-3 The structure of R-PSFs taxonomy

The structure displays the engineering and psychologists perspectives of the system and has been established on the generic model introduced by Hammerl and Vanderhaegen (2009) (Section 4.10.1.3). The structure clearly indicates how the performance of operators

is influenced by the different R-PSFs as well as how the R-PSFs interact with each other. As shown in Figure 5-3, blue factors reflect the dynamic R-PSFs, while orange represent the static. Although the structure was developed for the railway industry, it can be claimed generic, as it stems from the review of taxonomies from several domains. Therefore, it can be argued that both the structure model of the R-PSF taxonomy as well as the included R-PSFs categories are transferable to any other transport mode or industry. However, the individual R-PSFs might change according to the attributes and features of the domain or/and industry in which they will be implemented.

Finally, it is clear from the taxonomy's structure and the provided definitions that R-PSFs not only directly but also indirectly affect human performance. That is when a factor A may influence a factor B and subsequently affects operators level of competence. For instance, unclear procedures may increase operators' levels of workload and stress and consequently indirectly influence their performance. The interdependencies among R-PSFs are in detail described in Section 6.3.3.

Having identified the suggested R-PSFs taxonomy and its constituents, a validation process based on real data and Subject Matter Experts (SMEs) consultation was than conducted as described in the remainder of this chapter and Chapter 6.

## 5.7 Consolidation of taxonomy with real data

To consolidate the taxonomy, that is, crosscheck, verify, and confirm literature findings with real data, worldwide human error related accident and incident investigation reports were analysed. The literature shows that up to 80-90% of all serious railway accidents are attributed to human error due to degraded human performance (Rail Safety and Standards Board, 2004b). Thus, this thesis focuses on the study of factors that affect the performance of operators resulting in major accidents or high-risk incidents. Therefore the selected reports describe events that by legislation should be investigated. In Chapter 3 it was shown that different definitions are used to describe occurrences internationally. The European (European Commission, 2004) definitions are deemed to be more comprehensive, hence are adopted for the purposes of this thesis. Subsequently, all reports are classified into serious accidents, accidents and incidents. The reports describe occurrences relevant to operational errors caused by train drivers, signallers or controllers while trains move from origin to destination (including shunting operations). Railway occurrences caused by maintenance personnel are not included in the analysis.

## 5.8 Data characteristics

In total 2278 publicly available and 418 confidential reports were gathered from several worldwide stakeholders. Reports were studied and only those describing human error related operational occurrences were finally retained. Subsequently, 479 reports were selected and classified according to the EU Directive 2004/49/EC (European Commission, 2009), with 70 serious accidents, 189 accidents and 220 incidents. The reports describe events that occurred in the period 1997-2011 in 24 different countries includings 21 European countries such as U.K., Norway, Germany and Switzerland; U.S.A; Canada and Australia. The timeframe was decided upon consulting summaries of railway activities issued by institutions of national and international reach, e.g. the International Union of Railways (International Union of Railways, 2007), the Federal Railroad Administration (Federal Railroad Administration, 2012) and the European Railway Agency (European Railway Agency, 2008). Data represent occurrences only from countries with highly developed railway systems. Undoubtedly, it would be beneficial if data from developing countries could also be gathered; however, this was not possible due to access and resources limitations.

## 5.9 Data requirements

In Chapter 4 was highlighted the importance of examining human error by accounting for the overall context within which an error occurs, including any underlying or hidden factors that may affect operators performance. Hence, the selected reports should contain a limited amount of information that allows researchers to analyse in detail the causes of a human error. Such information contains the:

- description of the event;
- involved personnel (responsibility), e.g. individual characteristics;
- type of railway operation, e.g. passenger (intercity or regional), freight or metro;
- features of system, e.g. type of train, on-board and infrastructure equipment;
- severity of event, e.g. accident or incident;
- type of occurrence, ,e.g. train derailment or trains collision;
- location and time of event, e.g. open line, station;
- immediate cause, e.g. train unable to stop;
- causal factors, e.g. train driver falls asleep;
- contributing factors (R-PSFs) that affect operators performance, e.g. adverse weather conditions, inadequate training, communication errors, and

- number of casualties or material damages, which defines the severity of events.

To capture such level of information, the analysis presented in this thesis is based on detailed official railway investigation reports, as descriptive reports contain limited amounts of information (particularly with respect to the human factors aspect). The importance of analysing detailed reports also arose from the author's personal communication with two railway organisations, i.e. the Swiss Federal Railways (SBB) and the UK Network Rail. Both the SBB (Swiss Federal Railways, 2010) and Network Rail (Gibson et al., 2013b), have acknowledged the need to collect specific information with respect to human factors as significant contributors to railway occurrences.

Although the selected reports are official documents completed by authorised and eligible personnel, their quality was verified to ensure the reliability and robustness of the analyses. Therefore certain quality criteria were considered and examined, as presented in the following section.

## 5.10 Data quality criteria

The literature (Wang and Strong, 1996, Brackstone, 1999, Dupuy, 2011) identifies up to eight criteria (dimensions) that can be used to test the quality of collected data. These are: accessibility, accuracy, completeness, consistency, credentials, interpretability, relevance, and timeliness. All eight criteria are described in more detail in this section, while their summary is shown in Table 5-4.

Table 5-4 Data quality dimensions review (Dupuy, 2011)

| Quality dimension | Characteristics | Related notions |
|---|---|---|
| Accessibility | Availability of data to the user | Cost, availability |
| Accuracy | Extent to which data values are correct, close to reality and free from error | Correctness |
| Completeness | Quantity of reported data for each unique record | Missing values |
| Consistency | Extent to which data values are coherent and compatible with one another within their corresponding framework and consistent with the framework definition | Coherence, regularity, definitions, validity conformance |
| Credential | Extent to which data are obtained from reliable sources | Reliability |
| Interpretability | Extent to which data are completed or perceived in an unambiguous manner | Clarity |
| Relevance | Suitability of the reported data for a specific domain and absence or essential data | |

| Quality dimension | Characteristics | Related notions |
|---|---|---|
| Timeliness | Currentness of the data in terms of the delay in providing and processing the information | Up-to date |

### 5.10.1 Accessibility

*Accessibility* refers to the ease with which data are obtained by any user. This encompasses the extent of ease to which the existence of this data can be ascertained and the suitability of mediums through which the data can be retrieved (Dupuy, 2011). The cost of the information may also be an aspect of accessibility for a number of users (Brackstone, 1999). Regarding railway investigation reports, accessibility is associated with confidentiality considerations. That is the willingness of railway organisations to share and allow access to their data (Dupuy, 2011). For this research, all collected investigation reports refer to occurrences that railway organisations are obliged to report according to the relevant safety regulations. Subsequently the majority of reports are publicly available and railway organisations, i.e. train operators, infrastructure managers, national investigation bodies, were not involved in the collection process. In addition, two of the databases were collected directly from an infrastructure manager and an integrated company (train operator also being an infrastructure manager) on a confidentiality agreement. Thus, accessibility is not relevant as quality dimension for the reports collected for this research. However, accessibility should be considered as quality criterion if access to data of developing countries was tested.

### 5.10.2 Accuracy

Data *accuracy* accounts for the degree to which data correctly describe the elements they are aimed to measure (Dupuy, 2011). Accuracy is often expressed in terms of error in statistical estimates and is classified into bias (systematic error) and variance (random error) components (Brackstone, 1999). It may also be defined in terms of the main sources of error that could potentially lead to inaccuracy (e.g., coverage, sampling). For this research, due to the format of data, accuracy is not considered as quality dimension, since all investigation reports correctly describe the elements they are aimed to measure.

### 5.10.3 Completeness

*Completeness* is the degree to which a database provides information for all data fields that are supposed to contain information (Dupuy, 2011). This dimension is directly related to the predefined minimum requirements. Therefore, for this thesis, completeness is measured by the degree to which investigation reports satisfy the expectations of the predefined minimum requirements. Any report that did not include the minimum requirements was no analysed, therefore completeness, for this study, does not account for quality criterion.

### 5.10.4 Consistency

*Consistency* of data indicates the degree to which data could be considered coherent and compatible with one another within a broad analytic framework and over time (Brackstone, 1999, Dupuy, 2011). Coherence is enhanced by using standard concepts, classifications, methodology and practices, target populations and common procedures across surveys. Within the context of this research, all investigation reports included demonstrate high levels of consistency, as their format and detail of information are identical. However, depending on the severity of events the quantity of included information may vary. Therefore, reports were classified and analysed according to their classification. Subsequently, consistency is not considered a quality dimension for this research.

### 5.10.5 Interpretability

*Interpretability* reflects to which extent data are presented in a clear manner and indicates the availability of a common and shared definition for the collected data (Dupuy, 2011). For investigation reports, this typically implies any underlying concepts, classification and collection methodologies (Brackstone, 1999), where similar responses demonstrate well-interpreted collected data. Interpretability was selected as quality dimension, since many reports do not clearly indicate the contributing factors that influence the performance of operators. Subsequently, the author, while reviewing the selected reports, was extracting the implied information. Therefore, a level of bias, subjectivity and uncertainty may arise concerning the extracted information and the final findings could be questioned. To account for this, three final year PhD students in transport safety, familiar with analysis of safety reports, reviewed a random sample of the 479 reports to identify the contributing factors included in them based on the given definitions (Appendix IV). The findings of the PhD students confirm those of the author, and therefore the author concludes that whilst bias or

subjectivity issues cannot be eliminated, these do not compromise the quality of the findings. Secondly, SMEs verified the final findings, as described later in the Section 6.3.2.

### 5.10.6 Relevance

The *relevance* dimension represents the degree to which the reported data are suitable for a specific domain or study (Dupuy, 2011). For this thesis, relevant investigation reports are considered those that refer to railway accidents and incidents due to human errors. In particular, as stated earlier, this thesis accounts only for occurrences caused by train drivers, signallers and controllers while trains operate from origin to destination, including shunting operations. Therefore, all other reports were considered irrelevant for the purposes of this research. Subsequently relevance is not considered as quality criterion.

### 5.10.7 Timeliness

*Timeliness* reflects the delay between a reference point to which an information refers to, and the date when that information becomes accessible (Brackstone, 1999). In some cases, such delay may have an impact on the accuracy and completeness quality of a database, since the memory of interviewees may have weakened. However, time delay for reporting railway occurrences is mandated and necessary (Dupuy, 2011), therefore timeliness is not considered as quality dimension within the context of this research.

### 5.10.8 Credential

Finally, *credential* reflects the overall quality of the provided information (Dupuy, 2011). Credential is strongly associated with the trust factor that can be put into the data. That is, to what extent data are obtained from reliable (well-reputed) sources and whether researchers are confident in using the specific data for their analysis. Currently credential is measured qualitatively through surveys or audit of the organisation that owns the data (Dupuy, 2011). There is no known method for quantifying the credential of data. This thesis thus proposes such a method (described below) based on the concept of safety maturity.

In this research, credential is considered as relevant quality criterion only for data obtained from railway undertakings (RUs) and infrastructure managers (IMs). It is not the aim of this thesis to challenge NTSBs and NIBs trust in their data management and reporting reliability in particular, since the selected reports describe occurrences that by legislation should be

extensively reported and investigated. On the other hand, the reliability of data derived from RUs and IMs may be disputed; although such reports are again referred to occurrences that by law should be investigated, they constitute outcomes usually obtained from internal investigation processes. However, since it is not possible to assess directly the investigation processes of RUs and IMs, the reliability of reports is deductively measured by evaluating the level of trust in these organisations. Such assessment is based on the concept of organisations' safety maturity, as described in the next section.

## 5.11 Background of safety maturity

Safety maturity is the term used to describe organisations ability to achieve excellence in risk control. To assess such ability several approaches have been introduced over the years. One of the most widely recognised approaches, referred to as the Capability Maturity Model (CMM), has been introduced by Humphrey (1989). The CMM encompasses all the major aspects of process maturity, including behavioural and cultural issues. Although originally developed for software effectiveness, the CMM has been extended to cover a range of business processes including project management, asset management and human capital management (Paulk et al., 1995). It covers key practices and process areas or elements as well as behavioural differences and includes such aspects as process and technology change management, which are highly relevant to the safety and reliability of railways. These elements should ideally be included in any safety maturity model adopted by technology-dependent organisations such as railways. Therefore, it can be claimed that the appropriate way to translate the CMM is by a multi-faceted approach which accounts for all the major elements important to safety such as, human behaviour, hazards identification and quantification, explicit management and mitigation of all types of risk, including those involving equipment. However, the original CMM does not account for risks due to equipment inconsistencies; while by nature the model does not provide or support any quantitative features to assess safety maturity.

Further to the CMM, but partially based on it, the U.K. Keil Centre, in a report for the UK Health and Safety Executive (Fleming, 2000), developed a Safety Culture Maturity Model (SCMM) which focuses exclusively on the assessment of organisational behavioural and cultural issues. It expresses the maturity of an organisation in terms of safety as consisting of the following ten elements: communication; management commitment and visibility; productivity versus safety; learning organisation; safety resources; participation; shared perceptions about safety; trust; industrial relations and job satisfaction; and training. It can

be observed that these ten elements have great similarities with the individual organisational R-PSFs, as listed in the R-PSFs taxonomy. However, although the SCMM may seem suitable to the railway industry, its inventor admitted that in its current format it should be used as framework to facilitate discussion of safety culture issues rather than as assessment tool for organisations' safety maturity (Fleming, 2000). To date no further evolution of the model could be found in the literature. Moreover, the literature (Guldenmund, 2000, Sorensen, 2002, Gunderson, 2005, Choudhry et al., 2007) identifies that a consensus has yet to be reached on the different aspects of safety maturity. For instance, the United Kingdom's railway safety regulator, the Office of Rail Regulation (ORR), demonstrated their belief that a new safety maturity model was needed when they developed their own qualitative model, referred to as the Railway Management Maturity Model (RM3) (Ferguson and Anderson, 2010). The RM3 assesses safety maturity by evaluating organisations management arrangements such as the quality of policies, control and communication, planning and implementation, monitoring, audit and review and finally cooperation and competence (Office of Rail Regulation, 2011a). However, although highly elegant, the RM3 is very complex and thus far purely qualitative. On account of those limitations a new approach is suggested to assess safety maturity referred to as the safety maturity model (SMM).

### 5.11.1 The Safety Maturity Model (SMM)

The Safety Maturity Model is a framework developed for the evaluation of the safety maturity of railway organisations. At the time of its development, the model was initially based on information and data received by metro organisations, and subsequently adapted to the aims of this research. Given the lack of railway specific data, this adaptation is reasonable, as the SMM's main focus is on organisations of safety standards, and not specific types of operations. The SMM's main aim is to introduce railway stakeholders to aspects of organisational structures which they may not have addressed adequately in the past (Guldenmund, 2000). Therefore, the SMM aims to address not only the behavioural and/or attitudinal culture of organisations, as the SCMM, but also technical, operational and methodological elements and actual achievements in terms of safety outcomes. Moreover, compared to both the RM3 and SCMM, the SMM supports the quantitative assessment of safety maturity.

The SMM has been developed on the concept of the CMM; hence where possible the level definitions of the CMM have been incorporated. However, this was not always found to fit

logically with the overall safety maturity levels in the new model. The new model consists of five levels of maturity as illustrated in the Figure 5-4. Its detailed features are described in the Appendix V.

The SMM aims to assess, on a numerical scale, organisations maturity on the following criteria: (i) publication of safety reports, (ii) periodicity of safety monitoring, (iii) prioritisation of safety related incidents, (iv) efforts to mitigate risks and avoid incidents, (v) description of safety procedures including the registration of hazards and finally, (vi) monitoring of residual risks. For its development, a questionnaire (Appendix V) was created based on certain factors that have proved to reduce risk, as highlighted by the CMM, SCMM and RM3 models. The questionnaire was distributed to eleven metro organisations, with the aim of testing metros' safety maturity against their actual performance, as measured by the reported number of fatalities and injuries. By definition it is assumed that high safety maturity scores if negatively related to actual achievements in terms of safety, would indicate organisations with high standards of risk control and safety performance, hence highly reliable organisations.




Figure 5-4 The proposed safety maturity model

For the actual performance of railway organisations, data for precursors[7], top events, injuries and fatalities from 18 metro organisations were analysed for the period 2002-2009. The

[7] Precursors refer to events or conditions that may cause top events or indicate that higher levels of risk exist (e.g. near misses). Top events indicate serious incidents that may result in injury or death (Hirsch, 2006).

questions constitute a representative sample testing the characteristics contained in the SMM. Functional experts involved in risk management, namely those in the company who know most about the reporting procedures, hazard evaluation and development of safety policies completed the questionnaires. At this stage of the model's development, it is assumed that the differences in answers to the questions relate primarily to real differences in procedures and only in very few cases to variances arising from national cultures or from the expert's imperfect knowledge of their own organisation or attempt to give the "appropriate" answer. However, it is yet to be investigated and identified whether such an assumption can be reliably tested.

Each of the questions has the same "weight", i.e. the same importance. For every question a score between 0 and 4 is allotted, where 0 is the "worst" and 4 the "best" scenario. The values for assessing the scale are congruent with the levels of the model, as shown in the Appendix V. The responses of each participating metro are then scored to determine the metro's aggregate safety maturity level on a scale between 0 and 40. This score is then monitored against the metros' levels of actual performance. That is, the monitored and recorded number of precursors, top events, injuries and deaths per million passenger journeys. However, due to the fact that not all participants were monitoring the same number of precursors, a second test measured whether there is a correlation between safety maturity and the numbers of precursors adjusted for the number of missing observations.

### 5.11.2 The SMM results for metro organisations

Initial findings show that the safest railway organisations in terms of safety outputs (low injuries and fatalities) achieve the highest maturity scores, while organisations with a significant number of injuries are assessed with lower maturity scores. Maturity scores equal or higher than 26 indicate organisations that tend to manage risk efficiently. Moreover, scores higher than 33 indicate organisations that tend to optimise both their safety standards and risk management. On the other hand, safety maturity scores lower than 26 represent organisations that should improve their safety standards.

To define any relationships between safety maturity and organisations safety performance, correlation tests were performed. To determine the appropriate test, first the distribution of the four variables that express organisations actual performance - precursors, top events, injuries and fatalities - was investigated. Data showed that none of the variables satisfy the criteria of normality. Therefore, non-parametric correlation tests should be conducted. For this, the Kendall's tau and Spearman's rho statistics, which both use the ranking of data to

calculate correlation coefficients, could be applied. However, although the two coefficients are equivalent, there is much reason to indicate that Kendall's statistic is a better estimate of the correlation in relative small samples (Field, 2009, p.186); hence, it was chosen in this case[8].

The value of Kendall's tau, $\tau$, coefficient lies between -1 and 1. Its sign represents the direction of the relationship between the two variables (either positive or negative) whilst its absolute value reflects the strength of the relationship. The higher the value of $\tau$ the stronger the agreement between the two variables, while 0 value indicates that the two variables are completely independent (Field, 2009). Kendall tau is expressed by:

$$\tau = \frac{c - d}{c+d} = \frac{S}{\binom{n}{2}} = \frac{2S}{n(n-1)} \quad (5\text{-}1)$$

where, *S* represents the number *c* of concordant pairs minus the number *d* of discordant pairs within the *n* sample of observations. For a pair of bivariate observations (*Xi*,*Yi*) and (*Xj*,*Yj*), if (*Xj* -*Xi*) and (Yj - Yi) have the same sign (i.e. both positive), then that pair is concordant, while if they have opposite signs the pair is considered discordant (Nelsen, n.d.).

The performed statistical test showed only one significant correlation (Table 5-5, Figure 5-5). The significant relationship is found between the maturity score and the number of recorded injuries. In the surveyed sample, the maturity score was appreciably negatively correlated ($\tau$ = -.357, p<.01) with the total number of injuries. This indicates that the higher the maturity score of metros, the lower the number of the recorded injuries.

[8] Results in Table 5-5, however, show that similar findings are derived even if Spearman's rho correlation coefficient was used.

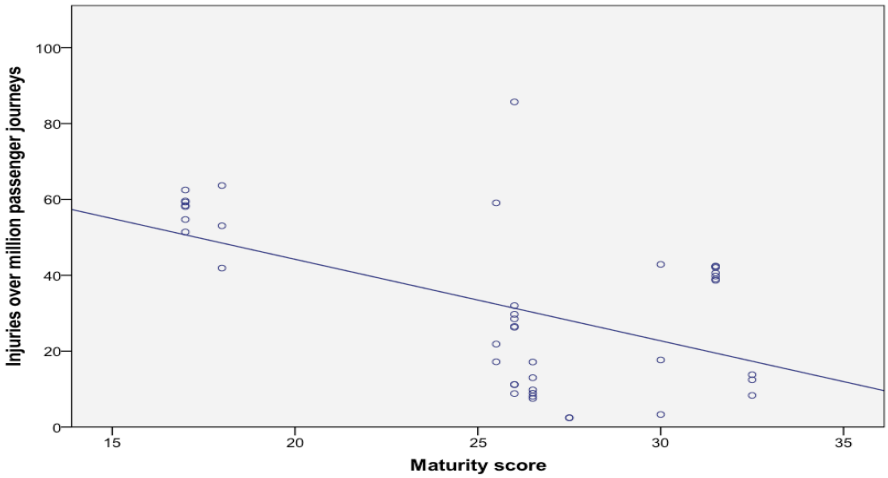



Figure 5-5 Maturity score versus reported injuries

Although correlation is not causality, it may be postulated that the characteristics that lead to a higher maturity index would be likely also to lead to a safer metro performance. It is noteworthy that data for the number of injuries is more likely to be accurate than for precursors or top events. This is explained because monitoring injuries is a legal requirement in most countries, as noted in Chapter 3.

On the other hand, data show that no correlation is identified between the maturity score and precursors, top events or fatalities. For the precursors and top events, this is explained due to the fact that not all metros monitor the same number of precursors and top events. Regarding fatalities, the very small number of observations across all metros could not lead to any useful findings. However, observations of metro behaviour over the years indicate that the more safety-conscious metros tend to record more accident precursors and top events and do so more rigorously, compared to less mature organisations.

To adjust for metros that did not provide a full set of precursors data, the number of precursors that metros could observe was calculated based on the Equation 5-2:

$$P_A = \frac{P}{O/O_{max}} \qquad (5\text{-}2)$$

where, $P_A$ is the adjusted number of precursors, P is the number of precursors observed for a specific year, O is the number of observations (including months with zero incidents) and Omax is the number of maximum observations (listed precursors × 12) for that year.

Table 5-5 Correlations between maturity score and recorded variables

| | Maturity Score | | | |
|---|---|---|---|---|
| **Variables** | **Kendall's tau correlation coefficient** | **Statistical significance at 99% confidence level** | **Spearman's rho correlation coefficient** | **Statistical significance at 99% confidence level** |
| Precursors | -.156 | .146 | -.210 | .157 |
| Adjusted Precursors | -.130 | .234 | -.173 | .256 |
| Top Events | .092 | .425 | .076 | .638 |
| Injuries | -.357 | **.001** | -.492 | **.001** |
| Fatalities | .002 | .987 | .002 | .992 |

where, precursors sample equals to 45, top events sample equals to 41, injuries sample equals to 43, fatalities sample equals to 30,

However, even after the adjustment, results indicate no significant correlation between maturity score and adjusted precursors. Subsequently, for both the monitoring of precursors and top events reliability considerations arose. Despite many efforts made to collect data with a similar degree of rigour across all metros, results verify that the safety-conscious metros are more objective than others in monitoring and registering them. Subsequently, having reviewed the relationships between safety maturity and organisations actual safety performance, it can be claimed that if an organisation has high safety maturity score then any data, which have been by law reported, derived by this organisation can be considered trustable and reliable for further analysis.

## 5.12 Summary

This chapter has presented the framework for the development of the R-PSFs taxonomy. To do so, a literature review of five existing railway PSF taxonomies was first conducted. The findings of the review were then augmented with the results of eleven additional PSF taxonomies from well-known and widely implemented techniques. These identified factors were tested for overlap and then compared with the hierarchical task analysis of railway operators outlined in Chapter 2. This resulted in a list of 43 factors divided into seven main categories, based upon their common characteristics. For each category and factor, specific definitions have been provided and examples are given. This R-PSF taxonomy will be validated using the analysis of 479 accident and incident reports gathered from several railway stakeholders worldwide. Therefore this chapter has addressed a major limitation of existing railway PSF taxonomies.

The latter part of this chapter also introduced a safety maturity model (SMM), which provides a novel approach to assessing (i) the safety levels of railway organisations and (ii) the quality of their reporting systems. This SMM was tested using safety data from a number of metro railways around the world. It is also implemented for the purposes of this thesis to assess the safety maturity of the organisations that provided data to this research, and subsequently the quality of such data.

# Chapter 6 THE RAILWAY PERFORMANCE SHAPING FACTORS TAXONOMY – CONSOLIDATION, RESULTS AND ANALYSIS

The previous chapter presented the process for the development and validation of the Railway-Performance Shaping Factors (R-PSFs) taxonomy. This chapter consolidates this taxonomy by analysing 479 accident and incident reports. The credibility of the organisations that provided this accident and incident data is first examined using the SMM, as presented in the previous chapter. This is followed by the demonstration of data extraction from the investigation reports using the R-PSF taxonomy. The review of the initial dataset is then presented, while the validation of the results with subject matter experts is also detailed. The rest of the chapter presents the analysis of the accident and incident data and interprets the findings.

## 6.1 Credential assessment of railway organisations

As described in Section 5.10.8, credential in this thesis is considered as relevant quality criterion only for data obtained from railway undertakings and infrastructure managers. In this research only two data sets were not derived from NTSBs or NIBs. One of them was derived from an integrated company and the other from an infrastructure manager. To assess the reliability of the obtained data, the SMM framework is used. In the previous section it was stated that if data, which have been by legislation recorded, were derived by highly mature organisations, then this data can be considered reliable. The reports obtained from both organisations meet the requirements of mandatory reporting. Thus, the estimation of their maturity score would, or not, provide confidence with respect to the reliability of data.

As it is shown in the Figure 6-1, both organisations maturity score is higher than 30. This suggests that the level of trust for those railway companies should be considered high. Hence, the derived data can be claimed to be reliable and expedient for further analysis.

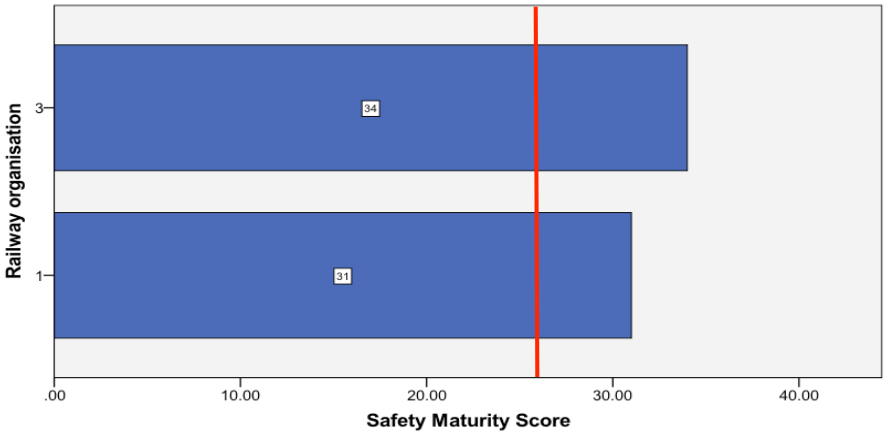



Figure 6-1 Maturity scores of the integrated company and infrastructure manager

In addition to the SMM findings, the literature (Valonen, 2000) also confirms the reliability of the collected reports. In his study, Valonen (2000) shows that the reporting system in the countries where the selected reports were completed, despite any limitations, is effective; hence it can be further suggested that those reports contain reliable information, sufficient to extract accurate outcomes.

Having determined the quality of selected reports, before presenting any data analysis, it is shown in the next section how the minimum requirements of a typical report were extracted.

## 6.2 Extracting information from investigation reports

The report presented in this section describes an occurrence at Hanger Lane junction investigated by the Rail Accident Investigation Board (RAIB). Part of the report is provided in the Appendix VI, while the full report is accessible online (Rail Accident Investigation Board, 2009). The report comprises four main parts: (i) executive summary, (ii) factual information of the event, (iii) detailed analysis and (iv) conclusions and recommendations.

Firstly, the executive summary describes briefly the sequence of events that led to the occurrence, while it also indicates the immediate cause, the causal and contributory factors and any underlying causes. For each of the reports, the author firstly reviewed the executive summary, to confirm whether the report is relevant to the research and if this was the case, to obtain substantial information regarding the event, e.g. location, time of event, involved personnel.

For this particular case, a potential collision between two trains is described, which was caused by the wrong instructions given by the responsible signaller. The event took place within the area of responsibility of Earls Court Control Room, which supervises the services of the west end of Piccadilly and District lines operations. After a District Line train passed a signal at danger (SPAD), the responsible signaller while trying to resolve the incident as soon as possible to avoid any severe delays on the network, authorised the train to proceed towards a junction; however, he overlooked the presence of another train resulting to a near miss and potential collision between the two trains.

The immediate cause of the event was the signaller's wrong authorisation to the train to proceed towards a junction before ensuring that the move is safe. Two were the causal factors for such occurrence. The first, that the District Line train had a SPAD. However this event, if isolated, would not have caused any safety concerns apart from potential delays on the network, therefore was not in detail investigated in the report (§99–102, Appendix VI). The second and most important causal factor was that the signaller did not halt all trains in the vicinity before giving to the District Line train authorisation to proceed.

In addition to the immediate cause and causal factors, the executive summary clearly states a number of factors that contributed to the occurrence. That is, signaller's *workload*, *time pressure* and *fitness for duty*, *communication* arrangements, *design of the system*, *quality of procedures* and *training methods* (highlighted in Appendix VI). Subsequently, seven R-PSFs were identified, covering almost all R-PSFs categories listed in the Table 5-3.

However, although the executive summary provides a constructive illustration of the event, if researchers had only reviewed the summary section, they would have missed at least three additional factors that significantly contributed to the occurrence. As stated, signaller's workload was identified as contributing factor to the event. In the summary, it is also written that this caused "*as he was operating two control desks during the incident*", but no further information is provided. However, in the analysis part of the report it is mentioned that the signaller was operating on two control desks, as it was common process within the control centre in order for two signallers to take their meal breaks at the same time (§33, in Appendix VI). According to the given definitions for the R-PSFs, this is an issue of compromised *safety culture*, since it directly affects the safety of the operations. In addition, it is clearly indicated in the report that signaller was *distracted* since he was in charge of two control desks (§109) and subsequently forgot that a second train was moving towards the junction. Finally, the analysis indicates limitations concerning the context of the course that the signaller attended to refresh his skills. This, again based on the given definitions, is

related to the *supervision* provided to the signallers. Table 6-1 demonstrates the extracted from the report information for the analysis.

Table 6-1 Information extracted from investigation reports (refer to Appendix VI)

| **Occurrence type** | **Type of railway** | **Year** | **Date** |
|---|---|---|---|
| ▪ Near miss (title) | ▪ Metros (§6,7) | ▪ 2009 (§6) | ▪ 27 March (§6) |
| **Location** | **Area** | **Time** | **Responsibility** |
| ▪ Hanger Lane junction (§6, 20,21) | ▪ London (§6) | ▪ 5:26 pm, (§7)<br>▪ (sunset 6:26pm) | ▪ Signaller (§8),<br>▪ 10 years of experience, instructor for new signallers (§80) |
| **Description** | **Immediate cause** | **Causal factors** | **PSFs (contributing factors)** |
| ▪ Eastbound District Line train 103 passed signal WM1 at danger at low speed (§6)<br>▪ Signaller authorised train operator to proceed across junction (§7)<br>▪ The same time westbound Piccadilly train crossed the junction creating potential for collision. Signaller overlooked the presence of this train and been unable to contact train operator of 103 train once became aware of the situation (§7) | ▪ Signaller gave approval to train 103 to proceed towards Hanger Lane junction before it was safe to do so (§8)<br>▪ Wrong track authorisation | ▪ Signaller did not bring all trains to a halt (§9)<br>▪ Signaller was taking prescribed medication (§9) | ▪ Time pressure (§11)<br>▪ Workload (§11,33)<br>▪ Communication (§11,103)<br>▪ System design (§12, 116,117)<br>▪ Quality of procedures (§12)<br>▪ Supervision (§81,123)<br>▪ Fit to work (§10,110)<br>▪ Distraction (§109)<br>▪ Safety culture (§33)<br>▪ Training (§12,14,81) |
| **Feature of the system** | | **Consequences** | |
| ▪ Signalling information (§25–27)<br>▪ Train information (§23,24)<br>▪ System design further information (§116,117) | | ▪ No injuries or material damage | |
| **Recommendations** | | | |
| ▪ Guidance should be provided to assist signallers in dealing with similar incidents<br>▪ Use of simulation techniques during training | | ▪ Training of safety critical communication skills<br>▪ Issuing of medical advice to managers | |

Having presented how the necessary information were extracted from the relevant investigation reports, the next sections detail the validation of the R-PSFs taxonomy based on real data and the consultation with SMEs.

## 6.3 Review of initial dataset

As shown in Figure 5-2, initially, 192 accident and incident railway reports were reviewed, gathered from the European Railway Agency Database of Interoperability and Safety (ERADIS), the Swiss Federal Railways (SBB) and the Australian Transport Safety Bureau (ATSB). The aim of the analysis is to enhance the results from the literature, to identify whether any factors should be added, replaced or removed, and to indicate the dependencies between the factors including in the R-PSFs taxonomy. Therefore the analysis focused primarily on extracting the relevant to the occurrences R-PSFs. To further verify the findings an expert's elicitation assessment was then conducted in collaboration with the Swiss Federal Railways (SBB).

### 6.3.1 Identification of R-PSFs from reports

The first set of data includes 105 reports (23 serious accidents, 46 accidents, 36 incidents) from the ERADIS database, 69 reports (6 serious accidents, 45 accidents, 18 incidents) from the SBB, and 18 reports (2 serious accidents, 6 accidents, 10 incidents) from the ATSB for the period 1997-2011. The sources of the reports represent many of the situations experienced in railway operations, for example, different types of networks, varying geography (Australia and Switzerland), different regulations, and different staffing (two drivers on the Australian trains instead of one on most European trains). Therefore, it can be argued that the sampled data are representative of the population of interest, to confirm the completeness of the identified from the literature R-PSFs taxonomy. Based on the given definitions (Appendix IV), 597 R-PSFs were finally extracted, as follows: 172 R-PSFs refer to incidents, 272 to accidents and 153 to serious accidents. That is approximately 1.5 R-PSFs per incident report, 3 per accident and 5 per serious accident report. The identified R-PSFs per type of event are illustrated in the Figure 6-2.

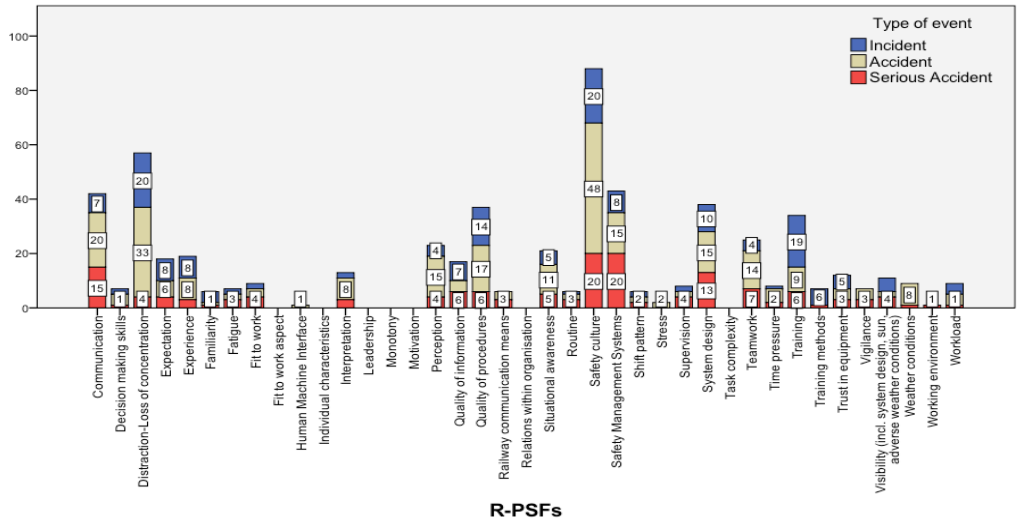



Figure 6-2 The identified R- PSFs per type of event from the initial set of reports

At first glance it is observed that no additional R-PSFs compared to those identified from the literature were either extracted or implied in the reports. Therefore, confidence is gained that the suggested list of 43 R-PSFs is representative for railway operations. However, it can be also observed that the extracted factors do not include all the factors that are contained in the R-PSFs taxonomy. In particular, 11 out of 43 factors were not found within the reports. These are: (i) *individual characteristics* and *motivation* (personal factors); (ii) *monotony*, *task complexity* and *task instructions* (task factors); (iii) *relations within the team* (team factors); and (iv) *communication within organisation*, *fit to work aspect*, *incentives for employees*, *leadership* and *relation within organisation* (organisational factors). The non-identified factors could be further classified into two main categories. The first includes the factors: individual characteristics, leadership, task complexity, task instructions and monotony. This category can be assumed to represent those factors that were not identified simply because no occurrence was related to them. For the remainder non-observed factors, however, such assumption may not be accurate. In contrast, it is likely to represent cases that are difficult to be identified within the investigation process as these factors are, by nature, implied or hidden; therefore it is not straightforward for them to be identified either by inexperienced investigators or by an investigation process that does not consider for such factors. On account of this, the author instigated further whether the unidentified factors were reported in the additional reports and secondly pointed out the significance of those factors to railway organisations.

In addition, findings show that not all factors have been equally identified within the reports. In other words, it is rather likely that not all factors have the same significance with respect to their influence on human performance and subsequently to safety of railway operations.

Therefore, the Pareto principle (Backbaus, 1980), which shows that the majority of the results come from a minority of inputs on an 80-20 basis, was applied to determine those R-PSFs that most frequently occur considering the severity of their consequences. For this, factors with common characteristics were first combined, in order to capture as many cases as possible. Subsequently, 17 combined R-PSFs (Table 6-2) were identified as responsible for more than 85% of the occurrences. In addition to them, s*tress* was also incorporated as it was identified in most of the reviewed taxonomies.

Table 6-2 The most frequent identified R-PSFs from the first set of reports

| R-PSF | Category | R-PSF | Category |
|---|---|---|---|
| Distraction – Loss of Concentration | DP | Experience – Familiarity | P |
| Expectation – Routine | DP, T | Time pressure (time to respond) | T |
| Communication | Te | Fit to work – Health | P, O |
| Safety culture – SMS | O | Workload | T |
| Training – Competence | P, O | Supervision | O |
| Perception | DP | Teamwork | Te |
| Quality of procedures | O | Weather conditions – Visibility | E |
| System design – HMI | S | Stress | DP |
| Fatigue (sleep lack, shift pattern) | DP, O | Situational awareness | DP |

DP: Dynamic personal, E: Environmental, P: Personal, T: Task, Te: Team, O: Organisational, S: System

This shorter list of factors was then compared against the complete list of R-PSFS by SMEs to indicate if any modifications are required, regarding the factors considered of higher importance for human performance.

### 6.3.2 Validation and assessment with SMEs

To confirm that no significant R-PSF is missing from the suggested list of factors, a study was conducted in collaboration with a widely recognised European railway organisation, the Swiss Federal Railways (SBB).

In the study, 18 SBB employees participated including 2 risk analysts, 3 passenger train drivers, 6 freight train drivers and 6 signallers. All employees have at least seven years of operational experience in their current position (μ=9.35, sd=7.09) and eight within the SBB organisation (μ=18.76, sd=9.18), while their average age is 42 years old (sd=9.45). It should be also noted that train drivers are dedicated either to passenger or freight trains. The SMEs were asked to mark if any of the 18 identified factors should either be replaced by any of the

remainder R-PSFs or removed from the list. None of the SMEs replaced any of the R-PSFs in the list; on the contrary they underpinned that these are the factors that predominantly affect their performance. However, recommendations were given for removing factors, as well as for changing some of the fusions of factors on the list. For instance, it was suggested to incorporate the factor *experience* with *training* instead of *familiarity* as, based on the definitions, such a combination was deemed more compatible. Similar, it was proposed that the factor *familiarity* should be combined with the factors *expectation* and *routine*. Furthermore, it was recommended to unite the factors *workload* and *stress*, as well as the factors *fatigue* and *fit to work.* It was suggested, finally, to remove the factor *weather conditions*; nevertheless, it was highlighted the importance of asking operators to express their general opinion of the influence of adverse and intermediate weather conditions on their performance.

Having confirmed that the majority of the most significant R-PSFs were captured, the study was also used to indicate whether the frequency of the extracted from the reports R-PSFs matches to how SMEs rank those factors with respect to how they influence their performance. As such influence may change for different tasks or operational scenarios; the assessment was associated with particular tasks and operational scenarios. Hence, a preliminary analysis based on the gathered reports was conducted. The analysis focused on identifying the factors that affect the types of operators, to test the hypothesis of independency between the factors and operators and finally to determine the most frequent and severe events with respect to responsible personnel. The Figure 6-3 shows the identified factors per type of employee, as extracted from the reports.

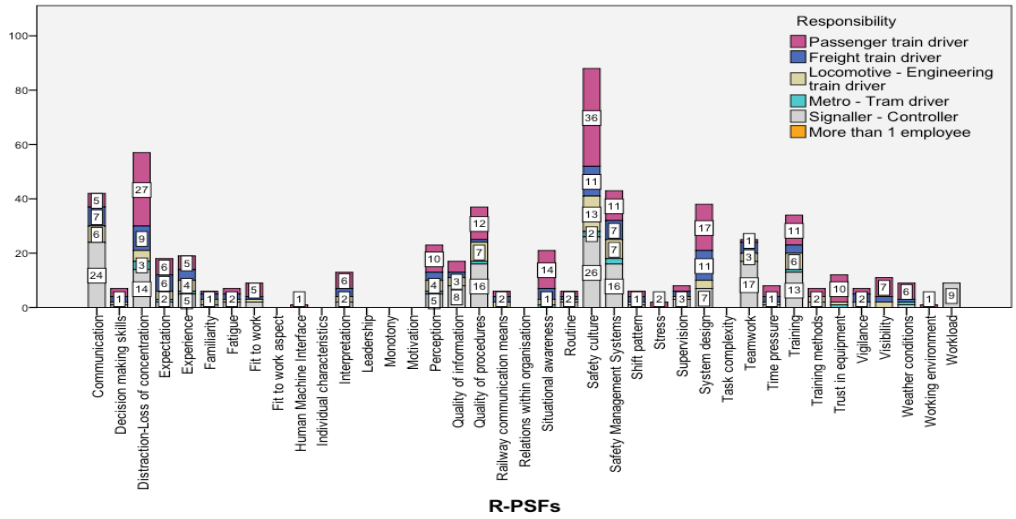



Figure 6-3 Most frequently identified R- PSFs per type of operator

If the hypothesis of independency is true, the difference $\{n_{ij} - \mu_{ij}\}$ should be small. In contrast, the larger the difference $\{n_{ij} - \mu_{ij}\}$, the stronger the evidence against the null hypothesis, in which case the two variables are deemed "*associated*" with each other (Agresti, 2007). The Pearson's chi-squared statistic is calculated as:

$$\chi^2 = \sum \frac{(n_{ij} - \mu_{ij})^2}{\mu_{ij}} \qquad (6\text{-}1)$$

where, $n_{ij}$ and $\mu_{ij}$ are the observed and expected frequencies, and *i* and *j* represent rows and columns in the contingency tables respectively (Field, 2009). The test requires that less than 20% of the cells in such tables should have expected frequencies below 5, while no single expected frequency should be lower than unity. However, if expected frequencies are too low, usually the case when samples are small, Fisher's exact test is used (Field, 2009). This is a method for computing the exact probability based on the *hypergeometric* distribution of cell counts (Agresti, 2007).

To run the tests, as well as any statistical tests in this research, the statistical package IBM SPSS Statistics 19.0.0 has been used. For both variables, the type of operators and the identified R-PSFs, clustering of categories was needed to avoid cell frequencies less than five and/or null. For instance, the signallers-controllers category was formed from the union of signallers and controllers. Similarly, the category locomotive running solo drivers and engineering train drivers formed the category locomotive-engineering drivers. Metro and tram drivers were neglected due to their limited representation. Regarding the R-PSFs the eighteen most common factors were used to satisfy both requirements of Pearson's chi-

squared test. Within those factors, *workload* and *stress* were united, as well as *fatigue, fit to work* and *shift pattern.* In contrast, weather conditions were not taken into account.

Results indicate that there is a significant association between the type of identified R-PSFs and the types of operators, if these are divided into train drivers and signallers-controllers ($\chi^2(14)^9$=81.671, p<.001), e.g. *communication* related errors are more typical for signallers. Furthermore, association was observed between the different types of drivers (passenger, freight and locomotive) and R-PSFs ($\chi^2(28)$=28.503, p<.05), e.g. *fatigue* affects primarily the freight train drivers. However, no association was identified between the different types of passenger (long-distance and regional) train drivers and R-PSFs ($\chi^2(14)$=48.624, p>.05). Therefore, the category passenger train drives was also formed, including those two types of drivers. Finally, from the collected reports no variation was identified concerning the frequency of the severity of events.

Based on the outcomes, it was decided in cooperation with the SBB risk analysts to define four scenarios deemed to be of great importance for the railway industry. Three of them are related to train drivers, while the remaining to signallers, as follows:

1. Scenario 1: Long-distance passenger train derails due to excessive speed.
2. Scenario 2: Regional passenger train passes a signal at danger although the driver was aware of it.
3. Scenario 3: Freight train passes signal at danger because of the wrong interpretation by the driver.
4. Scenario 4: During a shunting operation a shunter engine collides with a commercial train as the result of a signaller error.

The scenarios contain information, such as the type of train and occurrence; responsible and involved personnel; system information, for instance, technical failures; weather conditions; operator's level of experience; location and time of the event. All scenarios were then refined and checked for their simplicity, clarity and accuracy with the SBB risk analysts and the Organisation, Work and Technology Group of ETH Zurich and subsequently a questionnaire was created, which can be found in Appendix VII. For the R-PSFs assessment no alterations with respect to the initial list of 18 factors was made. Subsequently, the questionnaire was distributed to the SMEs, who assessed the factors on a scale of 1 to 18, ranging from most

[9] The number in parenthesis indicates the degrees of freedom for the statistic, which is calculated as (i-1)·(j-1), where *i* is the number of rows and *j* is the number of columns in the contingency table.

(indicated with 1) to least (indicated with 18) important for each factor. The assessment was conducted on an individual basis, apart from one case in which two participants completed one ranking assessment together. Participants ranked the R-PSFs by employee and primary cause, as included in the four scenarios. The author's physical presence when 10 out of 18 participants assessed the ranking, together with videocall contact with the rest, assured the validity of the responses in relation to any misunderstandings or vagueness experienced by the participants.

Due to the ordinal nature of responses, the non-parametric Kruskal-Wallis test (McCrum-Gardner, 2008) was applied to test the null hypothesis whereby the distributions of R-PSFs ranking across *k* experimental conditions (i.e. types of operators) were the same for each of the scenarios. For this test, observations for all groups are jointly sorted, ranked and then scored from lowest to highest, while the group to which each score belongs is ignored. The sum of ranks $R_i$ and the number of observations $n_i$ for each group is then derived. For a sample of size *N*, the test statistic, *H*, is calculated as:

$$H = \frac{12}{N(N-1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1) \qquad (6\text{-}2)$$

where *H* satisfies the chi-square distribution with (*k*-1) degrees of freedom (Field, 2009). Results for all scenarios (Table 6-3), show that the types of operators do not affect the ranking of R-PSFs, since $p>.05$ (Field, 2009).

Table 6-3 Test for differences between different types of operators for R-PSFs rankings

| R-PSFs | Scenario 1 | | Scenario 2 | | Scenario 3 | | Scenario 4 | |
|---|---|---|---|---|---|---|---|---|
| | *H*(3) | *p* | *H*(3) | *p* | *H*(3) | *p* | *H*(3) | *p* |
| Communication | 3.944 | .268 | 3.297 | .348 | 0.199 | .978 | 0.927 | .819 |
| Distraction - Loss of | 2.189 | .534 | 1.720 | .632 | 0.165 | .983 | 1.193 | .755 |
| Expectation - Routine | 1.970 | .579 | 4.051 | .256 | 0.939 | .816 | 2.748 | .432 |
| Experience - Familiarity | 1.192 | .755 | 3.534 | .316 | 1.380 | .710 | 1.390 | .708 |
| Fatigue (sleep lack, shift | 0.897 | .826 | 5.212 | .157 | 4.623 | .202 | 3.301 | .347 |
| Fit to work - Health | 3.006 | .391 | 3.154 | .369 | 1.441 | .696 | 0.984 | .805 |
| Perception | 1.640 | .650 | 3.633 | .304 | 7.854 | .049 | 6.447 | .092 |
| Quality of procedures | 5.905 | .116 | 2.557 | .465 | 4.213 | .239 | 2.517 | .472 |
| Safety culture - SMS | 4.553 | .208 | 2.978 | .395 | 5.943 | .114 | 3.872 | .276 |
| Situational awareness | 4.659 | .199 | 2.794 | .424 | 3.967 | .265 | 8.954 | .065 |
| Stress | 0.479 | .923 | 0.660 | .883 | 0.924 | .820 | 5.921 | .116 |
| Supervision | 1.944 | .584 | 0.415 | .937 | 2.040 | .564 | 2.397 | .494 |
| System design - HMI | 1.004 | .800 | 6.663 | .083 | 4.989 | .173 | 5.423 | .143 |
| Teamwork | 4.256 | .246 | 4.471 | .215 | 3.873 | .275 | 5.280 | .152 |
| Time pressure (time to | 0.812 | .847 | 4.829 | .185 | 0.420 | .936 | 6.569 | .087 |
| Training - Competence | 4.400 | .221 | 4.257 | .235 | 2.060 | .560 | 4.423 | .219 |
| Weather conditions - | 3.824 | .281 | 4.149 | .246 | 4.012 | .260 | 1.506 | .681 |
| Workload | 5.132 | .162 | 3.275 | .351 | 3.336 | .343 | 3.656 | .301 |

Grouping variable: Types of operators

As R-PSFs rankings do not differ per group data were aggregated. A four-scale range was then created to indicate the importance of R-PSFs on human performance, as follows: (i) 1-4, very important, (ii) 5-9, important, (iii) 10-14, relatively important and (iv) 15-18, less important. Results from assessment confirm in most of the cases the findings from reports, as shown in Table 6-4. Five of the presented R-PSFs were identified from reports as too frequent (Figure 6-2), while *workload* and *stress* illustrate a noticeable contradiction between reports and assessment results. Table 6-4 contains the following information: the four scenarios, findings for seven R-PSFs, and their rankings as identified from the reports, the percentage of SBB employees' common responses (first figure in column), and the corresponding range of these responses (in parentheses).

Table 6-4 R-PSFs assessment per scenario

| R-PSFs assessment | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2*. Distraction | | 1. Safety culture – SMS | | 4. Communication | | 9. Expectation – Routine | | 18. Workload | | 7. Training | | 30. Stress | |
| **Scenario 1** | **86.4% (1)** | **1-4 (2)** | 24.63% | 1-4 | 6.25%<br>**75.00%** | 1-4<br>**15-18** | 52.94% | 1-5 | 70.59% | >6 | 52.94% | 10-14 | 35.3% | 2-5 |
| **Scenario 2** | **76.47%** | **1-4** | 35.29% | 1-4 | 11.76%<br>**52.94%** | 1-4<br>**15-18** | 52.94% | 1-5 | 70.51% | >6 | 58.82% | 10-14 | 29.4% | 2-5 |
| **Scenario 3** | **76.47%** | **1-4** | 11.76% | 1-4 | 11.76%<br>**70.59%** | 1-4<br>**15-18** | 52.94% | 1-5 | 76.47% | >6 | 64.71% | 10-14 | 29.4% | 2-5 |
| **Scenario 4** | 52.94% | 1-4 | **52.94%** | **1-4** | 52.94%<br>**17.65%** | 1-4<br>**15-18** | **47.06%** | **1-5** | **70.59%** | **3-6** | 52.94% | 10-14 | **70.6%** | **2-5** |

* indicates the ranking of R-PSFs based on the reports findings

Considering the example of *distraction - loss of concentration*, which was identified as the second most frequent R-PSF from the reports. The assessment of SBB employees shows that for the train driver scenarios, more than 75% of employees identify this particular R-PSF as the most significant (range 1–4) and at least 50% believe the same for the fourth scenario. The rest of the R-PSFs are illustrated accordingly, in particular, as follows:

- Regardless the network or the operation, *distraction–loss of concentration* has been determined to be very significant, which strongly justifies the findings from reports.
- *Safety culture - SMS* is found a significant factor for signallers but not equally important for drivers, which strongly contradicts the results from the reports, as shown in Figure 6-3. Although safety culture should be equally important for all types of employees and operations, it is possible that the observed difference is due to the given definition, which includes the cases of disregarded procedures. For the selected scenarios, SMEs believe that it is extremely important for signallers to comply with given procedures in order to keep the operation safe and secure; whilst they did not seem to consider that train drivers may disregard relevant procedures.
- *Communication between employees* was assessed as important contributor R-PSF for signallers, which confirms the analysis findings.
- *Workload* is an essential PSF for signallers but not for train drivers, which again justifies the results in Figure 6-3.
- Although *training* (lack of or inadequate) was identified as a fairly substantial R-PSF from the analysis, this was not confirmed from the assessors. It is possible that the participants assumed or considered that rail operators have sufficient training to execute their tasks, whereas it cannot be neglected the fact that within the investigation reports it is very often mentioned the factor training, as considered a tangible factor to assess performance.
- Finally, *stress* was identified from assessors as a serious R-PSF contributor for signallers, but not for drivers. However, within the reports stress was barely identified. Subsequently, it is likely that, again, the reporting system does not provide direct information about this R-PSF, so researchers have to assume its existence.

### 6.3.3 Dependencies amongst the R-PSFs

Further to the validation of factors included in the R-PSFs taxonomy, the dependencies amongst factors were also investigated. At first glance, the identified factors reveal possible

interactions as indicated by the R-PSFs structure in Figure 5-3. A poorly designed display (i.e. HMI), for instance, as well as inadequate knowledge of system modes (i.e. inadequate training) in addition to a preoccupied operator (i.e. distraction) may lead to delayed failure detection, e.g. signal failure or braking failure and subsequently increase the likelihood of an occurrence due to an operational error. Similarly, stress as a dynamic personal factor cannot be independent of workload. If an operator deals with increased levels of traffic or delays, it is reasonable to assume that stress levels will be higher (Subotic, 2007). The analysis of interactions makes it possible to gain a more accurate picture of the context and thus a better understanding of the operational errors due to degraded human performance.

The problem of interactions between PSFs has been addressed in the literature, but not extensively. Sträter (2000) tackles the issue by looking at the common appearance of different factors. The analysis is based on capturing the observed interactions between the reported PSFs. The availability of a detailed database is, however, a requirement for this approach. Hollnagel (1998), in contrast, establishes the interactions between the contextual conditions in CREAM by considering each one with respect to how it generally influences the others (there is no information however whether expert judgment or operational expertise have been used) (Subotic, 2007). Finally, Gertman et al. (2004), address the problem by adapting the existing THERP dependency assignment using a zero to compete dependency rating system.

In this research, the dependencies amongst R-PSFs were investigated and determined in three stages. In the first, the interactions were identified based on known relationships from operational experience, as indicated in railway operational manuals (Network Rail, 2008a, Network Rail, 2008b, Network Rail, 2008c) and task analyses (presented in Chapter 2). The determined interactions are marked with the symbol "✔" in Table 6-5. They represent the irreversible influence between two R-PSFs or how R-PSFs in the left hand column affect the R-PSFs in the row. The reason for irreversible influence lies in the characteristics of the railway system where one factor may influence the other one without any reverse effect. For example, complex traffic, simultaneous train manoeuvres and authorisations to trains to proceed could affect signallers' capabilities in terms of increased stress, and workload; whilst the opposite influence (impact of stress on traffic complexity) is simply not logical.

The second and third stages were performed to validate the identified dependencies. In the second stage, the dependencies were specified as found from the reports analysis and marked with "(a)", whilst in the third stage the SBB SMEs were asked to indicate on a blank matrix sheet (part two of Appendix VII) how they perceive the interactions between factors.

Their input was then incorporated in Table 6-5 and marked with "(b)". However, as the complete list of 43 R-PSFs creates a complex matrix of dependencies, which requires extensive time to be completed, it was decided to ask SMEs to indicate only the dependencies amongst the most frequent identified R-PSFs. These dependencies are illustrated in Table 6-5, where 95.23% (220/231) of the interactions have been validated by SMEs and 62.77% (145/231) from the reports.

Table 6-5 Dependencies amongst R-PSFs

| Most frequent R-PSFs<br>Direct influence of Y upon X | Communication between employees - Information | Teamwork - Supervision | Distraction - Loss of concentration | Expectation - Routine | Experience - Familiarity | Situational awareness | Fatigue | Fit to work - Health | System design - Human Machine Interface | Perception | Safety culture (disregard procedures) | Quality of procedures | Visibility | Stress | Time pressure - Time to respond | Training | Weather conditions | Workload |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Communication between employees - Information | | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) |
| Teamwork - Supervision | ✔ (a), (b) | | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) |
| Distraction - Loss of concentration | ✔ (a), (b) | ✔ (b) | | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (a), (b) | | ✔ (b) | ✔ (a), (b) | | | | ✔ (a), (b) | ✔ (a), (b) | | | ✔ (a), (b) |
| Expectation - Routine | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) |
| Experience - Familiarity | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (b) | | | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | | ✔ (a) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) |
| Situational awareness | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | | | | ✔ (b) | ✔ (a), (b) | | | ✔ (b) | ✔ (a), (b) | ✔ (b) | | | ✔ (b) |
| Fatigue | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | | ✔ (a), (b) | ✔ (b) | ✔ (b) | | | ✔ (a), (b) |
| Fit to work - Health | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (b) | | ✔ (b) | ✔ (a), (b) | | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | | | ✔ (b) |
| System design - Human Machine Interface | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | | | ✔ (b) | ✔ (b) | | ✔ (a) | ✔ (b) | ✔ (b) | ✔ (a), (b) | | ✔ (a), (b) |
| Perception | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (b) | | ✔ (b) | | ✔ (b) | | | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (b) |
| Safety culture (disregard procedures) | ✔ (a), (b) | ✔ (a) | ✔ (a) | ✔ (a) | ✔ (b) | ✔ (b) | ✔ (b) | ✔ (b) | ✔ (a) | ✔ (b) | | ✔ (b) | ✔ (a) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a) | | ✔ (b) |
| Quality of procedures | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | | ✔ (b) | ✔ (b) | ✔ (b) | | | ✔ (b) | ✔ (a), (b) | ✔ (a) | | ✔ (b) |
| Visibility | | | ✔ (b) | ✔ (a), (b) | | ✔ (a), (b) | ✔ (b) | | | ✔ (b) | | | | ✔ (b) | ✔ (b) | | | ✔ (b) |
| Stress | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | | ✔ (b) | ✔ (a), (b) | ✔ (b) | | ✔ (b) | | | | | ✔ (a), (b) | | | ✔ (a), (b) |
| Time pressure - Time to respond | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | | ✔ (a), (b) | ✔ (a), (b) | ✔ (a), (b) | | | ✔ (a), (b) | | ✔ (a) | | ✔ (b) |
| Training | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | | | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | | ✔ (b) | ✔ (a), (b) | | | ✔ (a), (b) |
| Weather conditions | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (b) | ✔ (a), (b) | | | ✔ (a), (b) | | | ✔ (a), (b) | ✔ (b) | ✔ (b) | | | ✔ (b) |
| Workload | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (b) | ✔ (a), (b) | ✔ (a), (b) | ✔ (b) | ✔ (b) | | | ✔ (a), (b) | ✔ (a), (b) | | | |

(a): reports, (b) SBB employees

Based on the dependencies between R-PSFs, the network of dependencies amongst the R-PSFs categories is developed, as illustrated in Figure 6-4. Dependency between two categories is established when at least one R-PSF included in one of the two categories has an influence or is influenced by at least one factor of the second category, e.g. familiarity on distraction. These types of dependencies are considered outer. In addition to them, also a number of inner dependencies are also established. Inner dependencies indicate that within the same category at least one factor has an influence or is influenced by another factor, e.g. fatigue on distraction.

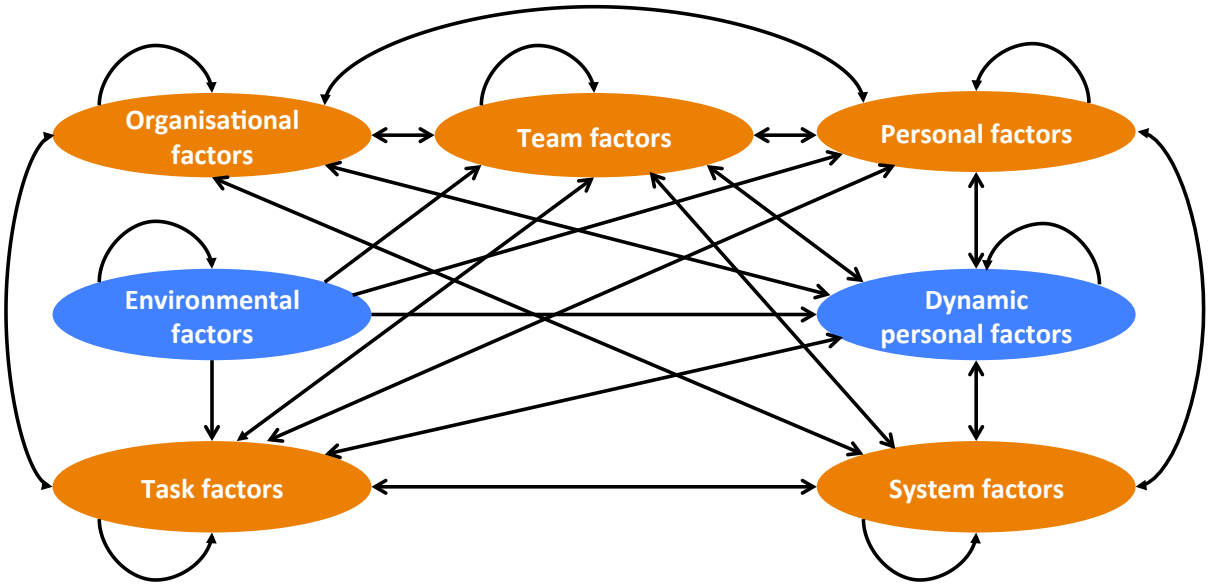



Figure 6-4 The network of interdependences between the categories of R-PSFs

### 6.3.4 Discussion of findings

The review of the first set of reports was conducted primarily to confirm that the list of the R-PSFs identified from the literature is complete and represents the majority of factors that affect the performance of railway operators. The results give confidence with respect to the completeness of the taxonomy; however they also show that factors are not equally identified on railway occurrences. Thus, based on the severity of occurrences, a set of 18 R-PSFs was established which includes those factors that mostly affect the performance of operators and subsequently the network in terms of safety. In turn, SMEs were asked to indicate the importance of those R-PSFs on human performance so as to compare (justify) the findings from reports with the opinion of specialists. Findings indicate that although for many R-PSFs the responses of SMEs were different, the range of such difference was marginal; subsequently in most of the cases the findings from reports are in general agreement with the perception of operators. In addition, the SMEs provided

recommendations with regard to how the list of the 18 factors could be changed so as to be more representative, capture additional occurrences and become more flexible. Finally, the problem of R-PSFs interactions was addressed and the dependencies between them established.

However, although the factors and their dependencies were validated, the analysis did not take into account certain variables that may be associated with the existence of particular R-PSFs. Such variables include, amongst others, the location of events, e.g. open line, the operational conditions, e.g. normal or degraded, the on board and infrastructure safety and protection systems, e.g. TPWS or the time of the event. Additionally, the immediate causes and associated events of the occurrences, e.g. signal passed at a danger, should be considered. Especially, since it is likely that specific causes are to be associated with certain R-PSFs (Rail Safety and Standards Board, 2004b).

The aim of this thesis is to introduce a R-PSFs taxonomy that represents and can be implemented within a broader context of operations. Therefore, additional reports were gathered and incorporated to the first set of reports. Analysis of all reports was performed aiming at: (i) extracting those R-PSFs that most frequently are involved in railway occurrences in a broader operational context, (ii) defining any relationships between the factors identified in (i) and the different types of events, (iii) indicating associations between the factors found in (i) and the variables that constitute the overall context of railway occurrences, (iii) determining any higher-order interactions amongst more than two associated variables, where the two of them are always the *type of event* and *R-PSFs*, e.g. *type of event – R-PSFs – variable 1*, and (iv) generating predictions for the types of events with respect to the contributing R-PSFs, as found in (i). Figure 6-5 shows the framework for the analysis.

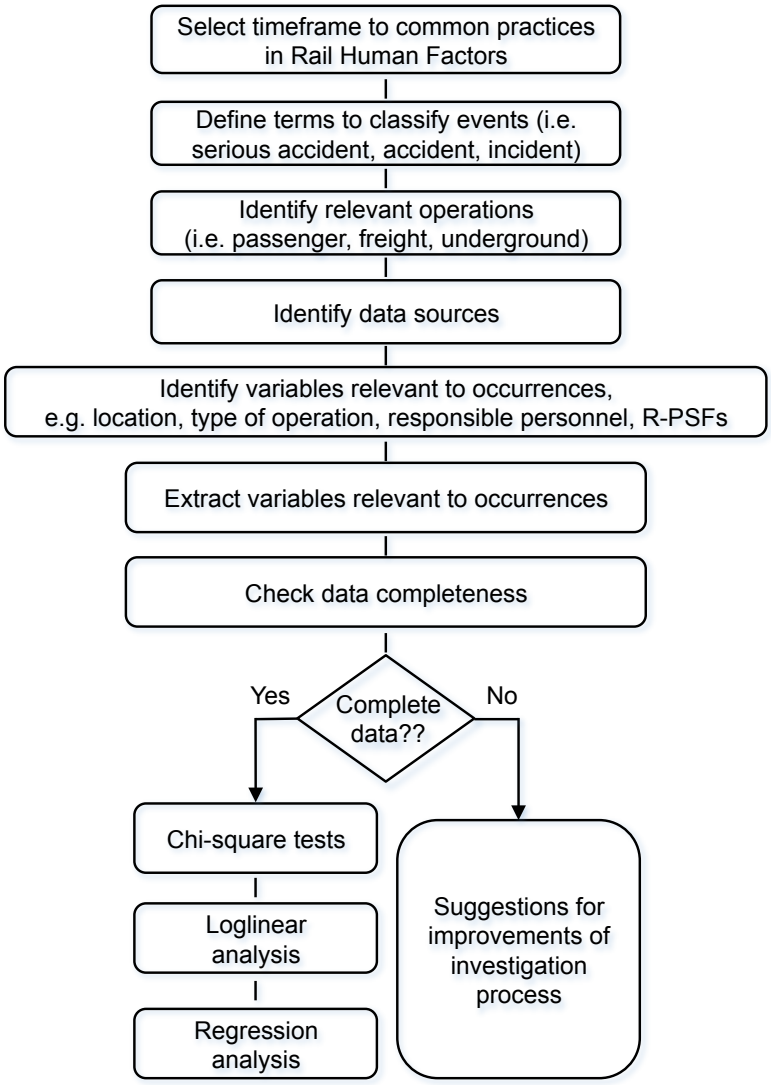



Figure 6-5 The methodology for the analysis of the occurrences

## 6.4 Analysis of the complete dataset

The complete review embraces 479 reports (287 in addition to the initial 192). Similar to the first set of analysis, the sources of the reports represent the majority of the situations experienced in railway operations. These include the different operations, for example, in the US where the majority of operations is freight, varying geography, e.g. the U.K. and Switzerland, and different regulations or staffing. In addition to the databases of Section 6.3, the supplementary reports have been obtained from the: French Land Transport Accident Investigation Bureau (BEA-TT), Canadian Transport Safety Board (CTSB), Dutch Safety Board (DSB), Finnish Safety Investigation Authority (SIA), U.K. Network Rail, Norwegian Accident Investigation Board (AIBN), U.K. Rail Accident Investigation Branch, Irish Railway Accident Investigation Unit (RAIU), Swedish Accident Investigation Authority (SHK) and U.S.

National Transport Safety Board (NTSB). Therefore, the sampled data can be argued to be representative of the population of interest and thus, suitable for a detailed analysis of the variables that affect railway safety in developed countries with well-established railway networks. The purpose of the analysis is to deliver information that can be used within a more generic context, incorporating all possible sources of information. Hence, as all reports satisfy the quality criteria, it is assumed that all have been derived from the same source and are treated accordingly. It does not compare the safety levels of the countries the reports originate from.

The analysis consists of 5 steps. In the first, the necessary information is extracted from reports, as described in Section 6.2. For this a master file, which contains all the data, is created. Based on the extracted information, step two involves a preliminary descriptive analysis of the data. This includes the frequencies of the different types of accidents, R-PSFs and all relevant variables that are associated with railway accidents. However, due to the unequal representation of several variables, clustering of the variables (except the R-PSFs) is performed in step two to enable further analysis.

The clustering of the R-PSFs is undertaken in step 3, based on the findings from the first set of analysis and the recommendations of the SMEs. Subsequently, the most important R-PSFs that affect human performance in railway operations are established, which account for more than 90% of railway occurrences, regardless the severity of the event. Step four examines the associations between the variables that describe railway occurrences. The predictions based on past data are generated in step 5. Based on the findings, the broad applicability of the R-PSFs taxonomy is confirmed. The factors that should be acknowledged within the investigation process are determined. And the most significant factors that should be further studied in order to assess human performance in railway operations are established.

### 6.4.1 Descriptive statistics

The reports, as shown in Figure 6-6, were initially classified into 70 serious accidents (47 fatal), 189 accidents and 220 incidents, based on the ERA definitions (Section 3.2.1). Since raw data were categorised by each source, when migrated according to the EU definitions, changes in the classification are observed, e.g. US incident reports are now indicated as accidents. Subsequently, the lack of events for particular data sources is explained by such migration.

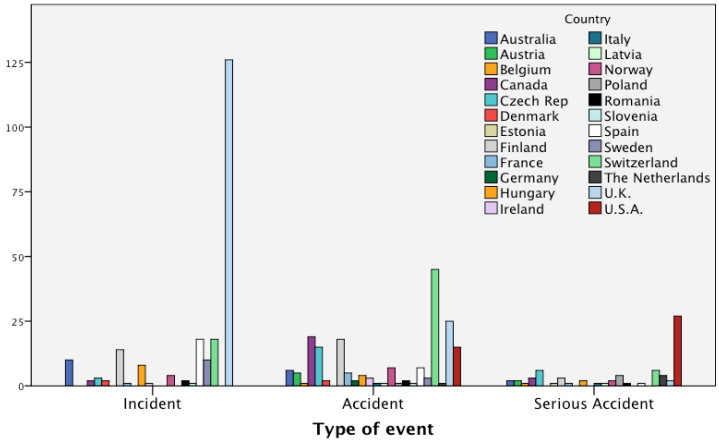



Figure 6-6 Type of collected reports per country of origin

Based on the information extracted from the reports (see Table 6-1) the set of relevant variables for the reference database was determined. Such variables are in agreement with the type of safety data collected by the FRA, ERA, CTSB, ATSB and the majority of railway national safety and investigation bodies, e.g. the U.K. RSSB and the Norwegian JBV. Subsequently, the events were also classified based on the: time and month of occurrence, responsible personnel, involved type of train, location and type of occurrence, immediate cause, contributing factors (R-PSFs), shunting operation, degraded operations, number of operators in train cabin and the installed protection systems.

Starting with the *time of day*, reports were divided into events that occurred during the day, night, at dawn and dusk, based on the local timing of events. However, as data on events at dawn and dusk was limited, these two day-periods were combined.

On rolling stock (*type of trains*), as displayed in Figure 6-7, the classes include: passenger trains, i.e. long-distance, regional; freight trains; metros; trams; engineering trains; engineering vehicles and locomotives running solo, with the majority of events to refer to passenger trains. However, due to the limited representation of particular types of trains within the selected reports, clustering of the rolling stock types was performed as follows: (i) passenger trains, (ii) freight trains, (iii) engineering trains, vehicles and locomotives running solo, and (iv) metros and trams.

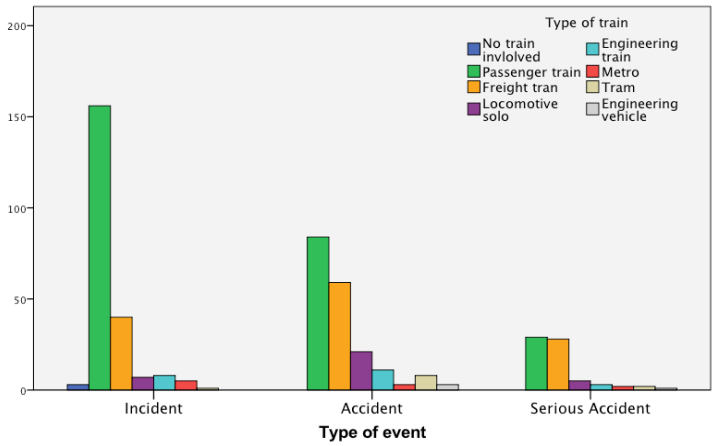



Figure 6-7 Type of train involved per type of event

With respect to the personnel involved (*responsibility*), as shown in the Figure 6-8, this was distinguished into passenger train drivers, i.e. long-distance and regional; freight train drivers; locomotive (solo) drivers, engineering train drivers, metro drivers, tram drivers, signallers and controllers. Again as some of the categories are limited, a new classification was generated based on the commonalities of the personnel as follows: (i) passenger train drivers; (ii) freight train drivers; (iii) locomotive (solo) drivers - engineering train drivers; (iv) metro - tram drivers; (v) signallers - controllers and (vi) more than one employees.

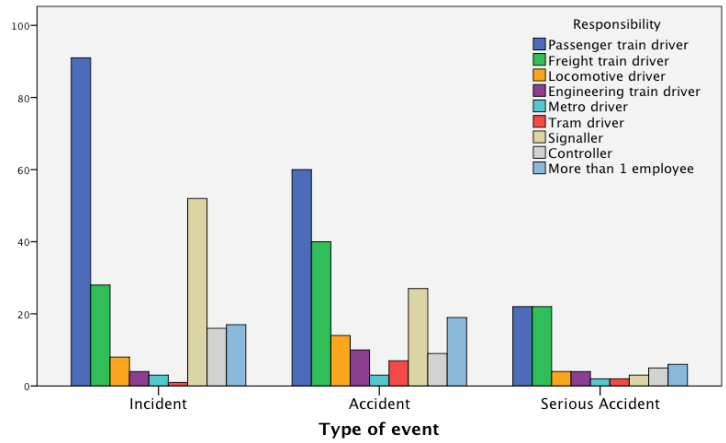



Figure 6-8 Type of responsible personnel per type of event

Concerning the *location* of events, reports revealed at least twenty different places that accidents may happen, such as: bridges, depots, depot points, junctions, level crossings, monorail lines, open lines, curvatures on open lines, open lines with gradient, points on open lines, run rounds, station platforms, stations (entrances and exits), points at stations, street overpasses, tunnels, underground lines, yards and viaducts. Similar to above variables, *location* was grouped based on common characteristics, in two different ways (Figure 6-9). In the first, location types were merged into: open lines (including curves, points or gradient),

stations (including station points, platforms, depots or depot points), yards, junctions, bridges or viaducts, tunnels (including entrance or the exit) and level crossings. In the second, a new category referred to as *points* was created, which comprises all cases of points and junctions. All other categories, not embraced within this new category, remained the same with the categories of location described in the first categorisation.

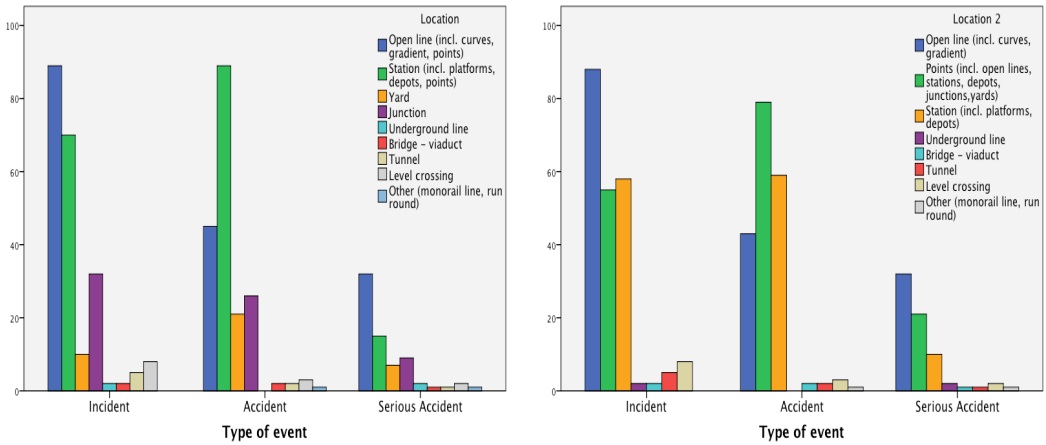



Figure 6-9 Type of location per type of event

The *type of occurrences* (Figure 6-10), similarly to ERADIS (2012) is classified according to the *type of events*: (i) for accidents and serious accidents are train collisions, train derailments, train collisions with obstacles and persons hit by train; (ii) whilst for incidents, are signals passed at danger (SPADs), SPADs (ground signal) and operational irregularities.

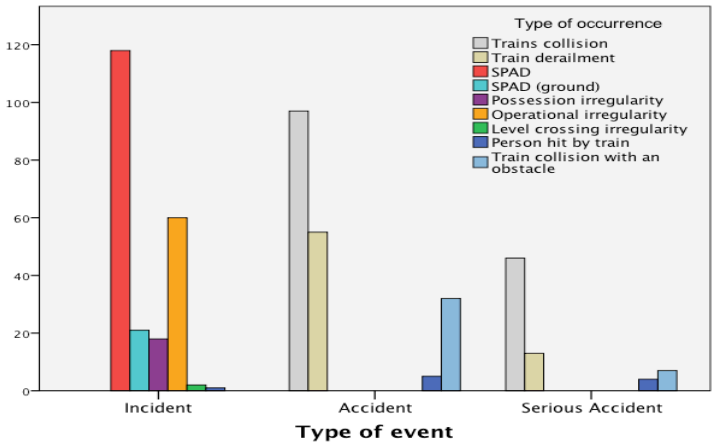



Figure 6-10 Type of occurrence per type of event

Operational irregularities include any type of irregularity caused by signallers or drivers, including possession irregularities or level crossing irregularities, which are not described as

SPADs and lead to an incident, e.g. wrong track authorisation by signaller. SPADs and operational irregularities are also used to describe the *associated events* to the *types of occurrences* for accidents and serious accidents, as displayed in the Figure 6-11.

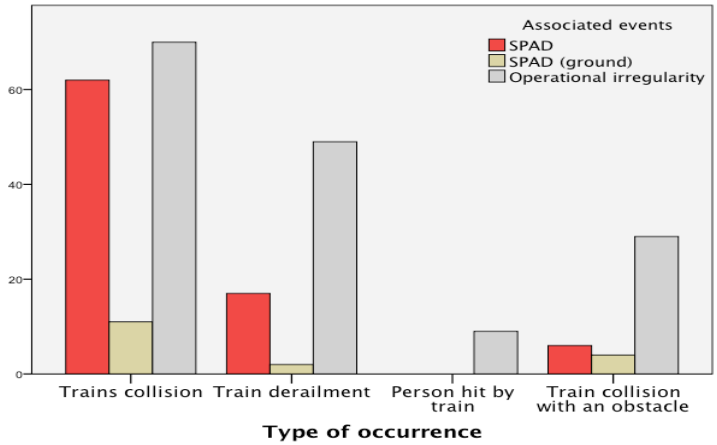



Figure 6-11 Associated events to types of occurrences

Subsequently, *immediate causes* (Figure 6-12) were distinguished into cases related to signalling (including no signal observation, wrong signal interpretation, disregard signal etc.), excessive speed, braking, distance estimation, routing, points setting and information. This classification is in general agreement with previous studies, e.g. (Evans, 2011).

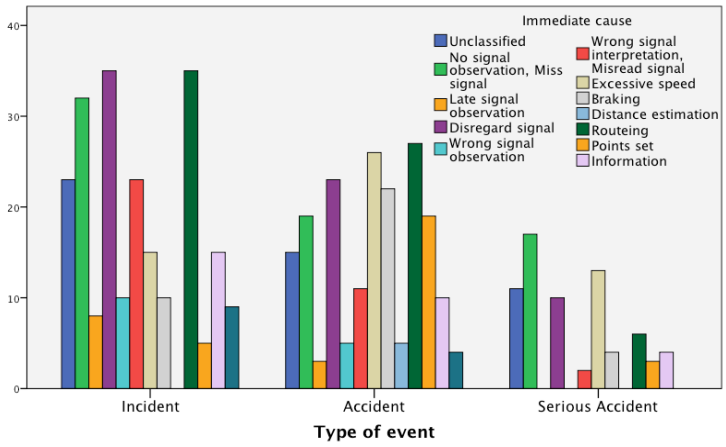



Figure 6-12 Type of immediate causes per type of event

With respect to the *railway protection systems* (Figure 6-13), these include the types of trains control used worldwide. However, as protection systems are implemented primarily on national level and their operational systems may, for this research their classification is based on their similar features to the: Automatic Warning System (AWS), Automatic Train Protection (ATP) and Train Protection Warning System (TPWS), as briefly described in Chapter 2.

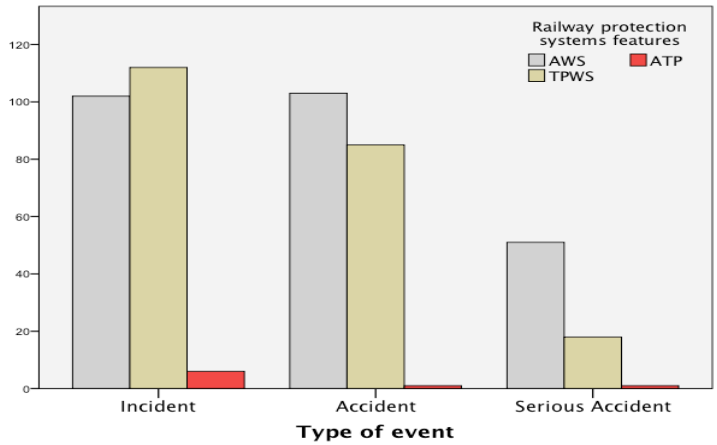



Figure 6-13 Type of railway protection system per type of event

In turn, *operational conditions (degraded operations)*, describe any type of emergency or special conditions applied at the time of an event, such as: rolling stock failure, lines blockage, infrastructure failures, technical communication failures, signalling failures, no signalling in the territory and speed restrictions. Possessions and maintenance work, although they can be characterised as *degraded operations*, they are not considered as such because they take place on daily basis and operators are familiar with operations under those conditions. Finally, the *number of operators* refers to train drivers and is distinguished into *one* or *more than one* drivers in the train cabin, while shunting operations describe any *"operation of moving a railway vehicle or set of railway vehicles inside a railway station or other railway installations (depot, workshop, marshalling yard, etc.) controlled by shunting signals"* (European Railway Agency, 2009, Eurostat et al., 2009).

Table 6-6 shows the reference database, which presents the findings after they were re-coded in order to reduce the subcategories of each variable to the relevant minimum number, when either were too many or their data was not sufficiently or applicable for further analysis.

Table 6-6 Descriptive statistics of reference database

| | | Type of event | | | | |
|---|---|---|---|---|---|---|
| | | Incidents (220) | Accidents (189) | Serious Accidents (70) | | Valid cases |
| Year | 1997 – 2005 | 25 | 55 | 30 | (110) | 99.8% |
| | 2006 – 2011 | 195 | 133 | 40 | (368) | |
| Month | Dec – Feb | 52 | 53 | 17 | (122) | |
| | Mar – May | 42 | 48 | 15 | (105) | |
| | Jun – Aug | 74 | 47 | 19 | (140) | 99.8% |
| | Sep – Nov | 52 | 40 | 19 | (111) | |
| Time of day | Day | 136 | 121 | 48 | (305) | |
| | Night | 9 | 52 | 19 | (130) | 100.0% |
| | Dawn – Dusk | 25 | 16 | 3 | (44) | |
| Responsibility | Passenger train driver | 91 | 60 | 22 | (173) | |
| | Freight train driver | 28 | 40 | 22 | (90) | |
| | Loco – Engineering train driver | 12 | 24 | 8 | (44) | 100.0% |
| | Metro – Tram driver | 4 | 10 | 4 | (18) | |
| | Signaller – Controller | 68 | 36 | 8 | (112) | |
| | More than one employee | 17 | 19 | 6 | (42) | |
| Type of train | Passenger train | 156 | 84 | 29 | (269) | |
| | Freight train | 40 | 59 | 28 | (127) | 99.4% |
| | Engineering – Locomotive | 15 | 35 | 9 | (59) | |
| | Metro – Tram | 6 | 11 | 4 | (21) | |
| Location | Open line (curves, gradient, points) | 89 | 45 | 32 | (166) | |
| | Station (platform, points, depots) | 70 | 89 | 15 | (174) | 92.9% |
| | Yard | 10 | 21 | 7 | (38) | |
| | Junction | 32 | 26 | 9 | (67) | |
| Location 2 | Open line (curves, gradient) | 88 | 43 | 32 | (163) | |
| | Station (platform, depots) | 58 | 59 | 21 | (138) | 92.9% |
| | Points (points, yards, junctions) | 55 | 79 | 10 | (144) | |
| Type of occurrence | Train collisions | - | 97 | 46 | (143) | |
| | Train derailments | - | 55 | 13 | (68) | |
| | Train collision with obstacle | - | 32 | 7 | (39) | |
| | Person hit by train | 1 | 5 | 4 | (10) | 99.6% |
| | SPAD | 118 | - | - | (118) | |
| | SPAD (ground) | 21 | - | - | (21) | |
| | Operational irregularities | 78 | - | - | (78) | |
| Associated event | SPAD | - | 52 | 33 | (85) | |
| | SPAD (ground) | - | 15 | 2 | (17) | 54.1% |
| | Operational irregularity | - | 122 | 35 | (157) | |
| Immediate cause | Signalling | 108 | 61 | 29 | (198) | |
| | Excessive speed | 15 | 26 | 13 | (54) | |
| | Braking | 10 | 22 | 4 | (36) | 86.0% |
| | Routing | 35 | 27 | 6 | (68) | |
| | Points | 5 | 19 | 3 | (27) | |
| | Information | 15 | 10 | 4 | (29) | |
| Operational conditions | Normal | 144 | 170 | 58 | (372) | 100.0% |
| | Special conditions | 76 | 19 | 12 | (107) | |
| Shunting operation | Yes | 21 | 38 | 9 | (68) | 100.0% |
| | No | 199 | 151 | 61 | (411) | |
| Number of operators | One | 217 | 164 | 47 | (428) | 100.0% |
| | More than one | 3 | 25 | 23 | (51) | |
| Railway protection systems | AWS | 102 | 103 | 51 | (256) | |
| | TPWS | 112 | 85 | 18 | (215) | 100.0% |
| | ATP | 6 | 1 | 1 | (8) | |

In addition to variables above, the contributing factors were also classified based on the R-PSFs taxonomy, where all R-PSFs categories and elements were taken into account. In total

1820 contributing factors were identified. From these, 1815 were allotted to the defined R-PSFs, while 5 were not classified; 800 reflect incidents, 668 accidents and the remainder serious accidents. That is about 3.7 R-PSFs per incident, 3.5 per accident and 5 per serious accident report. Figure 6-14 shows the complete R-PSFs frequencies per type of severity. It can be seen that in average the extracted R-PSFs compare to those displayed in Figure 6-2 and determined by the SMEs are same for almost 87.73%. That is, the predominant factors as indicated in Table 6-2 remain the same, e.g. safety culture is again the most frequent identified R-PSFs, while the increase in number of observations is proportional to the additional number of reports (varies from 1 to 35%).

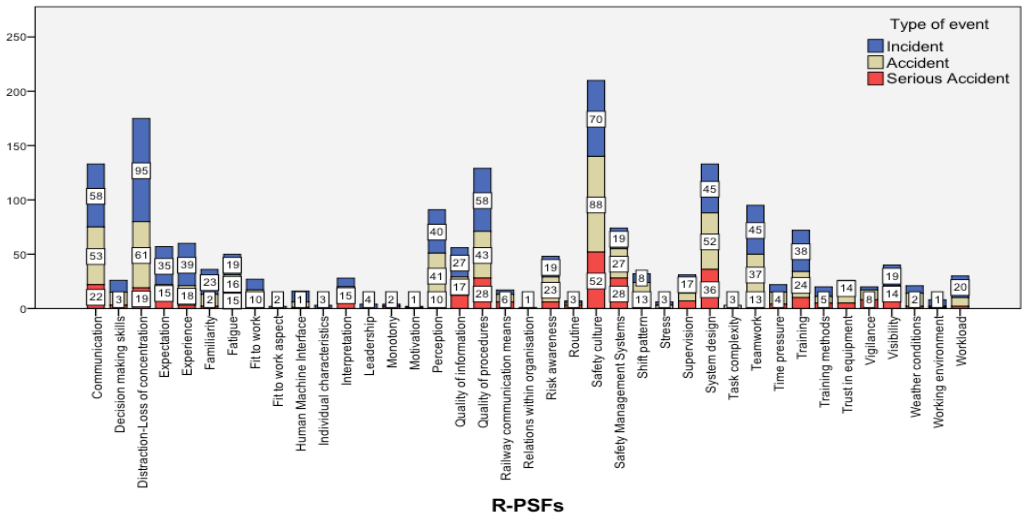



Figure 6-14 The complete list of the identified R-PSFs per type of event

In contrast, there are not anymore eleven non-identified R-PSFs, as shown in Section 6.3.1, but four. These are: *relations within the team, task instructions*, *and communication within organization and incentives for employees*. All but task instructions are hidden or implied factors and their constant absence from findings sustains the assumption of the gap in the investigation schemes concerning these types of factors. On the other hand, although *individual characteristics*, *motivation, monotony*, *task complexity, leadership, fit to work aspect* and *relation within organisation* were identified within the additional reports, they constitute only 1% of total identified R-PSFs (19 out of 1815). Moreover none of them was extracted more than four times, while all except monotony refer to incidents. However, even though monotony reflects two serious accidents, both of them were neither fatal or provoked material nor financial losses higher than 200,000 euros. Similar to the other relevant variables, due to the limited representation of particular R-PSFs elements, the clustering of factors was required for conducting any further analysis and excluding accurate results. Such clustering is described in the next sub-section.

### 6.4.2 The R-PSFs *lite* – combination of R-PSFs for further analysis

As the R-PSFs were again unequally identified, confidence is gained that the majority of railway occurrences it could be associated primarily with a limited number of factors that influence the performance of operators. Therefore, the Pareto principle (Backbaus, 1980) was re-applied to determine those factors that were most frequently indicated within the investigation reports considering the severity of events. Hence, starting with the list of factors as shown in Table 6-2, the 43 R-PSFs are re-combined based on the: given definitions, SMEs recommendations and frequency of observations. By merging the factors is aimed to (i) provide a simpler more flexible version of the R-PSFs taxonomy for both investigation purposes and statistical analyses, (ii) highlight the factors that primarily affect operators performance and keep both operators and investigators alert about these factors, and (iii) be used to quantify ("weight") the contribution of each significant R-PSF on the performance of operators.

Therefore, first, the R-PSFs were classified according to their frequency within the serious accidents then the accidents and finally the incident reports. In turn, the factors with higher frequencies and common characteristics were combined. For instance, with respect to *organisational factors*, it can be seen from Figure 6-14 that *safety culture* is the most often identified R-PSF related to serious accidents. Moreover, safety culture is by definition strongly linked to organisations *Safety Management Systems* (European Railway Agency, 2010), hence the union of these factors is not only suggested but also represents a large number of railway accidents. With respect to organisational factors no other factors were united, as either the factors were not frequently identified or their characteristics are not common. Moving towards *system factors*, *railway communication means*, *human machine interface* and *visibility (due to system design)* are absorbed by the predominant *system design*. Furthermore, for *team factors*, *teamwork* and *communication between employees* were migrated, whilst within *dynamic personal factors*, the primary *distraction-loss of concentration* embraces *situational awareness* and *vigilance* and *perception* is united with *interpretation* (Rail Safety and Standards Board, 2004b). Finally, accepting SMEs' recommendation (Section 6.3.2) *experience* is merged with *training*.

In addition to above combinations, which all refer to cases within same R-PSFs categories, factors from different categories were combined. For example, again based on SMEs' recommendations *workload* is united with *time pressure* and they both absorb *stress*, whilst *fatigue* includes *shift pattern* and *fit to work*. Similarly, *familiarity* is merged with *expectation* and *routine* and is integrated to a factor that primarily belongs to the personal category, as

familiarity is considered the most dominant factor among them. The combined factors represent 1460 out of 1815 or 80.4% of the identified R-PSFs.

Furthermore, the R-PSFs elements that were mostly extracted from reports were also determined, including *quality of procedures*, *quality of information* and *supervision*. Having added them to the combined list of factors, a list of 12 factors is established (Table 6-7) which represents in average 92.3% of total R-PSFs. In particular that list reflects 91.75% of R-PSFs related to incidents (Figure 6-15), 92.23% related to accidents (Figure 6-16) and 93.37 related to serious accidents (Figure 6-17). The 12 R-PSFs represent the most frequent identified contributing factors, regardless the severity of railway occurrences. Hereafter the list will be referred to as the R-PSFs *lite* taxonomy.

Table 6-7 The combined most frequent identified R-PSFs per type of event

| **The R-PSFs *lite*** | | | | | |
|---|---|---|---|---|---|
| R-PSFs elements | Categories | Incidents | Accidents | Serious Accidents | Total |
| Safety culture – SMS | Organisational | 89 | 115 | 80 | 284 |
| System design – HMI – RCM – Visibility | System | 71 | 71 | 45 | 187 |
| Fatigue – Shift pattern – Fit to work | *Dynamic personal* | 37 | 34 | 38 | 109 |
| Communication – Teamwork | Team | 103 | 90 | 35 | 228 |
| Quality of procedures | Organisational | 58 | 43 | 28 | 129 |
| Distraction – Loss of concentration – Vigilance – Situational awareness | Dynamic Personal | 117 | 93 | 33 | 243 |
| Perception – Interpretation | Dynamic Personal | 48 | 56 | 15 | 119 |
| Training – Experience | Personal | 77 | 42 | 13 | 132 |
| Familiarity – Expectation– Routine | *Personal* | 60 | 28 | 12 | 100 |
| Quality of information | Team | 27 | 17 | 12 | 56 |
| Supervision | Organisational | 17 | 7 | 7 | 31 |
| Workload – Time pressure – Stress | *Task* | 30 | 22 | 6 | 58 |
| | | 734 / 800 (803) | 618 / 668 (670) | 324 / 347 | 1676 / 1815 (1820) |

*italics* indicate clustering of factors that belong to different R-PSFs categories, parentheses show the total number of R-PSFs

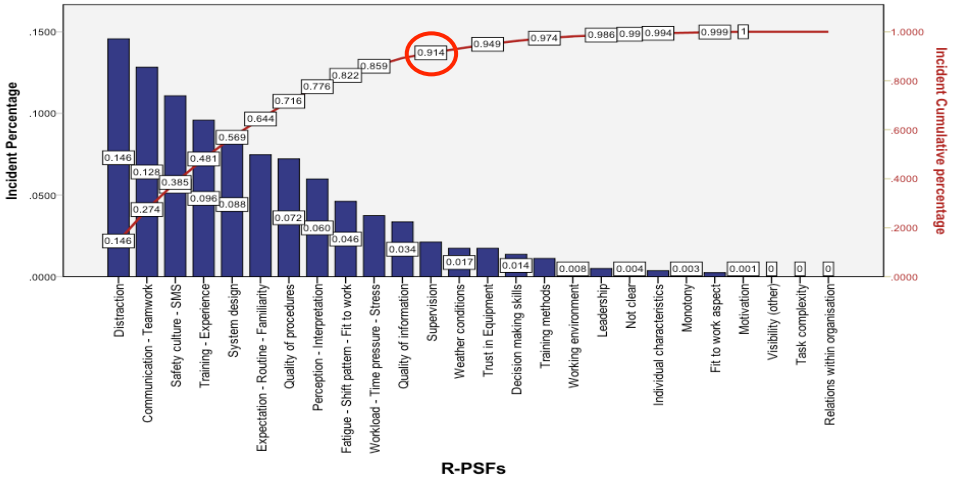



Figure 6-15 The identified R-PSFs for incidents

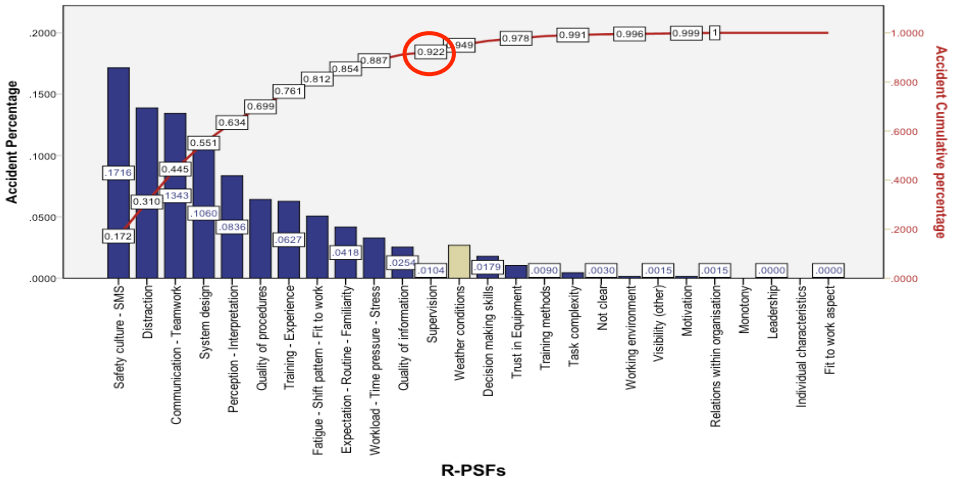



Figure 6-16 The identified R-PSFs for accidents

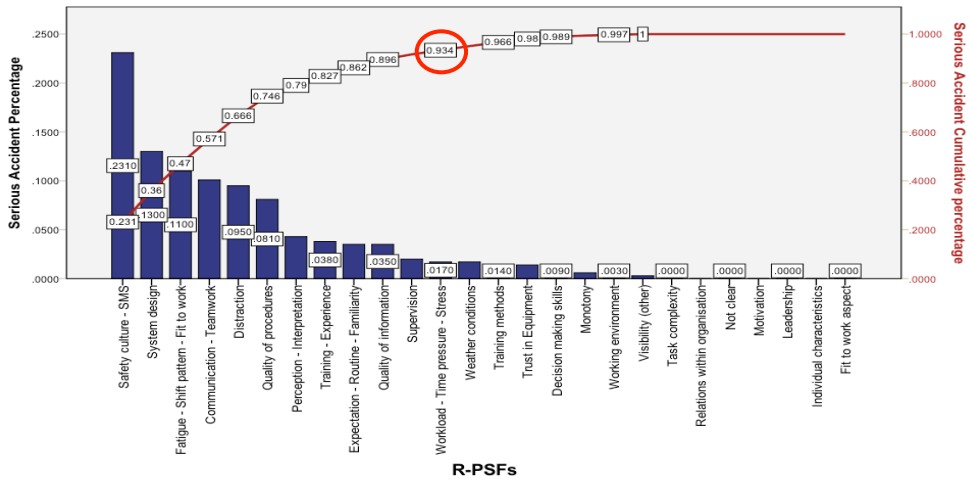



Figure 6-17 The identified R-PSFs for serious accidents

The R-PSFs *lite* taxonomy confirms entirely the Pareto principle (Backbaus, 1980) for all types of occurrences, while no one of the excluded factors is related to a fatal occurrence or material damage more than 200,000 €; this justifies why no more factors are included in the shorter version of the taxonomy. In turn, the R-PSFs *lite* was verified by two SMEs with profound knowledge in the field of Human Factors. The purpose of verification was to ensure that the combined factors constitute rational and acceptable (based on the given definitions) integrations that can be further used to describe and investigate railway accidents sufficiently and to study human performance within the context of such accidents. For this, two sessions with the HF experts, i.e. the Head of RSSB HFs group and the Chair of OWT group of ETH Zurich, were conducted, where the author presented the concept, suggestions and findings for the construction of the R-PSFs *lite.* Feedback from both experts verified that the combination of factors is rational and well justified, thus the R-PSFs *lite* can be used, to study the contribution and influence of each of the most significant R-PSFs on human performance. The R-PSFs *lite* is currently being implemented by the LUL (Appendix I), while it can be also used to define the impact of ERTMS technology on the performance of operators (Smith et al., 2013).

Having confirmed that the set of R-PSFs suggested in this thesis is representative for the railway industry and having also defined those factors that mostly contribute to railway occurrences by influencing the performance of operators, it is yet to be defined whether such list and factors can be broadly used in a variety of contexts, including different types of accidents, networks and involved personnel. Therefore, the examination of relationships between the variables that characterise the railway occurrences with the factors included in the R-PSFs *lite* is examined, as described in the next sections.

### 6.4.3 Examination of relationships between two variables

This section presents a series of statistical tests performed to explore associations between the factors that included in the R-PSFs *lite* and the variables that characterise railway events.

As variables are categorical, Pearson's chi-squared tests of independence are conducted, as described in Section 6.3.2. Once statistical significance is established, the Cramer's V test is used to measure the importance (strength) of the association. This test is applicable across contingency tables of size greater than 2x2 (Field, 2009) and calculated as:

$$V = \sqrt{\frac{\chi^2}{N \cdot m}} \quad (6\text{-}3)$$

where $\chi^2$ is the Pearson's chi-square statistic, $N$ the sample size and $m$ the smaller value between the number of rows reduced by one and number of columns reduced by one in the contingency table. Cramer's V varies between 0 and 1, where values close to 0 indicate little association and closer to 1 strong association. In particular, Cohen (1988) states that for $m$=1, Cramer's V≈0.1 indicates weak association, whilst V≈0.3 and V≈0.5 medium and strong respectively; as $m$ increases, associations are considered stronger for smaller V values. Table 6-8 includes the summary of findings, where the established associations are indicated in addition to Pearson's and Cramer's V statistics and their values of significance, whilst the complete results can be found in Appendix VIII.

Table 6-8 Chi-square associations between R-PSFs lite and other variables

| Two-paired associations | Association | Pearson's test statistic | Cramer's V test statistic | Significance |
|---|---|---|---|---|
| Type of event – R-PSFs *lite* | ✓ | $\chi^2(22) = 89.612$ | $V = .164$, m = 2 | $p < .001$ |
| Fatal-Non fatal – R-PSFs *lite* | ✓ | $\chi^2(11) = 27.306$ | $V = .128$, m = 1 | $p < .01$ |
| Year – R-PSFs *lite* | ✓ | $\chi^2(11) = 45.152$ | $V = .164$, m = 1 | $p < .001$ |
| Month quarters – R-PSFs *lite* | × | $\chi^2(33) = 46.202$ | $V = .093$, m = 3 | $p > .05$ |
| Time of day – R-PSFs *lite* | ✓ | $\chi^2(22) = 47.976$ | $V = .121$, m = 2 | $p < .05$ |
| Responsibility – R-PSFs *lite* | ✓ | $\chi^2(44) = 263.323$ | $V = .188$, m = 4 | $p < .001$ |
| Location – R-PSFs *lite* | ✓ | $\chi^2(33) = 58.198$ | $V = .111$, m = 3 | $p < .05$ |
| Location 2 – R-PSFs *lite* | ✓ | $\chi^2(22) = 35.265$ | $V = .106$, m = 2 | $p < .05$ |
| Type of occurrence – R-PSFs *lite* | ✓ | $\chi^2(66) = 179.213$ | $V = .135$, m = 6 | $p < .001$ |
| Associated events – R-PSFs *lite* | ✓ | $\chi^2(22) = 58.669$ | $V = .178$, m = 2 | $p < .001$ |
| Immediate causes – R-PSFs *lite* | ✓ | $\chi^2(55) = 175.141$ | $V = .156$, m = 5 | $p < .001$ |
| Operational conditions – R-PSFs *lite* | ✓ | $\chi^2(11) = 38.654$ | $V = .152$, m = 1 | $p < .001$ |
| Shunting operation – R-PSFs *lite* | ✓ | $\chi^2(11) = 39.074$ | $V = .153$, m = 1 | $p < .001$ |
| Number of operators – R-PSFs *lite* | ✓ | $\chi^2(11) = 53.658$ | $V = .179$, m = 1 | $p < .001$ |
| Protection systems – R-PSFs *lite* (excluding ATP) | ✓ | $\chi^2(11) = 46.570$ | $V = .168$, m = 1 | $p < .001$ |
| Protection systems 2 – R-PSFs *lite* (incorporating ATP with TPWS) | ✓ | $\chi^2(11) = 42.330$ | $V = .159$, m = 1 | $p < .001$ |

However, although the nature of an established association in a 2x2 contingency table can be easily interpreted by observing the cell percentages or counts, in large contingency tables such interpretation is not always straightforward. Thus, a cell-by-cell comparison of observed and expected frequencies helps to better understand the nature of the association (Agresti, 2007). For this purpose, the standardised residuals for each cell are computed. Residuals are the errors between the model's expected frequencies and the observed frequencies. The standardised residual is then derived by dividing the residual with the square root of the expected frequencies as:

$$standardised\ residual = \frac{observed_{ij} - expected_{ij}}{\sqrt{expected_{ij}}} \quad (6\text{-}4)$$

where *i* and *j* are again the rows and columns in the contingency table. By computing the standardised residuals one can determine which cells within the contingency table are the major contributors to the significant chi-square value. If the value of the standardised residual lies outside ±1.96 then it is significant at p<.05, while if it lies outside ±2.58 and ±3.29 it is significant at p<.01 and p<.001 respectively. Finally, a positive or negative value shows whether the observed counts in a cell are above (+) or below (-) the expected (Sheskin, 2007, Field, 2009).

Starting with the pair of R-PSFs *lite* and type of events, Table 6-9 shows that the confirmed association is primarily driven by seven cells (shown in light grey). These are: (a) *safety culture*[10] – incident, (b) *safety culture* – serious accident, (c) *fatigue* – serious accident, (d) *distraction* – serious accident, (e) *training* – incident, (f) *training* – serious accident and (g) *familiarity* – incident.

Table 6-9 Chi-square summary table for R-PSFs lite and Type of events

| | | | Type of event | | | |
|---|---|---|---|---|---|---|
| | | | Incident | Accident | Serious Accident | **Total** |
| **R-PSFs *lite*** | Safety culture – SMS | Count | 89 | 115 | 80 | 284 |
| | | Expected Count | 124.4 | 104.7 | 54.9 | 284.0 |
| | | Std. Residual | **-3.2** | 1.0 | **3.4** | |
| | Fatigue – Shift pattern – Fit to work | Count | 37 | 34 | 38 | 109 |
| | | Expected Count | 47.7 | 40.2 | 21.1 | 109.0 |
| | | Std. Residual | -1.6 | -1.0 | **3.7** | |
| | Distraction – Loss of concentration – Vigilance – Situat. awareness | Count | 117 | 93 | 33 | 243 |
| | | Expected Count | 106.4 | 89.6 | 47.0 | 243.0 |
| | | Std. Residual | 1.0 | .4 | **-2.0** | |
| | Training – Experience | Count | 77 | 42 | 13 | 132 |
| | | Expected Count | 57.8 | 48.7 | 25.5 | 132.0 |
| | | Std. Residual | **2.5** | -1.0 | **-2.5** | |
| | Familiarity – Expectation – Routine | Count | 60 | 28 | 12 | 100 |
| | | Expected Count | 43.8 | 36.9 | 19.3 | 100.0 |
| | | Std. Residual | **2.4** | -1.5 | -1.7 | |

For the first cell, the value of the standardised residual suggests that among the survey reports refer to incidents, there were fewer reports than expected (z=-3.2, p<.01) that indicate *safety culture* as contributor factor to railway incidents. In contrast, there were more reports than expected (z=3.4, p<.001) among those describe serious accidents that indicate

[10] For the sake of brevity only the first R-PSFs of each combination included in the *R-PSFs lite* is written.

*safety culture* as contributor factor for railway serious accidents. Similarly, within the serious accidents reports there were more than anticipated (z=3.7, p<.001) that imply *fatigue* as major contributor to serious accidents. Moreover, *distraction* was identified to contribute significantly less than expected to serious accidents, whilst *familiarity* to more than predicted incidents. Finally, *training* found to participate in significantly less than expected serious accidents and significantly more than expected incidents. In other words, results indicate that safety culture and fatigue are significant contributors for serious accidents, while training and familiarity for incidents. However the strength of the association for this sample, which expresses the influence of one variable on the other, is considered weak. The complete Table 6-9, as well as all relevant contingency tables can be found in Appendix VIII. In turn, if one classifies the types of events into fatal and non-fatal, then the established association stems primarily from the significant contribution of *fatigue* (z=2.6, p<.01) in fatal accidents (Table VIII-2). However, again this association is found weak.

Moving to the association between R-PSFs *lite* and the years of events (Table VIII-3), this is primarily driven by the relation between the factors *system design*, *fatigue*, *perception* and *training* with the period prior 2006. In particular, within the reports refer to 1997-2005, *system design* (z=2.5, p<.05), *fatigue* (z=2.4, p<.05) and *perception* (z=2.6, p<.01) were found significantly more times than expected to affect human performance and lead to railway events[11]. In contrast, results suggest that for the same period *training* (z=-3.0, p<.01), was identified in significantly less reports that one would predict. Findings also imply improvements on the design of the system along the years, which subsequently explains the identified improvement on the perception of operators. In addition, better scheduling of shift patterns as well as the investment on studies related to fatigue, e.g. (Härmä et al., 2002, Rail Safety and Standards Board, 2006), have undoubtedly decreased the incidents caused by fatigued operators. However, similar to the type of events, the strength of the association between R-PSFs *lite* and the period of those occurrences is weak. Moreover, regarding the time of railway events (Table VIII-5), the established association is almost exclusively driven, as probably expected, by the role of *fatigue* on human performance at night operations. Specifically, it was found that *fatigue* is involved to significantly more events (z=3.2, p<.01) than anticipated amongst the reports that describe night operations. On the contrary, within the reports that refer to events during the day, *fatigue* was indicated significantly less times

[11] Unless otherwise stated, hereafter the term *event* will be using to describe all types of railway occurrences, i.e. incidents, accidents and serious accidents.

(z=-2.2, p<.05) as contributing factor than expected. However, again the strength of the association is weak.

Concerning the determined association between the responsible personnel and R-PSFs *lite*, Table VIII-6 indicates that nineteen cells are primarily responsible for this relationship. Firstly, clear association is observed between R-PSFs and types of operators, i.e. signallers and train drivers. For instance, within the reports that refer to events caused by signallers or controllers, *workload* (z=4.1, p<.001), *communication* (z=5.5, p<.001), *supervision* (z=2.3, p<.05) and *quality of procedures* (z=2.0, p<.05) were found to contribute to significantly more events than expected. On the contrary, the factors *system design* (z=-3.6, p<.001), *fatigue* (z=-2.0, p<.05), *distraction* (z=-2.9, p<.01) and *familiarity* (z=-2.9, p<.01) were significantly less identified within the same reports. In contrast, among the reports refer to train drivers' errors; it is now *distraction*, *familiarity*, *system design*, *fatigue* and *perception* that found to contribute to railway events significantly more than expected. Moreover, association was again found between R-PSFs *lite* and the different type of drivers. That is, *distraction* (z=3.2, p<.05), *familiarity* (z=2.2, p<.05) and *perception* (z=2.1, p<.05) were involved to significantly more events than expected for passenger train drivers. *Fatigue* in contrast was found in significantly more reports (z=5.7, p<.001) refer to freight train drivers, but also in significantly less (z=-2.0, p<.05) reports involve passenger train drivers. In turn, *system design,* which primarily expresses the location of signals along the route, was almost equally identified to contribute to significantly more than expected to events involve passenger and freight train drivers. Finally, *safety culture* was found to be a significant (z=2.1, p<.05) contributor factor for locomotive-engineering train drivers. This may imply that when drivers operate without either transferring people or goods may be more prone to disobey rules and procedures. The strength of the association, between the two variables, is considered medium as the equivalent to 0.188 Cramer's V for m=4 is circa 0.33 for m=1 (Cohen, 1988, p.222).

With respect to the locations of events and their relationship with the factors that affect human performance, results show that for both classifications, i.e. *Location* and *Location 2*, the established associations are primarily driven by very few cells (Tables VIII-7 and VIII-8). In particular, for *Location 2* was found that for only one cell, *system design* – open line, the extracted frequencies were significantly more than the expected. This is explained by the fact that this classification includes all areas with points within one group and subsequently, information for specific areas that involve points cannot be revealed. On the other hand, using the classification *Location*, it was found that the association is mainly driven by the relationship between *system design*, *familiarity (routine)* and *training* with specific areas. In particular, *system design* was found involved in significantly more (z=2.0, p<.05) reports

about open lines than expected, which is in general agreement with the association of *system design* with passenger and freight train drivers. In addition, *routine* was extracted from significantly more ($z=2.7$, $p<.01$) reports than anticipated amongst events at yards, which is explained by the fact that the processes on those areas are standard while operators repeatedly execute the same tasks several times per hour. *Training*, finally, was derived from significantly more ($z=2.2$, $p<.05$) investigation reports, within those describe events in stations, than predicted; primarily this observation refers to events caused by new train drivers who miss signals within a complex station layout. However, again, results indicate that the strength of the association between the area of events and the relevant R-PSFs is weak.

On the contrary, the association between R-PSFs *lite* and types of railway occurrences is considered medium (Cohen, 1988, p.222), while results suggest that there are fourteen cells primarily responsible for this association (Table VIII-9). In particular, it is shown that within the reports that refer to train collisions, *safety culture* ($z=2.8$, $p<.01$), *system design* ($z=2.5$, $p<.05$), and *fatigue* ($z=2.2$, $p<.05$) were identified as contributing factors in significantly more reports than predicted. On the contrary, *safety culture* ($z=-3.5$, $p<.001$) was found to contribute to significant less SPAD events than expected. Furthermore, *communication* ($z=2.4$, $p<.05$) and *supervision* ($z=3.9$, $p<.001$) were both found to largely contribute to possession irregularities, justifying the association between operators and R-PSFs. Moreover, results show that among the survey reports refer to SPADs, *distraction* ($z=2.0$, $p<.05$), *familiarity* ($z=4.7$, $p<.001$) and *training* ($z=2.5$, $p<.05$) were identified as contributing factors significantly more times than expected. Finally, it is worth mentioning that no large discrepancies between observed and expected frequencies for any R-PSF *lite* amongst the reports that describe train derailments were identified.

Regarding the relationship between R-PSFs *lite* and associated events (Table VIII-10), it is the cells *fatigue* – SPAD, *communication* – SPAD and *perception* – SPAD (ground) that primarily suggest the established association. Specifically it was found that within the reports that describe train collisions, which were associated with SPADs, *fatigue* ($z=3.0$, $p<.01$) was indicated as contributing factor to significantly more cases that one could predict. Nevertheless, this association, opposite to the one between the type of occurrences and R-PSFs, is considered weak.

The association between R-PSFs *lite* and immediate causes was then examined (Table VIII-11), indicating an association with medium strength (Cohen, 1988, p.222). Fourteen cells mainly drive the association; in particular, within the reports refer to signalling, *familiarity*

(z=3.6, p<.001) and *system design* (z=2.1, p<.05) were indicated significantly more times than expected as contributing factors. Furthermore, *supervision* was extracted significantly more times (z=2.6, p<.01) than predicted amongst reports refer to excessive speed events, whilst *workload* amongst reports describe routing (z=2.3, p<.05) and setting of points (z=2.4, p<.05) occurrences. Especially with respect to routing, it is noticed that *safety culture* is indicated as significant contributor in significantly more (z=2.1, p<.05) cases than expected; this in conjunction with the established association between *workload* and *routing* and the fact that it is the signallers responsible for routing, implies that signallers under increased workload, may tend to not follow and obey (not necessarily intentionally) the relevant procedures and rules as required.

With respect to operational conditions (Table VIII-12), a weak association has been found, stemmed mostly from the relation between *quality of procedures*, *information* and level of *workload* under degraded operations. It was found that amongst reports that refer to degraded operations all three factors have been encountered significantly more than expected, at p<.05, to affect human performance. As regards shunting operations (Table VIII-13), on the other hand, results indicate that the established weak association is driven by four cells which all refer to occurrences during shunting operations. Specifically, it was found that *safety culture* (z=2.1, p<.05) and *communication* (z=3.3, p<.001) have been identified in significantly more reports than expected to negatively contribute to human performance, whilst *fatigue* (z=-2.5, p<.05) and *system design* (z=-2.0, p<.05) in significantly less.

Furthermore, concerning the number of operators (Table VIII-14), the association is derived mainly due to the relationship between *system design*, *fatigue*, *distraction* and *training* when two operators are on board, however is again considered weak. Results are complimentary and show that on operations with two train drivers, which usually refer to long haul freight train journeys, *fatigue* (z=4.5, p<.001) due to long and repetitive shift patterns was discovered to contribute significantly more than expected. In addition, as these types of operations include territories characterised by sections not controlled by signals, known as dark, *system design* (z=2.9, p<.01) was also encountered to contribute to significantly more reports than one would expect. On the other hand, operations with two drivers on board can be claimed to overcome and prevent occurrences due to *distraction* (z=-2.1, p<.05) or lack of *training* (z=-2.5, p<.05).

Finally, with respect to the railway protection systems, the ATP related cases, due to their limited representation were incorporated with the TPWS reports. Findings (Table VIII-16)

indicate that five cells are primarily responsible for the significant chi-square value. Particularly, it was observed that *quality of procedures* (z=-2.2, p<.05), *fatigue* (z=-2.4, p<.05) and *supervision* (z=-2.1, p<.05) were significantly less than expected identified as contributing factors within those events that involve either the TPWS or ATP systems. Although no direct explanation is possible for this finding, it could be only speculated that organisations which have installed ATP and TPWS systems, being more aware about the dangers during the operations, provide better supervision to their employees, while they also monitor and schedule more efficient shift patterns. On the other hand, findings show that *perception* was contributed to significantly more (z=2.5, p<.05) events related to ATP or TPWs and to significantly less (z=-2.1, p<.05) relevant to AWS. This can be explained by the fact that drivers feel safer, more confident and secure knowing that even if they make an error the consequences will not be severe, as the protection system will automatically stop their train. Nevertheless, again the determined association is classified weak.

In addition to the above tests, similar tests were performed to identify any associations between the type of events and the remainder variables, as shown in the Table 6-10.

Table 6-10 Chi-square associations between Type of event and other variables

| Two-paired associations | Association | Pearson's test statistic | Cramer's V test statistic | Significance |
|---|---|---|---|---|
| Type of event – Year | ✓ | $\chi^2(2) = 35.495$ | $V = .273$, m = 1 | p < .001 |
| Type of event – Month | × | $\chi^2(22) = 18.786$ | $V = .140$, m = 2 | p > .05 |
| Type of event – Time of day | × | $\chi^2(4) = 3.477$ | $V = .060$, m = 2 | p > .05 |
| Type of event – Responsibility | ✓ | $\chi^2(10) = 36.036$ | $V = .194$, m = 2 | p < .001 |
| Type of event – Type of train | ✓ | $\chi^2(6) = 41.858$ | $V = .210$, m = 2 | p < .001 |
| Type of event – Location | ✓ | $\chi^2(6) = 28.256$ | $V = .178$, m = 2 | p < .001 |
| Type of event – Location 2 | ✓ | $\chi^2(4) = 26.335$ | $V = .172$, m = 2 | p < .001 |
| Type of event – Type of occurrence | ✓ | $\chi^2(12) = 477.718$ | $V = .715$, m = 2 | p < .001 |
| Type of event – Type of occurrence (excluding incidents) | × | $\chi^2(3) = 7.087$ | $V = .165$, m = 1 | p > .05 |
| Type of event – Associated events | ✓ | $\chi^2(2) = 9.789$ | $V = .194$, m = 2 | p < .01 |
| Type of event – Immediate causes | ✓ | $\chi^2(10) = 42.517$ | $V = .234$, m = 2 | p < .001 |
| Type of event – Operational conditions | ✓ | $\chi^2(2) = 11.105$ | $V = .152$, m = 1 | p < .01 |
| Type of event – Shunting operation | ✓ | $\chi^2(11) = 9.428$ | $V = .140$, m = 1 | p < .01 |
| Type of event – Number of operators | ✓ | $\chi^2(11) = 57.548$ | $V = .347$, m = 1 | p < .001 |
| Type of event – Protection systems (excluding ATP) | ✓ | $\chi^2(2) = 14.514$ | $V = .176$, m = 1 | p < .01 |
| Type of event – Protection systems 2 (incorporating ATP with TPWS) | ✓ | $\chi^2(2) = 15.039$ | $V = .177$, m = 1 | p < .01 |

Results show that the *type of event* is associated to all relevant variables except *month* and *time of day*. However, not all of associations provide useful outcomes. For instance, the

majority of the selected incident reports are referred to the period post 2005, therefore association between the *type of event* and *year* of event is rather expected, while by definition is expected association between the types of occurrences and the type of events.

Findings indicate association between *responsible personnel* and *type of event*, with freight train drivers to get involved to significantly more ($z=2.4$, $p<.05$) serious accidents than expected. In addition, signallers were found to be both, significantly more ($z=2.3$, $p<.05$) responsible for incidents and significantly less responsible for serious accidents ($z=-2.1$, $p<.05$) than predicted. Similar findings were derived also when the association between the *type of event* and *type of trains* is examined. It was then found that incidents reports involve significantly more ($z=2.9$, $p<.01$) than expected passenger trains, while accident reports significant less ($z=-2.4$, $p<.05$). In addition, it was observed that freight trains were involved in significantly more ($z=2.2$, $p<.5$) than anticipated serious accidents and significant less ($z=-2.4$, $p<.5$) incidents. Finally, amongst the events involve locomotives-engineering trains, it was found that this category of trains was involved in significantly more ($z=2.4$, $p<.5$) accidents and significantly less incidents ($z=-2.3$, $p<.5$) than predicted.

With respect to location, the association is primarily driven by significant more cases than expected of accidents in stations ($z=2.4$, $p<.5$), less than expected incidents on open lines ($z=-2.6$, $p<.1$) and less than expected serious accidents in stations ($z=-2.1$, $p<.5$). Findings also show that 45% of accidents in stations take place at points, while only 15% occur at the platforms.

As already stated, by definition, association was expected between the type of occurrences and type of events. However, when incidents are excluded from the equation, then no association is observed between the different types of accidents and the severity of accidents. Nevertheless, train collisions provoke 67% of accidents and at least 30% of serious accidents, while derailments lead to less than 20% of serious accidents and almost 50% of accidents. SPADs and operational irregularities on the other hand, are responsible for 62% and 27% of the incidents respectively. Furthermore with respect to the associated events, it is the SPADs, which were found in significantly more ($z=2.1$, $p<.05$) than expected reports related to serious accidents, the prime reason of the established association.

Although only few of events occurred under *degraded operations* clear association between them and the type of events is observed. This association is driven by the significantly more ($z=2.2$, $p<.05$) than expected incidents and significant less ($z=-2.1$, $p<.05$) than excepted accidents under those conditions. This could be explained firstly by the fact that when one

operates under special conditions, strict restrictions are applied, e.g. speed restrictions, therefore the time and capability to react and/or correct errors increases. Secondly operators tend be more vigilant and alert, therefore the number of incidents that convert into accidents or serious accidents remains low. Finally, the established association between shunting operations and type of events, is forced primarily by the fact that within accidents reports, more shunting operations related events than expected ($z=2.2$, $p<.05$) were identified.

### 6.4.4 Examination of relationships between more than two variables

In the previous section it was shown that R-PSFs *lite* are associated with the different types of events, as well as with the majority of variables that describe the overall context of accidents. This section explores the simultaneous association of more than two variables, i.e. higher-order interactions. Therefore, loglinear analysis is used, which is an extension of the chi-square technique for multi-way contingency tables (Agresti, 2007, Field, 2009). In loglinear models all variables are treated as responses, i.e. no variable is selected as the independent, while only the associations between them are analysed. The term loglinear stems from the model's formation, where the natural logarithm of cell counts in the contingency table is expressed by the linear function of the effects of categorical variables as well as their interactions (Dupuy, 2011). For example, for three categorical variables, *X*, *Y* and *Z*, with *i*, *j* and *k* categories respectively the full (saturated) loglinear model, which describes all possible interactions is:

$$log\ (\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijK}^{XYZ} \qquad (6\text{-}5)$$

where, $\log(\mu_{ijk})$ is the log of expected cell frequency in cell *ijk*, $\lambda$ is the overall mean of the log of expected frequencies and the $\lambda$ terms represent the ***effects*** which the variables have on the cell frequencies. For example, the second term on the right hand side of the equation, $\lambda_i^X$, is the main effect for variable X, while a high value of $\lambda_i^X$ translates into a high value of the expected frequency is in row *i*. (Agresti, 2007).

Simpler models, containing a subset of the parameters involved in the saturated model are displayed in Table 6-11 (Dupuy, 2011).

Table 6-11 Examples of loglinear models for three-way contingency tables (Dupuy, 2011)

| Loglinear model | Description of model |
|---|---|
| $log(\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$ | all three categorical variables are mutually independent (*X, Y, Z*) |
| $log(\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY}$ | X and Y are associated but both are independent of Z (*XY, Z*) |
| $log(\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$ | X and Y are conditionally independent, given Z (*XZ, YZ*) |
| $log(\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$ | there is no three-way interaction (*XY, XZ,YZ*) |

To apply loglinear analysis, all possible meaningful pair associations between variables were investigated. This is to define which of those variables could be further involved into higher associations, as shown in Table 6-12. Subsequently, higher-order associations for the following set of variables were examined: *type of events and R-PSFs lite* with *time of day*, *responsibility*, *location*, *immediate causes*, *degraded operations*, *shunting operation*, *railway protection systems* and *number of operators*. The excluded variables were not taken into account, as (i) they are redundant variables, i.e. *type of occurrences – associated events,* and (ii) would provide expected outcomes, *e.g. type of occurrences* classification by definition distinguishes accidents and incidents cases.

Table 6-12 Pair-wise associations amongst variables

| Associated variables | R | ToT | L | $L_2$ | ToO | AE | IC | OC | SO | NoO | RPC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Time of day (ToD) | ✓ | ✓ | × | × | ✓ | × | × | × | × | ✓ | ✓ |
| Responsibility (R) | | *d* | ✓ | ✓ | × | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Type of train (ToT) | | | ✓ | ✓ | ✓ | ✓ | × | × | ✓ | ✓ | ✓ |
| Location (L) | | | | *d* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Location 2 ($L_2$) | | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | × |
| Type of occurrence (ToO) | | | | | | *d* | ✓ | ✓ | ✓ | ✓ | ✓ |
| Associated events (AE) | | | | | | | ✓ | ✓ | ✓ | ✓ | × |
| Immediate cause (IC) | | | | | | | | ✓ | × | × | × |
| Operational conditions (OC) | | | | | | | | | ✓ | × | × |
| Shunting operation (SO) | | | | | | | | | | × | - |
| Number of operators (NoO) | | | | | | | | | | | ✓ |
| Railway protection systems (RPC) | | | | | | | | | | | |

✓: association, × : indicates no association, *d*: dependency

The analysis focused on defining whether the established pair association between R-PSFs and types of events can be simultaneously related to one or more of the remaining variables. Within this context, seven tests were run to identify any three-way associations between the R-PSFs *lite*, the type of event and each one of the remainder variables. For any established

associations, four-way associations were then explored amongst the variables. For all tests, the goodness-of-fit statistic was calculated, to examine the hypothesis that the frequencies predicted by the model are significantly different from the observed frequencies. Results show that the difference between the expected and observed frequencies for all tests is not significant, i.e. p>.05, thus the hypothesis is rejected and the model accepted as a good fit good to the sample data.

Table 6-13 presents a typical output example of the *K-Way and Higher-Order Effects*, as derived by the IBM SPSS Statistics 19.0.0, for three variables. This table provides substantial information about which, if any, components of the model could be removed without having a significant impact on its ability to predict the observed data. The table has two parts. The first shows whether removing the K-Way, i.e. main effects, two-way interactions, three-way interactions, and higher-order effects will compromise the fit of the model. The second expresses the same, without however accounting for the removal of the higher-order effects (Field, 2009).

Table 6-13 K-Way and Higher-Order Effects for three variables

| | K | df | Likelihood Ratio | | Pearson | | Number of Iterations |
|---|---|---|---|---|---|---|---|
| | | | Chi-Square | Sig. | Chi-Square | Sig. | |
| K-way and Higher Order Effects[a] | 1 | 155 | 1659.762 | .000 | 1900.639 | .000 | 0 |
| | 2 | 138 | 322.441 | .000 | 317.821 | .000 | 2 |
| | 3 | 72 | 111.980 | .002 | 106.838 | .005 | 4 |
| K-way Effects[b] | 1 | 17 | 1337.321 | .000 | 1582.818 | .000 | 0 |
| | 2 | 66 | 210.461 | .000 | 210.983 | .000 | 0 |
| | 3 | 72 | 111.980 | .002 | 106.838 | .005 | 0 |

a. Tests that k-way and higher order effects are zero. b. Tests that k-way effects are zero.

Each row in Table 6-13 represents the influence of a particular effect on the model. For each row the Pearson chi-square statistics are calculated. When statistically significant, i.e. p<.05, then any removal of the effects that represent that row affects the fit of the model.

In this particular example, the first row of the first part of the table (K = 1) shows whether the removal of the one-way effects and any higher-order effects, i.e. everything is removed from the model, will significantly affect the fit of the model. Similarly, the second row (K = 2) indicates whether removing the two-way interactions and any higher-order effects will affect the model. Finally, the third row (K = 3) examines whether the removal of the three-way effect will influence model's fit. Since, the chi-square test for the first row is significant, the removal of everything from the model would significantly compromise its ability to predict the observed data. Similarly, if the statistic for the second row is significant, then removal of any two-way and three-way interactions would have a significant detrimental effect on the model.

Finally, if the statistic for the third row is significant, removal of the three-way interaction would affect the fit of the model.

The first row of the second part of the table examines whether removing the main effects from the model would have a significant effect on it. As can be seen in the table, the statistics for this row is significant with p<.05. Hence, if the main effects are removed the model will be significantly affected. The finding indicates that one or more of the effects are significant predictors of the data. Likewise, the statistic for the second row indicate that one or more of the two-way interactions is significant predictor of the data, and hence should not be removed from the model. Finally, the statistic for the third row provides evidence that the interaction between the three variables is a significant predictor for the data. Therefore, its removal would have a significant detrimental on the model.

Results of all loglinear tests can be found in the Appendix IX (Tables IX-1 to IX-7), whilst they are summarised in Table 6-14. Findings show that three tests indicate three-way interactions amongst variables. These are: (i) type of event – R-PSFs *lite* – location, (ii) type of event – R-PSFs *lite* – shunting operations and (iii) type of event – R-PSFs *lite* – railway protection systems. For those variables, four-way interactions were investigated, however for none of the combinations of variables such interactions were confirmed (K=4, p>.05) (Tables IX-8 to IX-10, Appendix IX). As a result, no higher-order interactions were explored.

Table 6-14 Three-way associations amongst variables

| Model | Three-way association | df | Chi-square for K=3 | Significance | Valid cases |
|---|---|---|---|---|---|
| Type of event – R-PSFs *lite* Responsibility | ✗ | 120 | 106.878 | p>.05 | 1676 |
| Type of event – R-PSFs *lite* Location | ✓ | *72* | *106.838* | *p<.05* | 1563 |
| Type of event – R-PSFs *lite* Immediate causes | ✗ | *168* | *128.823* | p>.05 | 1446 |
| Type of event – R-PSFs *lite* Shunting operations | ✓ | 24 | 57.150 | p<.05 | 1676 |
| Type of event – R-PSFs *lite* Special operational conditions | ✗ | 24 | 20.872 | p>.05 | 1676 |
| Type of event – R-PSFs *lite* Number of operators | ✗ | 24 | 28.266 | p>.05 | 1767 |
| Type of event – R-PSFs *lite* Railway Protection Systems | ✓ | 24 | 55.975 | p<.05 | 1676 |

✓ : association ✗ : indicates no association

### 6.4.5 Multivariate analysis

In the previous two sections the associations between the variables that characterize the different railway accidents were explored. Moving towards, this section examines further relationships between variables in a more structured manner. In particular, this section aims at generating predictions for the types of events (i.e. fatal – non-fatal, incidents – accidents – serious accidents) based on the involved R-PSFs. Thus, regression analysis is used to identify relationship between an outcome variable (the dependent variable, DV) and an independent variable (IV).

Multivariate logistic regression is broadly used in transport safety studies when the dependent variable is categorical and describes mutually exclusive states, e.g. incident/accident. The multivariate logistic regression is applied to model the linear relationship between the natural logarithm of the odds of the dependent variable and the independent variables (Dupuy, 2011). For a π probability of event, the odds are defined as the ratio between the probability of occurrence and non-occurrence, i.e. $\frac{\pi}{1-\pi}$, while the logistic regression model is expressed as:

$$log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \sum_{i=1}^{n} b_i\, X_i \qquad (6\text{-}6)$$

where π is the probability of the event of interest of the dependent variable, *α* is the constant, $X_i$ are the independent variables and bi their corresponding logistic regression coefficients (Agresti, 2007). The model uses the maximum likelihood method for the estimation of the parameters. This method maximises the probability of obtaining the observed values given the fitted regression coefficients (instead of the least square deviation criterion in linear regression analysis) (Field, 2009, p.267). In general, when a DV has only two categories the binomial logistic regression analysis is applied. On the other hand, if the DV has more than two response categories the multinomial logistic regression is used. Finally, in the cases where the DV is multi-categorical and ordinal, the ordinal logistic regression is used.

Two regression tests were conducted to generate predictions. In the first, the dependent variable is classified into *fatal – non-fatal* accidents, whilst the independent variable is the multi-categorical R-PSFs *lite* variable. For this test the binomial logistic regression is used. In the second test the DV is classified into *incidents - accidents - serious accidents*, while the IV is again the R-PSFs *lite* multi-categorical variable. Since this classification is not only nominal but also ordinal, the ordinal logistic regression is implemented.

### 6.4.5.1 Binomial logistic regression

Using the factors that affect operator performance predictions for *fatal – non-fatal* accidents were generated. Therefore, a test of the model that includes the predictors against the model that contains only the constant was performed. The test was found to be statistically significant ($\chi^2(11)=28.810$, $p<.01$), providing evidence that overall, the predictors can reliably distinguish whether an accident is fatal or not within the 479 reports. The overall percentage of the model's prediction was 87.1 (100 per cent for non-fatal and zero for fatal accidents), and the number of observations 1,676.

Table 6-15 shows the typical output for the predictors included in the model. The outcome variable is coded from 0 to 1, where DV(0) and DV(1) indicate the non-fatal and fatal accidents respectively. Similarly, each of the factors included in the IV is coded from 0 to 11, with IV(0) defining the reference category. Since *safety culture* was identified in the previous section as the most frequent contributing factor to both fatal and non-fatal accidents, it was chosen as the IV reference category.

Table 6-15 Variables in the equation for binary logistic regression

Compared to reference category DV(0)

b coefficients

Importance of the contribution of each predictor in the model

Odds ratio

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) Lower | Upper |
|---|---|---|---|---|---|---|---|---|---|
| DV(1) | IV | | | 25.295 | 11 | .008 | | | |
| | Workload - Time pressure - Stress | -1.740 | .737 | 5.574 | 1 | .018 | .176 | .041 | .744 |
| | System design - Visibility - HMI - RCM | -.062 | .255 | .060 | 1 | .806 | .939 | .571 | 1.547 |
| | Fatigue - Shift pattern - Fit to work | .328 | .280 | 1.371 | 1 | .242 | 1.388 | .802 | 2.404 |
| | Communication - Teamwork | -.547 | .268 | 4.183 | 1 | .041 | .578 | .342 | .977 |
| | Quality of procedures | -.163 | .295 | .308 | 1 | .579 | .849 | .477 | 1.513 |
| | Distraction - Loss of concentration - Vigilance - Situational awareness | -.487 | .258 | 3.551 | 1 | .060 | .615 | .370 | 1.020 |
| | Perception - Interpretation | -1.180 | .421 | 7.872 | 1 | .005 | .307 | .135 | .701 |
| | Training - Experience | -.319 | .304 | 1.100 | 1 | .294 | .727 | .400 | 1.320 |
| | Familiarity - Expectation - Routine | -.721 | .384 | 3.532 | 1 | .060 | .486 | .229 | 1.031 |
| | Quality of information | -.528 | .460 | 1.315 | 1 | .252 | .590 | .239 | 1.454 |
| | Supervision | -.317 | .559 | .322 | 1 | .571 | .728 | .244 | 2.177 |
| | Constant | -1.593 | .158 | 101.173 | 1 | .000 | .203 | | |

DV(1): Fatal, DV(0): Non fatal,
IV(0): Safety culture

Compared to reference category IV(0)

The results in Table 6-15 are interpreted as follows. Firstly, the Wald statistics indicate the importance of the contribution of each factor to the prediction of the outcome, considering the statistical significance (p-value given in the Sig. column). The larger the deviation of the Wald value from zero, the greater the importance of the factor in predicting the outcome. In this case, only workload, communication and perception are significant predictors ($p<.05$). The Exp(B) column provides the odds ratios (OR). This column indicates the change in the

odds resulting from a unit change in the predictor. If OR>1 then as the predictor increases the odds of the outcome also increase. In contrast, if OR<1 then as the predictor increases the odds of the outcome to occur decrease. Subsequently, compared to the reference category of safety culture, workload is 82.4% less likely to contribute to a fatal accident than non-fatal. Similarly, communication and perception are 42.2 and 69.3% respectively less likely, to contribute to a fatal accident than non-fatal.

However, although the model can be reliably used to predict fatal – non-fatal accidents, the contributing factors explain only a very small proportion, i.e. $R^2_{Nagelkerke}$ = 3.0% of the variation between the accidents. Given the fact that the model predicts more accurately non-fatal rather than fatal railway events, a limitation in the application of regression analysis to accident data might have been revealed. Fatal events along the years have become significantly more rare than non-fatal occurrences. Therefore, sample size might be too low to support such analysis, especially with respect to the factors that affect human performance.

#### 6.4.5.2 Ordinal logistic regression

To generate predictions for *incidents – accidents – serious accidents* using the R-PSFs *lite*, again a test of the model that includes the predictors against the model that contains only the intercept was performed. This test was statistically significant ($\chi^2(11)=70.182$, $p<.01$), indicating that overall, the predictors can reliably determine the severity of each of the railway accidents within the 479 collected events. The number of observations was again 1,676. In addition, the goodness-of-fit statistics was calculated (Section 6.4.4). The test was not significant, ($\chi^2(11)=18.573$, $p=.069$), and therefore, the predicted values do not significantly differ from the expected. Table 6-16 presents the output of the parameter estimates for ordinal logistic regression. The outcome DV is coded from 1 to 3, where DV(1), DV(2) and DV(3) represent the incidents, accidents and reference category respectively. *Serious accidents* were used as reference category for the DV since they constitute the most severe and critical events in terms of prevention. On the other hand, *safety culture* was again the reference category for the IV.

Table 6-16 Parameter estimates for ordinal regression analysis

Compared to reference category DV(3)

Coefficients

Importance of the contribution of each predictor in the model

| | | | | | | | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|---|
| | | Estimate | Std. Error | Wald | df | Sig. | Lower Bound | Upper Bound |
| Threshold | Incident | -.790 | .113 | 48.533 | 1 | .000 | -1.013 | -.568 |
| | Accident | .943 | .114 | 67.955 | 1 | .000 | .719 | 1.167 |
| Location | Workload - Time pressure - Stress | -.925 | .276 | 11.223 | 1 | .001 | -1.466 | -.384 |
| | System design - Visibility - HMI - RCM | -.264 | .174 | 2.285 | 1 | .131 | -.605 | .078 |
| | Fatigue - Shift pattern - Fit to work | .100 | .208 | .230 | 1 | .631 | -.308 | .508 |
| | Communication - Teamwork | -.641 | .166 | 14.814 | 1 | .000 | -.967 | -.314 |
| | Quality of procedures | -.510 | .198 | 6.631 | 1 | .010 | -.898 | -.122 |
| | Distraction - Loss of concentration - Vigilance - Situational awareness | -.760 | .164 | 21.369 | 1 | .000 | -1.083 | -.438 |
| | Perception - Interpretation | -.557 | .204 | 7.450 | 1 | .006 | -.957 | -.157 |
| | Training - Experience | -1.149 | .205 | 31.410 | 1 | .000 | -1.550 | -.747 |
| | Familiarity - Expectation - Routine | -1.173 | .228 | 26.563 | 1 | .000 | -1.618 | -.727 |
| | Quality of information | -.613 | .274 | 4.995 | 1 | .025 | -1.151 | -.075 |
| | Supervision | -.804 | .360 | 4.996 | 1 | .025 | -1.508 | -.099 |
| | Safety culture - SMS | 0[a] | . | . | 0 | . | . | . |

Link function: Logit
a. This parameter is set to zero because it is redundant
DV(3)=Serious Accident, IV(12)=Safety culture

Compared to reference category IV(12)

The results in Table 6-16 are read in a similar manner to those of binary regression. However, as the OR column is absent, it is the estimate column that is of interest in this case. In addition, threshold coefficients represent the intercepts, which are the points (in terms of a logit[12]) where factors might be predicted into the higher categories. For categorical IVs, when estimates are positive, the relevant IV category has a greater likelihood of higher score on the ordinal dependent compared to its reference level, i.e. the highest IV score.

In this test, all the factors except *system design* and *fatigue* are significant predictors for the model, and all the factors negative coefficients. With respect to *workload*, for example, the negative coefficient means that workload is more likely than safety culture to have lower dependent scores, i.e. to be involved in accident or incidents. However, coefficients do not directly provide information on the odds of each contributing factor. To derive the odds, $\exp^{-b}$ is calculated which in the case of workload is $\exp^{-(-.925)}$ = 2.52. In other words, *workload* is 2.52 times more likely to have answer of accidents or incidents compared to safety culture. In addition, it is also possible to calculate the likelihood of factors to be in a specific outcome

[12] A cumulative probability for Y is the probability that Y falls at or below a particular point. For an outcome category j, the cumulative probability is expressed by $P(Y \leq j) = \pi_1 + \cdots + \pi_2$, j=1,..,J. It is also $P(Y \leq j) + P(Y > j) = 1$. The logits of the cumulative probabilities are then defined as: $\log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \left(\frac{\pi_1 + \cdots + \pi_j}{\pi_{j+1} + \cdots + \pi_J}\right)$ (Agresti, 2007, p.180).

category using the category probabilities as derived after calculating the cumulative probabilities. The cumulative probability for workload, for example, to belong to the category serious accident is computed as:

$$P(Y > accident) = \frac{1}{1+exp(a_j - bx)} = \frac{1}{1+exp(0.943+.925)} = 0.133 \quad (6\text{-}7)$$

where $\alpha_j$ is the coefficient for the intercept where factors might be predicted in serious accidents and b the coefficient of workload.

In turn, the cumulative probability of workload to belong to “accident +” category is:

$$P(Y > incident) = \frac{1}{1+exp(-0.79+0.925)} = 0.466 \quad (6\text{-}8)$$

where the sign “+” indicates that the probability refers also to events with severity higher than accidents. Subsequently the category probability of workload to be exactly at the level of accidents is:

$$P(Y = accident) = P(Y > incident) - P(Y > accident) = 0.333 \quad (6\text{-}9)$$

Following the same rationale, the odds and probabilities of the rest factors can be derived.

However, although the model can be used to generate predictions for the three different types of accidents, similar to the case of fatal – non-fatal accidents, the used predictors explain only a relatively small proportion, i.e. $R^2_{Nagelkerke}$ = 4.7%, of the variation between the accidents. Although the proportion of events under the second classification is more equally distributed, the finding is rather expected, as the current investigation schemes are primarily focused on exploring the direct factors that lead to an accident rather the contributing and in particular those related to human factors and human errors.

### 6.4.6 Discussion of findings - Recommendations

The review of the complete set of reports aimed to confirm that the suggested taxonomy of 43 R-PSFs represents and can be implemented within a broader context of operations. Descriptive statistics show similarities with the analysis of the first set of data in terms of the identified R-PSFs and the frequency of their extraction for the different type of railway accidents. Thus, confidence is gained that the proposed taxonomy and the factors included in it capture the majority of operational railway accidents caused by human errors.

In addition, it was found that if the 43 factors are combined based on the given definitions and their common characteristics, 12 account for more than 90% of all railway events, regardless of their severity. These factors are: distraction, perception, fatigue, training, familiarity, workload, communication, safety culture, system design, supervision, quality of procedures and quality of information. A series of tests was performed to identify potential associations between those factors and the variables that characterise the different types of occurrences, including the location of events or the immediate cause if such events. Results show associations between almost all combinations of variables. For example, safety culture is mostly associated with serious accidents and fatigue primarily affects freight train drivers. The associations are however weak, supporting the concept that R-PSFs can be used for a broad context of operational accidents caused by human errors as indicators of the need for the development of effective solutions for improving operator performance.

In addition, predictions were generated in terms of the contribution of each R-PSF to a specific event type. Although R-PSFs can be efficiently applied to generate predictions, they explain only a very limited variance of the difference between the types of events. This finding was expected, as most of the current investigations schemes, apart from limited cases as implemented by Network Rail and SBB, do not contain a standard format in order to capture those factors that have contributed to accidents. Therefore, a simple check list (Smith et al., 2013) is introduced in this thesis. Such a list, should be included in the investigation schemes to capture and identify the most common contributing factors in a structured and consisted manner, as illustrated in the Table 6-17.

The list, due to its simplicity can be used by investigators even with limited experience in the field of human factors. However, all potential users should read first and understand the definitions of factors and what each represents. This will enable less biased data to be gathered, as all investigators will use the same list of factors, in a consistent format, using the same definitions and examples. Overall, the list will enhance the quality of collected data and reduce the biases amongst investigators and researchers. In addition it will improve the understanding and accuracy of the factors that affect human performance and may lead to railway incidents/accidents.

Table 6-17 The R-PSFs lite table to assist investigation reporting schemes

| Contributor factors Level I | | Main Contributing factors Level II | | Involved personnel |
|---|---|---|---|---|
| **Personal factors** | Yes No | Training – Experience | Yes | |
| | | Familiarity – Expectation | Yes | |
| | | Other (please indicate) | | |
| **Dynamic Personal factors** | Yes No | Distraction – Loss of concentration | Yes | |
| | | Perception – Interpretation | Yes | |
| | | Fatigue – Shift pattern | Yes | |
| | | Other (please indicate) | | |
| **Task factors** | Yes No | Workload – Time pressure – Stress | Yes | |
| | | Other (please indicate) | | |
| **Team factors** | Yes No | Communication – Teamwork | Yes | |
| | | Quality of Information | Yes | |
| | | Other (please indicate) | | |
| **Organisational factors** | Yes No | Safety culture – SMS | Yes | |
| | | Supervision | Yes | |
| | | Quality of procedures | Yes | |
| | | Other (please indicate) | | |
| **System factors** | Yes No | System design – HMI | Yes | |
| | | Other (please indicate) | | |
| **Environmental factors** | Yes No | Please indicate | | |

**SMS: Safety Management Systems, HMI: Human Machine Interface**

## 6.5 Summary

This Chapter presented validation of the Railway Performance Shaping Factors (R-PSFs) taxonomy. For this, 479 railway accident and incident reports were analysed while the findings were also verified by SMEs. Results indicate that no additional R-PSFs compared to those identified from the literature were either extracted or implied in the reports. Therefore, confidence is gained that the list of the 43 R-PSFs is representative for railway operations. However, it was also seen that the extracted factors do not include all the factors that are contained in the R-PSFs taxonomy. This could be explained either because no occurrence was related to such factors, or because they represent cases that are difficult to be identified within the investigation process.

In addition, the dependencies and interactions between the R-PSFs were determined, whilst, a comprehensive list of 12 factors was also defined, accounting for more than 90% of railway occurrences regardless the level of their severity. For the 12 factors, a series of statistical tests were conducted showing particular associations between the R-PSFs and the variables that characterise railway occurrences. However, such associations are weak,

and therefore, it can be claimed that the factors can be used for a broader context of railway operations.

There are possible uses of the 12 R-PSFs. The first is to incorporate them into the current investigation schemes to help inspectors define and describe the elements that contribute to railway accidents. Subsequently, the derived factors can be used to analyse accidents trends and determine safety critical areas and develop appropriate countermeasures. The second use is to estimate the proportional contribution of each R-PSF on human performance. For this, due to the lack of data and limitations of existing techniques, a new framework is developed and presented in the next Chapter. This framework, referred to as Human Performance Railway Operational Index (HuPeROI) allows researchers to assess the impact of the R-PSFs on the performance of the operators.

# Chapter 7 THE HUMAN PERFORMANCE RAILWAY OPERATIONAL INDEX

The previous chapter presented the Railway Performance Shaping Factors (R-PSFs) taxonomy and the twelve most important factors that affect the performance of railway operators. Based on these factors, this chapter introduces a new approach, referred to as the Human Performance Railway Operational Index (HuPeROI), which can be used to: (i) assess the impact of each of the R-PSFs on human performance and (ii) to estimate the relative likelihood of human error for several operational scenarios.

The chapter starts by presenting the background to the HuPeROI. This is followed by justifying the significance of the Analytic Network Process (ANP) in the development of HuPeROI. The ANP methodology is then detailed, and the calculation of the HuPeROI demonstrated. Finally, the implementation of the HuPeROI for a scenario of significant interest to the railway industry is presented, and discussed in detail.

## 7.1 The HuPeROI background

The previous chapters described the substantial role and significance of the R-PSFs on the performance of railway operators. It has also been shown that the factors do not equally affect human performance. In addition, as there is no complete taxonomy tailored to the railway industry, it has not been possible to date to quantify the influence of the contributing factors on human performance. Therefore, this thesis proposes a framework to quantify the influence of each of the R-PSFs on human performance. In contrast to existing analytical approaches (see chapters 4 and 5), this new framework accounts for interdependencies amongst factors.

HuPeROI identifies two types of dependencies, the direct and indirect. The former describes the impact that a particular R-PSF has on human performance without taking into account the presence of other R-PSFs. The indirect dependencies represent the gradual influence of a particular R-PSF on human performance due to its relationship with other R-PSFs. Due to the lack of empirical data, subject matter experts (SMEs) are employed to assess the influence of the R-PSFs on human performance. Based on their assessments, a dataset is

generated to support researchers and railway stakeholders to gain a better understanding of how the R-PSFs influence railway operators. In addition, the approach supports researchers to estimate the relative likelihoods of different human operational errors. This should help to direct resources more efficiently to the development of solutions that enhance the safety of operations by mitigating the influence of the contributing factors on human performance.

HuPeROI is based on the integration of two techniques, the Analytic Network Process (Saaty and Vargas, 2006) and the Success Likelihood Index Methodology (Embrey et al., 1984). The former is one of the most well-known and widely used multi-criteria decision making (MCDM) techniques. It is used to assess the contribution of each R-PSF to operators' performance. The latter is a widely cited and broadly used HPA technique. It is applied to rate the quality of each R-PSF for different operational actions. In turn, it used to determine the relative likelihoods of the different human errors for those actions, as function of the R-PSFs. The following section explains why one of the MCDM methodologies is used to assess the R-PSFs. In addition, justifies why the ANP methodology is the most appropriate to be applied to this research. Finally, although the justification for choosing the SLIM methodology was provided in chapter 4, additional explanation is given in the section 7.4.

## 7.2 An overview of the multi-criteria decision making methods

Multiple-criteria decision making (MCDM) analysis deals with problems where decision makers are asked to choose amongst a set of alternatives on the basis of an evaluation according to several criteria. Belton and Stewart (2002, p.2) define MCDM as "*a term to describe a collection of formal approaches which seek to take explicit account of multiple criteria in helping individuals or groups to explore decisions that matter*". Thus, within the context of this research the decision makers, i.e. the SMEs, are asked to indicate the weight of each of the R-PSFs on human performance.

The MCDM approaches were initially developed for the business sector. However, they have subsequently been widely applied to several domains including engineering, energy, air traffic management, transportation, and natural resources management (Triantaphyllou, 2000). Different classifications of the methodologies have been used. Triantaphyllou (2000), for example, distinguishes them according to: (i) the type of data, and (ii) the number of decision makers involved in the process. The first classification includes the deterministic, stochastic and fuzzy methodologies, whilst the second divides MCDM into single and group decision makers. However, recent literature, e.g. (Mendoza and Martins, 2006, Pirdashti et

al., 2011), shows that one of the most precise and well-defined categorisations for the MCDM approaches was given by Hwang and Yoon (1981). This classification suggests two main concepts of methodologies based on the different problems settings. One works with an infinite set of possible continuous alternative[13] solutions, whilst the other with a finite set of discrete alternatives (Xu and Yang, 2001, Buchholz et al., 2009). The first concept is used to identify an optimal solution, which satisfies the decision-making constraints and preferences priorities. Its relevant techniques referred to as multi-objective decision making (MODM) (Triantaphyllou, 2000). The second concept supports the decision maker to understand the complex judgments assists them towards an acceptable solution (Roy and Bouyssou, 1986, as cited in Pirdashti et al., 2011). These problems represent selection and assessment questions with the related techniques referred to as multi-attribute decision making (MADM).

Based on Hwang and Yoon (1981), Belton and Stewart (2002) later classified the MCDM approaches into: (i) value measurement, (ii) goal, aspiration or reference level and (iii) outranking models (Figure 7-1). This categorisation has been widely adopted and cited in the literature, e.g. (Mendoza and Martins, 2006, Løken, 2007, Buchholz et al., 2009, Gervásio and Simões da Silva, 2012). It is also considered appropriate to justify why the MCDM approach is chosen for this research. Thus, the main features of each model are described below.

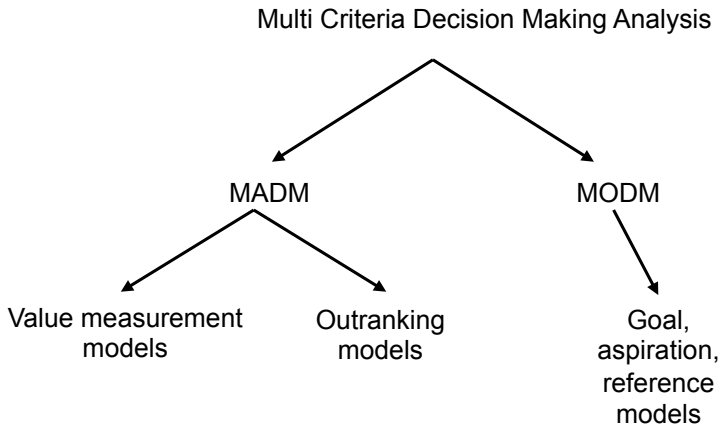



Figure 7-1 Classification of the MCDM methodologies (based on Belton and Stewart, 2002)

[13] Alternatives represent the different choices of action available to the decision maker. They are characterised by a set of attributes (or decision criteria) based on which decision makers evaluate or assess them Triantaphyllou (2000).

### 7.2.1 Value measurement models

The value measurement models use "*numerical scores that are constructed in order to represent the degree to which one decision option may be preferred to another* " (Belton and Stewart, 2002, p.9). The scores (values, V) show an order of preference for the alternatives according to a list of criteria. That is, alternative A is preferred to alternative B (A>B), if V(A)>V(B). Moreover, weights (w) are assigned to the criteria, representing their partial contribution to the overall score. They are based on the decision makers perception of the importance of each criterion (Løken, 2007).

The main properties of these models are that they: (i) provide clear guidance on how the preference models should be built; (ii) support decision makers to gain better understanding of the values they appoint to each alternative and to explain their decisions when required; and (iii) require comparisons between and within criteria and encourage explicit or implicit trade-offs (Mendoza and Martins, 2006, Pirdashti et al., 2011). Typical examples of value measurement models are the Multi-Attribute Utility Theory (MAUT) (Keeney and Raiffa, 1976) and Analytic Hierarchy Process (Saaty, 1980).

### 7.2.2 Goal, aspiration or reference level models

The goal, aspiration or reference level models describe cases where ''*desirable or satisfactory levels of achievement are established for each criterion and options which are closest to achieving these desirable goals or aspirations are sought''* (Belton and Stewart, 2002, p.9). These models are recommended when a theoretically infinite number of continuous alternatives are defined by a set of constraints on a vector of decision variables (Mendoza and Martins, 2006, p.2). They are goal programming (GP) methods where ''*a mathematical programming algorithm is used to approach these goals as closely as possible*" (Belton and Stewart, 2002, p.105) and are used when it is cumbersome for the decision makers to indicate tradeoffs or importance weights.

### 7.2.3 Outranking models

The outranking models are applied to the cases where "*alternative courses of action are pairwise compared, initially in terms of each criterion, in order to identify the extent to which a preference for one over the other can be asserted*" (Belton and Stewart, 2002, p.9). Similar to value measurement, they deal with evaluation problems and have a finite number of alternative solutions. Outranking models aim to identify which of the pairwise compared alternatives is preferred with regard to each criterion, by aggregating decision makers

preferences (Løken, 2007, Gervásio and Simões da Silva, 2012). An alternative A outranks alternative B when all criteria have been taken into account and there is enough evidence to justify that A is at least equally good to B.

However, Greening and Bernow suggest (2004) that outranking methods should not be used for the actual selection of alternatives. However, they should better be applied to the initial screening process, e.g. to categorise alternatives into acceptable or unacceptable. In this case after the screening process another method should be used, i.e. from the value measurement models, to get the full ranking amongst alternatives. The ELECTRE (Benayoun et al., 1966) and PROMETHEE (Brans et al., 1986) are the most frequently used amongst the outranking models.

### 7.2.4 Characteristics of the appropriate MCDM approach

As stated in Section 7.1, the HuPeROI is established on a finite and determined set of R-PSFs. This immediately rules out the use of the "*goal, aspiration or reference level*" models as they address problems with an infinite number of continuous alternatives. Subsequently, the choice of the appropriate MCDM approach is between the MADM categories. However, the requirement for the consideration of dependencies amongst the R-PSFs almost eliminates the choices. In particular, Öztürk (2006), Taslicali and Ercan (2006) and more recently Lee (2010), indicate that only the Analytic Network Process (ANP) introduced by Saaty (1996) provides the mathematical theory that makes it possible to deal systematically with all kinds of dependence. Furthermore, De Ambroggi and Trucco (2011) and De Ambroggi (2010) recently introduced an approach, based on the ANP methodology, to model and assess dependent PSFs that affect the performance of air traffic controllers. This provides additional confidence that the ANP is the most appropriate approach to be used for the development of the HuPeROI.

## 7.3 The Analytic Network Process

The Analytic Network Process (ANP) (Saaty, 1996) supports decision making and allows researchers to systematically study all types of dependencies within the elements of a system. It has been used in several problems that involve the interaction and dependence of higher-level elements on lower-level elements (Saaty and Vargas, 2006). These include the decision of research and development projects (Meade and Presley, 2002), selection of product planning alternatives (Karsak et al., 2003), evaluation of the environmental impact of

various project alternatives (Chen et al., 2005), and quantification of the contributing factors that affect the performance of air traffic controllers (De Ambroggi and Trucco, 2011).

The ANP is a theory of relative measurement used to obtain priority scales from individual judgements that represent relative influence of factors which interact with respect to clearly established control criteria (Saaty, 2004). The methodology is developed on the concept of pairwise comparisons. In addition, it introduces the notion of supermatrix, a square matrix whose elements are themselves matrices of column priorities. The supermatrix represents the result of dependence within and between the elements of the system. In this research, the ANP is used to define the contribution (priorities) of each one of the identified R-PSFs *lite* on human performance. The underpinning theory and the steps for the implementation of the ANP are described in the next subsections.

### 7.3.1 The ANP framework - Developing the structure of the decision model

To structure the ANP model, first the objective of the decision process is defined (Lombardi et al., 2007). This objective is then decomposed into elements which, based on their common characteristics, are grouped into clusters (components). To achieve the objective of the decision process, the established elements and clusters are assessed (weighted) with respect to a control criterion. Within the context of this research, the 12 identified factors represent the elements and the six categories the clusters (the environmental factors are excluded), whilst human performance is the control criterion. Hence, to identify how each of the R-PSFs influences the railway operators, i.e. the goal of the decision process, the factors are judged and compared themselves with respect to how they affect human performance.

The generic structure of an ANP model is developed on a network of clusters, their constituent elements and interdependencies, as shown in Figure 7-2. In general, for a system of *N* clusters, with *n* elements that interact, there are three types of clusters: (i) the *source*, which are considered as origins of influence, hence arrows originate them only, (ii) the *sink*, which are destinations of influence, hence arrows can only enter to them, and (iii) the *transient*, whose arrows can both enter and leave from (Saaty and Vargas, 2006).

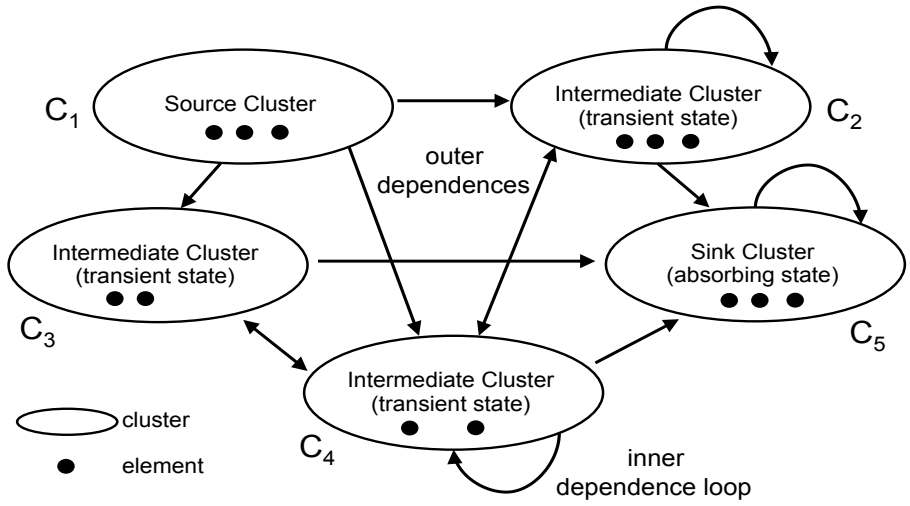



Figure 7-2 Types of components in the network scheme (Saaty and Vargas, 2006)

Once classified, the dependencies between clusters are identified also. That is, when elements of one cluster are either connected with other elements within the same cluster (*inner* dependence) or with elements of another cluster (*outer* dependence) (Saaty and Vargas, 2006). For example, in the network scheme shown in the Figure 7-2, the clusters $C_2$, $C_4$ and $C_5$ indicate loops that connect them to themselves. These loops demonstrate inner dependencies, whilst all the remaining connections, e.g. from $C_1$ to $C_3$, represent outer dependencies**.** Moreover, the clusters $C_3$ and $C_4$ illustrate a mutual dependency, which is described as cycle of the two clusters. In this research, the inner and outer dependencies are defined based on the Table 6-5 and Figure 6-4 in the Section 6.3.3. Finally, a cluster (or element) is influenced by one or more other clusters (or elements) is called a *parent*, while those (or elements) that influence the parent are known as *children*. Pairwise comparisons are conducted based on the relationships between the parent and children clusters (or elements) of the system.

### 7.3.2 The ANP pairwise comparisons and relative weight estimation

The pairwise comparisons are conducted for those elements and clusters that indicate dependencies and share a common parent with respect to the control criterion (Saaty and Vargas, 2006). Such comparisons are performed by answering the question: ”*given a control criterion, a parent cluster (element) of the network and a pair of clusters (or elements) that both are linked to the parent cluster (element), how much more does a given member of the pair influence that parent cluster (element) with respect to the control criterion than the other member?”* (Saaty and Vargas, 2006). To indicate such a relationship, the *Fundamental Scale* in Table 7-1 is used (Saaty, 1980).

Table 7-1 The Saaty's fundamental scale

| Intensity of importance | Definition | Explanation |
|---|---|---|
| 1 | Equal importance | Two factors contribute equally to the objective |
| 2 | Weak | Compromise between values 1 & 3 |
| 3 | Moderate importance | Experience and judgment slightly favour one factor over another |
| 4 | Moderate plus | Compromise between values 3 & 5 |
| 5 | Strong importance | Experience and judgment strongly favour one factor over another |
| 6 | Strong plus | Compromise between values 5 & 7 |
| 7 | Very strong or demonstrated importance | A factor is favoured very strongly over another; its dominance demonstrated in practice |
| 8 | Very very strong | Compromise between values 7 & 9 |
| 9 | Extreme importance | The evidence favouring one factor over another is of highest possible order of affirmation |

Judgments are first given verbally and then each is assigned a corresponding number. The numerical assessments established at each level of the network are used to construct the pairwise matrices, as illustrated in Table 7-2. Such matrices display the strength of dominance of an element in the column on the left over an element in the row on top. For example, in Table 7-2 it is shown that *Element 1* influences the *Parent Element* **3 times more** than *Element 3*.

If the element on the left is less important than the one at the top of the matrix, then the reciprocal value in the corresponding position in the matrix is entered. The approach takes as input the pairwise comparisons and use them to produce the relative weights of the elements as output, using the *eigenvalue* approach (Lombardi et al., 2007). For this, the following eigenvalue formulation is used:

$$Aw = \lambda_{max} w,\ e^T w = 1 \qquad (7\text{-}1)$$

in order to obtain the desired ratio-scale priority vector (or weights) *w* of *n* elements, where $A=a_{ij}$ is the positive reciprocal pairwise comparison matrix, $\lambda_{max}$ is the principal eigenvalue of the matrix A and $e^T$ = 1,1,..,1 is a unit row (or summation) vector (Promentilla et al., 2008).This output is the *priority vector* and corresponds to the main eigenvector of the comparison matrix. The calculation of the priority vectors is described in detail in Section 7.5.7.

Table 7-2 Pairwise comparison matrix for Elements 1-3

| Parent element | Element 1 | Element 2 | Element 3 |
|---|---|---|---|
| Element 1 | 1 | **1** | **3** |
| Element 2 | 1 | 1 | **6** |
| Element 3 | 1/3 | 1/6 | 1 |

However, associated with these judgments is a level of inconsistency. For example, it is assumed that for the comparison of the three elements in Table 7-2, one compared *Element 1* with *Elements 2* and *3*, with the derived judgements to be 1 and 3 respectively (bold red font). It is clear that $if\ x = y\ and\ x = 3z\ then\ y = 3z$. Thus, if the assessor was consistent, *Element 2* when compared to *Element 3* would be assigned the value 3 instead of 6. Inconsistent matrices may significantly affect the reliability of the derived results. Hence, the acceptable level of their inconsistency must be determined. The consistency of a matrix is expressed by its consistency index, which is derived from Equation 7-2:

$$CI = \frac{\lambda_{max} - n}{n-1} \tag{7-2}$$

where, *λmax* is the principal eignevalue of the pairwise square matrix and *n* the order of this matrix. Once the consistency index is calculated, the result is compared with an average consistency index generated by Saaty (1987) for a very large sample of randomly filled matrices. This is called the Random Index (RI) and its values for n≤10 are given in Table 7-3.

Table 7-3 The Random Index

| The Random Index | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| RI | 0 | 0 | 0.58 | 0.90 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 | 1.49 |

The ratio of the CI to the average RI for the same order matrix is called the *consistency ratio* (CR). According to Saaty (2008) a matrix is considered consistent when CR≤ 0.10 for any matrix larger than 5x5. If the CR is larger than 0.1, then the following three steps should be followed: (i) define the most inconsistent (larger) assessment in the matrix, (ii) determine the range of values to which that assessment can be changed to improve inconsistency, (iii) ask assessor to change their assessment to a plausible value in that range (Saaty, 2008). The matrix is rejected if remains inconsistent.

### 7.3.3 The ANP supermatrices

To estimate the final priorities (importance) of the elements against the control criterion, the initial ANP supermatrix is created. This supermatrix is developed by entering in its columns the priorities derived from the set of pairwise comparisons. The elements within this matrix allow for a resolution of the interdependencies that exist between the elements of the system (Bottero et al., 2007). It is a portioned matrix where each of the sub-matrices shows the relationships between and within the clusters and elements as defined by the decision making model. The generic form of a supermatrix is illustrated in Table 7-4. The $C_N$ refers to the $N^{th}$ cluster, while $e_{Nn}$ indicates the $n^{th}$ element in the $N^{th}$ cluster. In addition, the $W_{ij}$ block sub-matrix comprises the priority vectors (w) of the influence of the elements in the $i^{th}$ cluster with respect to the $j^{th}$ cluster. Finally, when no influence exists between elements, the corresponding entries in the supermatrix become zero (De Ambroggi and Trucco, 2011). The initial supermatrix is commonly referred to as *unweighted*.

Table 7-4 The generic structure of an *unweighted* supermatrix

| | | $C_1$ | | | | $C_2$ | | | | … | $C_N$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $e_{11}$ | $e_{12}$ | …. | $e_{1n}$ | $e_{21}$ | $e_{22}$ | …. | $e_{2n}$ | | $e_{N1}$ | $e_{N2}$ | ⋯ | $e_{Nn}$ |
| $C_1$ | $e_{11}$<br>$e_{12}$<br><br>$e_{1n}$ | $W_{11}$ | | | | $W_{12}$ | | | | ….. | $W_{1N}$ | | | |
| $C_2$ | $e_{21}$<br>$e_{22}$<br><br>$e_{2n}$ | $W_{21}$ | | | | $W_{22}$ | | | | ….. | $W_{2N}$ | | | |
| | | | | | | | | | | | | | | |
| $C_N$ | $e_{N1}$<br>$e_{N2}$<br><br>$e_{Nn}$ | $W_{N1}$ | | | | $W_{N2}$ | | | | ….. | $W_{NN}$ | | | |

The priority entries in the unweighted supermatrix do not indicate the priority of the elements in the entire set of clusters. This is explained by the fact that each of the clusters should have a highest ranked element and by definition not all elements can be first in the system (Saaty and Vargas, 2006). In addition, in many cases the clusters in a network may not be equally important (Saaty, 2003), which in turn affects the importance of their elements. Thus, the clusters themselves are compared to establish their relative influence with respect to the control criterion.

A cluster has an impact on another cluster when at least one element in the source cluster is linked to one or more elements in the target cluster (Saaty, 2008). The children clusters connected to a parent cluster are pairwise compared resulting in a new matrix, known as *cluster* matrix. The priorities of the cluster matrix are then multiplied with those of the corresponding elements in the unweighted supermatrix in order to weight the block of matrices of the eigenvector columns under that cluster (Saaty, 2004). The first entry of the vector is multiplied by all the elements in the first block of that column, the second by all the elements in the second block of the column and so forth. This process results in a stochastic supermatrix i.e. all its columns are equal to one which is called *weighted*. Based on the weighted matrix, the desired priorities of elements are estimated.

### 7.3.4 The ANP final priorities

The entries of the weighted supermatrix illustrate only the direct influence of an element on any other element. However, due to the dependencies, an element A can influence (or be influenced by) a second element B indirectly through its influence on a third element C which also affects the second element B. In this case all indirect influences of pairs of elements are derived by squaring the weighted supermatrix (Saaty and Vargas, 2006, Saaty, 2008). Similarly, the influence of one element on a second may also be established due to a third element that influences a fourth element, which in turn affects the second element. All such influences are derived from the cubic power of the weighted supermatrix. Following the same rationale, the weighted supermatrix is raised to limiting powers, where each power of the matrix represents all transitivities of an order that is equal to that power (Saaty, 2008). Since the weighted supermatrix is stochastic, the process of raising to powers stops when all the columns of the new matrix are identical (Saaty and Vargas, 2006). The values of this new matrix, referred to as the *limit* supermatrix, constitute the desired final priority vectors, known as limit priorities. These priorities represent the relative measurement of the elements after capturing the transmission of influence along all possible paths of the weighted supermatrix (Saaty and Vargas, 2006). The values of the *limit* supermatrix in this research represent the weights of the R-PSFs *lite*.

## 7.4 The integration of ANP and SLIM methodologies – the HuPeROI calculation

As mentioned in Section 7.1, in addition to the assessment of the influence of the R-PSFs on human performance, the HuPeROI framework is designed also to enable researchers to

estimate the relative likelihoods of actions caused by human error for a variety of operational scenarios. For this reason, the values derived from the ANP approach are integrated with the SLIM methodology (Embrey, 1983).

The main features of SLIM were addressed in the Section 4.10.2.2. In summary, the methodology is a decision-analytic approach that uses expert judgement to assess human performance. The basic rationale of the approach lies in the assumption that the occurrence of an error during the execution of a task in a particular situation depends on the combined effects of a set of contributing factors. Thus, it is assumed that a panel of experts are able to indicate the relative importance of each PSF on the performance of the operator for that task. It is also assumed that the same panel can rate those PSFs with respect to how good or bad they are in the task under consideration. Within this context, “good” or “bad” refers to whether the PSFs will either enhance or degrade the ability of the operator to perform the task respectively (Bell and Holroyd, 2009). Subsequently, a success likelihood index (SLI) for the particular task is calculated, as a function of the relevant set of factors. This index represents the overall belief of the panel concerning the effect of each PSF on the likelihood of success for the task under consideration. One of the main attributes of the SLIM approach is that, unlike other HPA techniques, it allows researchers to decide which PSFs are the most important and should be used in every case. This feature is of great importance for this research, as it enables the implementation of the R-PSFs for the computation of the success indices known as HuPeROIs. The HuPeROI is derived from the equation:

$$HuPeROI_j = \sum_{i=1}^{n} w_i \times r_{ij} \quad (7\text{-}3)$$

where, $w_i$ is the weighting for the ith R-PSF and $r_{ij}$ the rating of task j on the ith R-PSF. As shown in the previous section, the weighted values are derived from the implementation of the ANP.

For the ratings the SMEs are asked to indicate to what extent the identified factors assist the operators to accomplish a certain action. For each of the actions, the SMEs rated each of the R-PSFs on a scale of 0 to 100 (Gertman and Blackman, 1993). The closer a rating is to zero for a factor, the more suboptimal the condition for the action under examination. However, for factors such as workload, since it can be claimed that both extreme workload and no workload may have an impact on the performance of the operators, it is assumed that a zero workload value indicates normal conditions. The Table 7-15 in Section 7.5.8 illustrates the summary of the R-PSF scales used to rate each of the actions in the question. For the R-PSFs ratings, in a manner similar to the ANP approach, the following question was distributed to the SMEs: “*considering that each R-PSFs affects the ability of the*

*operator to conduct a certain action, please rate how good or bad do you think that each R-PSFs is with respect to the action required*", where actions refer to potential errors. Having obtained the R-PSFs weightings and ratings and by implementing Equation 7-3, the HuPeROI for each action was then calculated, and the greater the HuPeROI value the greater the probability of success of the task under consideration. The framework for the development of the HuPeROI is illustrated in Figure 7-3.

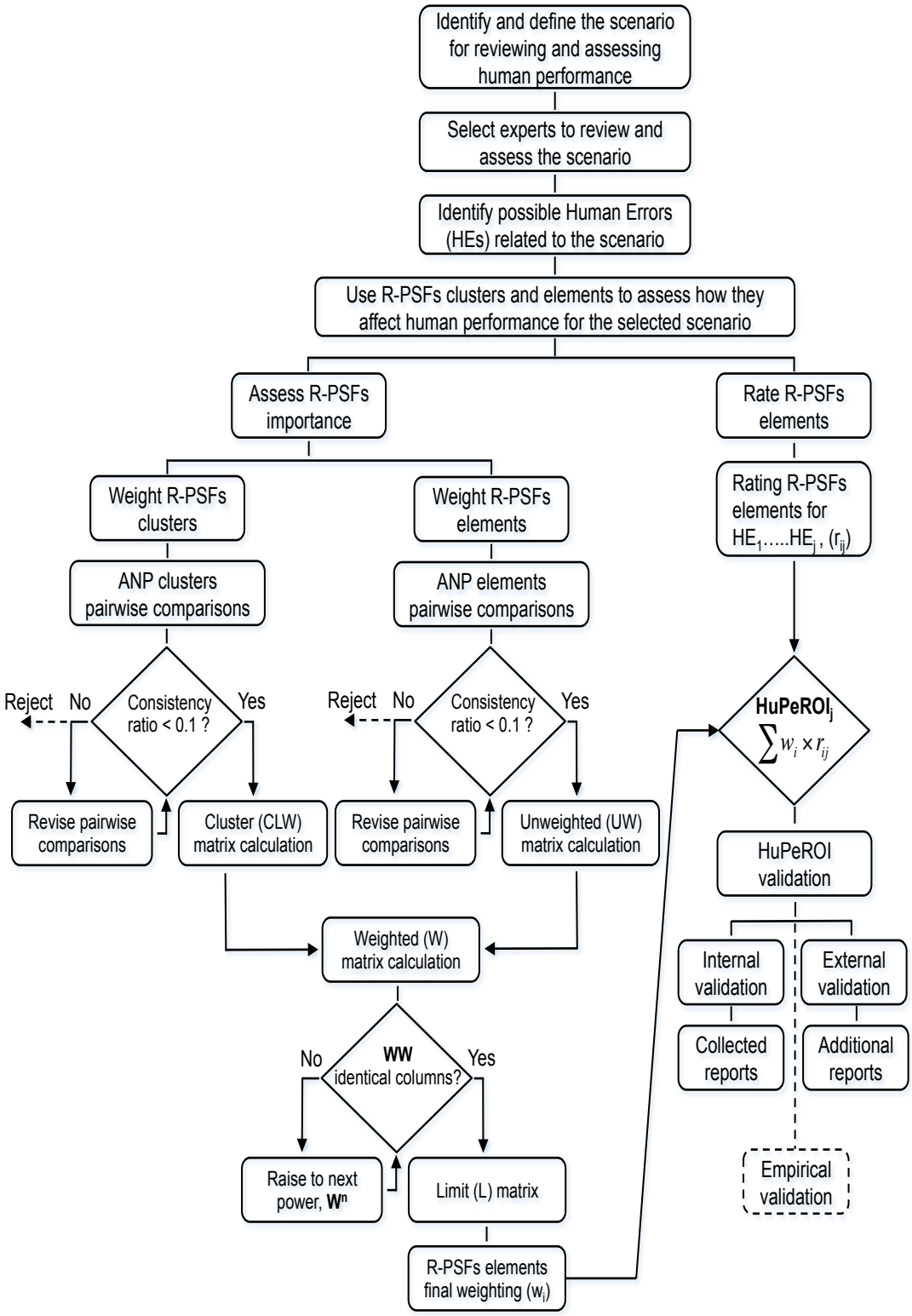



Figure 7-3 The framework to develop and calculate the HuPeROI

As shown in Figure 7-3 the first step in the process is to identify and define the scenario for which human performance should be assessed. This is followed by the selection of experts who review the scenario and assess human performance. The potential human errors for the scenario are determined and the R-PSFs are applied to compute the HuPeROI for each

human error. Once the HuPeROIs are determined, results are internally and externally validated. For the former, the reports related to the scenario under investigation are reviewed and the findings compared with those of the HuPeROI. For external validation, additional reports, which describe the same human errors as the ones to be assessed, should be collected and analysed. The findings of the additional reports are again compared with the HuPeROI. In addition, simulation observations could be also used to validate and consolidate the accuracy of findings.

The integration of the ANP and SLIM methodologies has been cited in the literature in at least two different studies. Park and Lee (2008) first introduced the AHP-SLIM, as a novel approach to estimate the probabilities of the occurrence of car driver related human errors. The AHP-SLIM is developed on the Analytic Hierarchy Process (AHP), a special case of the ANP where elements are considered to be independent. Although the AHP-SLIM is unsuitable for this research, as it does not account for dependencies amongst PSFs, it indicates that the integration of the two approaches can be successfully performed.

The second study, introduced by De Ambroggi (2010) is used to model and assess the factors that affect the performance of air traffic controllers (ATCOs). The approach combines the ANP and SLIM methodologies, where the former is applied to quantify the impact of particular PSFs on the performance of ATCOs. SLIM is then used to determine the success likelihood index of certain human errors. Despite the fact that De Ambroggi (2010) improves Park and Lee's (2008) methodology by taking into account the dependencies amongst the PSFs, the technique has two main limitations. The first is related to the fact that for the assessment of the factors that affect human performance, only the main categories of factors are taken into account, whilst the PSFs elements are disregarded. De Ambroggi justifies the choice to use only the categories of factors due to the large number of elements within the taxonomy used in the study, i.e. the HERA taxonomy. In particular, De Ambroggi argues that if the PSF elements were used, this would lead to a rather large and complex set of relationships between the PSFs, making the approach cumbersome and user-unfriendly. However, by using only the categories of PSFs this approach fails to determine how each of the elements affects human performance, which is essential for organisations to direct resources more efficiently towards the mitigation of those factors. The second limitation of the approach, an issue tackled by Park and Lee, is related to the numerical rating of the PSFs for the tasks under consideration. Instead of asking SMEs to rate the PSFs, De Ambroggi considers three operational scenarios where all or some of the PSFs categories, in a Boolean logic scheme, play a negative role on the task. However, this simplification

provides only limited insights with respect to how the SMEs consider a PSF to be in reality (Kirwan, 1994).

The HuPeROI addresses both limitations. Regarding the weighting of the R-PSFs, it assesses both the categories and elements of the R-PSFs *lite* with respect to their importance on human performance. In rating the R-PSFs, the SMEs indicate how good or bad the R-PSFs are for certain operational tasks. Based on the judgements, the success indices of the tasks are then calculated, with the results indicating which of the tasks is more prone to fail. In the remainder of the thesis, the implementation of the HuPeROI for a scenario of significant interest for the railway operations is presented.

## 7.5 The implementation of the HuPeROI

This section presents a case study developed for the purposes of this research to demonstrate the application of HuPeROI to the railway operations. In particular, the selected scenario and the sampling of the SMEs are introduced, and the design of the survey and the relevant calculations are demonstrated. Finally, this section it addresses also the HuPeROI's limitations and potential sources of error.

### 7.5.1 Signals passed at danger (SPADs) scenario

The first step requires the identification of a scenario and for this a signal passed at danger (SPAD) scenario was selected. SPADs occur when trains pass a signal displaying a stop aspect without authority (Office of Rail Regulation, 2011c) and whilst they may occur due to a failure in infrastructure, e.g. signal failure, unexpected events, e.g. failure of train in block ahead, or a train out of control, they are primarily related to operational errors caused by train drivers (Rail Accident Investigation Board, 2003, Independent Transport Safety Regulator, 2011). Thus, the main human error of interest in this situation is "failure of the operator to stop the train before passes a stop signal".

Chapter 3 highlighted the Ladbroke Grove multi-fatal railway accident, probably the most cited SPAD in recent years. The accident, which caused 31 fatalities and 523 injuries, occurred because the driver of a passenger train passed a red signal without authority. This action caused a near head-on collision with a train heading to opposite direction (Cullen, 2001). The inquiry concluded that the accident occurred due to a combination of contributing factors including driver inexperience, training limitations, complexity of the layout and location of the signal (Health and Safety Executive, 2000). It was stated clearly that if the

rolling stock and infrastructure were equipped with more sophisticated protection systems, e.g. TPWS, the accident would have been prevented.

With the continuous evolution of technology and the implementation of more advanced railway protection systems, such as the TPWS and ATP, it could be argued that SPADs would not be a priority for the railway industry. However, such an argument is flawed because SPADs today are a significant precursor of risk due to their ability to still cause fatal accidents especially in the cases where advanced protection systems are not installed. Typical examples of recent SPAD accidents, as mentioned in Chapter 3, are: (i) the multi-fatal trains collision in California in 2008 where a train driver missed a red signal because he was using inappropriately his mobile phone (National Transportation Safety Board, 2010) and (ii) the multi-fatal trains collision in Hordrof of Germany in 2011 (EUB, 2011). In addition, one could also refer to the fatal accident in Granges-pres-Marnand of Switzerland in July 2013, where evidence indicate that was caused due to a SPAD (Swissinfo, 2013).

Furthermore, the literature identifies the significant role of SPADs in the safety of railway operations (Wilson et al., 2005, Wilson and Norris, 2005, Wilson and Norris, 2006, Dhillon, 2007, Wilson et al., 2007a, Wilson et al., 2007b), where numerous studies have been conducted over the years (Rail Accident Investigation Board, 2003, Pasquini et al., 2004, Rail Safety and Standards Board, 2004a, Hamilton and Clarke, 2005, Blanchard and Lowe, 2009, Independent Transport Safety Regulator, 2011, Wright et al., n.d.). In addition, worldwide statistics, e.g. (European Railway Agency, 2012c), as well as the review of accident and incident reports presented in the previous chapter, indicate SPADs as the main cause of railway incidents and the primary cause of most of the severe train collisions.

Finally, five European train operating companies (TOCs) and an infrastructure manager (IM) were also asked to indicate the type of error that involves human performance and considered of great interest to their organisations. Three scenarios were distributed, including occurrences of the same severity caused by SPAD, excessive speed and wrong routing. Respondents suggested that the SPAD scenario is the most critical, for the following reasons:

(i) the consequences of SPADs nowadays are usually not severe, i.e. due to protection systems; hence, although train drivers claim to acknowledge the importance of SPADs prevention, in reality they do not consider them as serious occurrences;
(ii) subsequently SPADs constitute a large proportion of railway incidents, which may lead to catastrophic events;

(iii) although they usually result in minor consequences, SPADs may cause: (a) significant disruptions on the network, especially in the cases of metro services, (b) inconvenience to passengers and (c) financial losses for the organisations as each SPAD requires official investigation;

(iv) TOCs claim that, cope more efficiently with incidents due to excessive speed or routing; however, the respondents did not provide any justification with respect to this statement.

Taking into account the literature, worldwide statistics, review of investigation reports and feedback from the railway stakeholders, a SPAD scenario was created for the purposes of this research (see Appendix XI). The scenario describes a SPAD incident within a specific operational context, commonly seen in current railway operations.

### 7.5.2 Sampling of experts

The next step in implementing the HuPeROI, is to determine the group of experts that weight and rate the influence of the R-PSFs on human performance. The population of experts should primarily include train drivers, who are likely to be involved, or may have already been involved, in a SPAD. However, if only train drivers are selected as SMEs, the possibility of biased results, which may indicate specific factors, e.g. fatigue or workload, as the most significant contributors to human performance could not be disregarded. Thus, a wider population was targeted including operations managers and human factors (HFs) experts. The main reason for this is to indicate whether the collected responses are related to a particular type of SMEs and to identify if any differences are established between frontline and regional management staff. The responses of HFs experts are complementary and used to highlight similarities or differences between operational and non-operational experts.

To obtain results that represent the majority of railway operations, the targeted population should be formed of randomly selected SMEs with a range of operational experience, i.e. years in service, position in the organization, from various TOCs worldwide. The operating companies should also reflect different types of service, levels of traffic, network characteristics and protection systems. However, collecting representative data (responses) from many countries would pose a significant challenge due to time and resource constraints. Moreover, railways operate primarily on national regulations and standards, as well as on signalling and controlling systems. Thus, it is likely that responses collected from several countries to be biased given national differences. In addition, responses from several countries may also be biased due to cultural, social and financial differences. Given

these limitations, it was decided that the target population would consist of SMEs operating in the same country. Furthermore, this country was the UK and its network was selected for three reasons: (i) the broad network of railway services, including freight, intercity, regional and metro passenger services; (ii) despite the fact that it is considered one the safest railway networks in Europe, a substantial number of SPADs is reported every year on both its main line (Rail Safety and Standards Board, 2013) and underground operations, and (iii) direct accessibility to train operating companies (TOCs).

The UK railway network comprises 28[14] passenger and seven freight train operating companies (National Rail Enquiries, 2012, Network Rail, 2012a, ATOC, 2013). Contact was established with 22 companies (6 freight and 16 passenger) with the aim of including SMEs from all types of services. Six companies expressed an interest in participating in the research, satisfying the criterion of incorporating all types of services in the target population. Eventually, three TOCs agreed to contribute to the survey, while the remainder were unable to do so due to time and resources constraints. The three participating TOCs are: the First Hull Trains (intercity services), the First ScotRail (regional and commuter services) and the Piccadilly Line of the London Underground (underground services). This population represents 10.7% of U.K. passenger train companies. The First ScotRail reflects 95% of Scottish passenger railway services and the Piccadilly Line is the longest and fourth busiest manually operated line of London Underground. The population is completed with the HFs experts from the RSSB.

Having described the target population, it is important to define the size of the sample. The number of potential SMEs from the population is approximately 800. However, due to resource and scheduling constraints each TOC could support the study with up to 20 volunteers, while the HFs groups could provide maximum 5. Salant and Dillman (1994) indicate that for a population with 800 members, a sample size of at least 57 responses is required to obtain results with an error of no more than ± 10% at 95% level of confidence. Thus, the total sample of 65 participants satisfies Salant and Dillman (1994), while it is considerably larger compared to samples generally used in similar studies, e.g. Khan et al., (2006), De Ambroggi and Trucco (2011). However, as the target population is not homogeneous, i.e. the types of employees are not equally represented in the organisations, a random selection of experts could potentially lead to a non-representative sample (Hinkle et al., 1979). Hence, based on the proportion of train drivers and regional management staff

[14] Including the 24 train companies of the Association of Train Operating Companies and the London Underground Limited, Heathrow Connect, Gatwick Express and Stansted Express.

in the population of interest, the aim was to achieve the following allocation: 70% train drivers and 30% regional; that is 42 train drivers and 18 regional management staff. Finally, acknowledging that due to limited resources obtaining statistically representative sample at large was not achievable, representative diversity was sought, as shown in Section 8.1.

### 7.5.3 Survey methodology – Responses collection

Surveys are widely acknowledged as an accurate methodology to measure attitudes, beliefs or facts. In particular, one of the most common uses of surveys, is to quantify the behavioural past experiences of individuals (Weisberg et al., 1996). For the purposes of this research, a questionnaire survey was conducted to collect information on the influence of particular R-PSFs on the performance of railway operators.

To conduct the survey, two approaches could be used: either to gather information from face-to-face interviews with the SMEs, or remotely by self-completion. The first approach is more time consuming, requiring visits to railway organisations and direct access to the SMEs. However, it is considered more reliable since it supports the clarification of any ambiguous issues in person, either prior to or during the interview (Subotic, 2007). Furthermore, since the schedule of interviews can be set up in advance, this approach facilitates representative sampling within railway organisations. However, apart from the time limitations, a main weakness of the in-person interviews is the financial requirements for the visits to railway organisations.

In contrast, a remote survey enables the online distribution of questionnaires. Compared to in person interviews, this approach saves time and supports the increase of the sample, since the physically presence of the moderator is not required. However, it has three main drawbacks. Firstly, the remote surveys are strongly dependent on the response rate. In particular, Subotic (2007) states that for the collection of 100 samples, between 200 to 1000 questionnaires should be distributed, which renders the approach rather cumbersome. Secondly, it relies entirely on the willingness and motivation of respondents to participate in the survey. Finally, as the physical presence of the moderator is not necessary, the quality of answers cannot be monitored, and participants may either misinterpret some of the questions or need additional information on the subject under investigation.

After careful consideration of the advantages and disadvantages of both approaches, the face-to-face strategy was selected to facilitate the survey, for two reasons. Firstly, the defined sample size and location of respondents facilitated a visit to all the four railway organisations that agreed to participate in the survey. Secondly, it is vital for the accuracy of

the results to ensure that the respondents understand the context of the survey and respond accordingly. This would be enhanced by the presence of the interviewer.

### 7.5.4 Questionnaire design

The questionnaire was designed in a way to account for three certain criteria:

(i) clarity and representativeness to engage respondents in the usefulness of the study;
(ii) questions clarity, to prevent any ambiguities between respondents;
(iii) R-PSFs clear and precise definitions and examples, to ensure that all respondents have a common understanding of the components they are asked to evaluate.

The questionnaire consists of four parts, as shown in the Appendix XI. The first part includes questions that describe the profile of SMEs, i.e. years of experience, duties, and age. In addition, this part includes questions with respect to the previous SPAD experiences of the SMEs. The second part contains the questions for the assessment (weighting) of the R-PSFs categories and elements with respect to the defined scenario. The third part addresses the ratings of the R-PSFs with respect to the three typical actions (human errors) that lead to SPADs. Such actions were defined based on literature, e.g. (Pasquini et al., 2004, Dhillon, 2007, Wright et al., n.d.), the analysis of the investigation reports, whilst they were also verified by three SMEs. Finally, in the last part of the questionnaire the SMEs were asked to indicate how external conditions, e.g. weather conditions, may affect their performance.

The questionnaire, i.e. layout, questions, scenario, was validated with a HF expert from the RSSB and two train drivers. Initially, the questions were found confusing, thus the SMEs were asked for clarifications, in particular with respect to the definitions of R-PSFs. Thus, cards (Appendix X) were distributed to the reviewers, which precisely describe the R-PSFs, provide definitions and illustrate relevant examples. Using the cards, the reviewers agreed that it was clear how to answer the questions in terms of weighting and rating the R-PSFs. Moreover, the two train drivers highlighted that the use of cards eliminates the interpretation of the R-PSFs based on personal experiences and perception. In addition, the reviewers confirmed that the scenario was clearly written, well-justified and contained sufficient information for its completion. Finally, they confirmed that the including human errors are those which most likely lead to SPADs, compared to other human errors, e.g. late signal observation. Subsequently, confidence was gained that the questionnaire satisfies the criteria of clarity, usefulness and common understanding amongst respondents.

### 7.5.5 Preliminary survey

A trial preliminary set of interviews, before conducting the full survey, was conducted in order to confirm the clarity of the questionnaire, to certify that the respondents understand the definitions and the given examples and finally to assess the necessary time to complete the questionnaire. For this, one HFs expert and two train drivers were interviewed (different from those who validated the questionnaire). Results indicated the need of minor amendments with respect to the layout of the questionnaire and the way that the respondents should compare the R-PSFs, e.g. circling only one factor per row, while the duration of each session was approximately 55 minutes. The necessary alterations were incorporated in the final questionnaire, which is provided in the Appendix XI.

### 7.5.6 Completion of survey

For the completion of the survey, the author visited the premises of the four railway organisations to distribute the questionnaires and collect the responses of the randomly selected SMEs. Using the face-to-face interviews approach, responses were received first from the First ScotRail SMEs in December 2012. Then, interviews were conducted in February 2013 with the SMEs from FirstHull Trains. Piccadilly Line and HFs SMEs, were interviewed in March of the same year. All interviews were performed on an individual basis, where the average duration of each session was approximately 65 minutes.

Based on sampling, 65 interviews were scheduled. However, due to a technical failure on the Piccadilly Line, three interviews were cancelled, because the SMEs were required to return on duty. In addition, six SMEs refused to answer the questionnaire. Due to time constraints none of the cancellations could be rescheduled. Finally, 56 SMEs took part in the study, and their characteristics are shown in the Section 8.1.

Having described the layout of the survey, the next subsections present the process of answering the questionnaire and how such responses were transformed to data.

### 7.5.7 The assessment (weightings) of the R-PSFs

Starting with the validated network of the R-PSFs dependencies, the influence of each R-PSF element on human performance is assessed. Based on such dependencies, the parent and children elements as well as clusters are firstly identified, as shown in the Table 7-5 and Figure 7-5. In the Table 7-5, in particular, the parent elements are listed in the left hand

column and the symbol (✓) is used to indicate the children elements who have an influence on the parent elements.

Table 7-5 The parent and children R-PSFs elements

| Parent (Y axis) and children (X) elements | T | F | D | P | Ftg | W | C | I | SC | S | P | SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Training | | ✓ | | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Familiarity | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Distraction | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Perception | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fatigue | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Workload | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Communication | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Information | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Safety culture | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Supervision | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| Procedures | ✓ | | | | | | ✓ | ✓ | ✓ | ✓ | | |
| System design | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |

D: Distraction, F: Familiarity, P: Perception, Pr: Procedures, SC: Safety culture, S: Supervision, Ftg: Fatigue, T: Training, SD: System design, W: Workload, C: Communication, I: Information

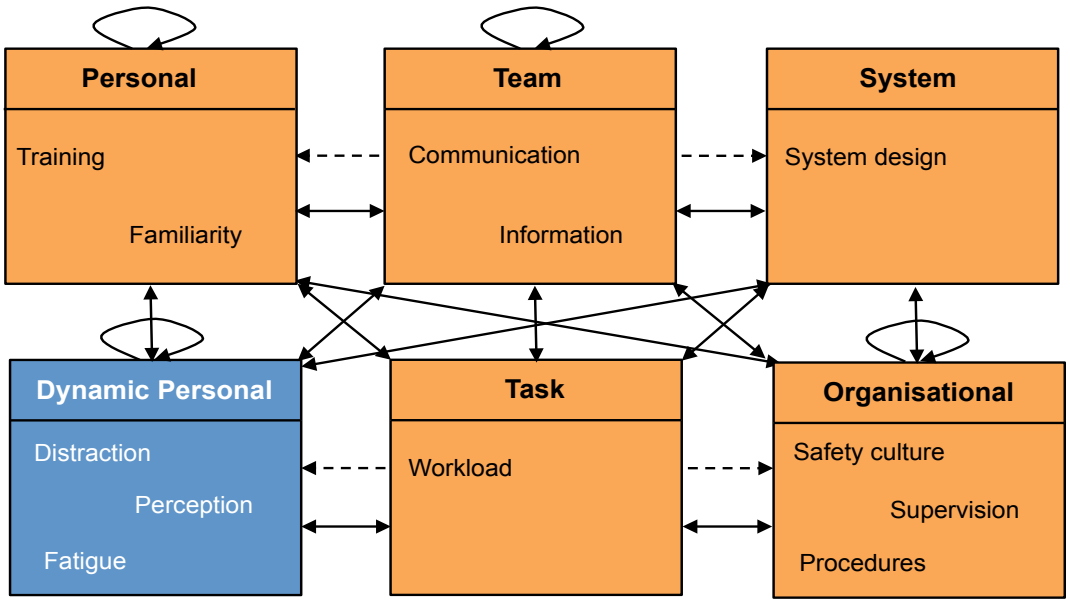



Figure 7-4 The cluster dependencies for the pairwise comparisons

Subsequently, the pairwise comparisons between the dependent R-PSFs are designed and the corresponding, unweighted supermatrix is created. Likewise, the dependent clusters are compared in a pairwise manner, and the cluster matrix is established, while the weighted supermatrix is calculated afterwards. Finally, the limit supermatrix, which includes the final priorities of the R-PSFs is computed. The four steps of the ANP calculations for the SPAD

scenario are presented below, while the responses of a random SME are used to illustrate the process.

#### 7.5.7.1 Elements pairwise comparisons – the unweighted supermatrix

For the pairwise comparisons of the R-PSFs elements, the SMEs were asked to judge all children elements with respect to their parents for their influence on human performance. For this, the following question was distributed: *"considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on the parent factor?*". As a reminder, for each parent element, the compared children elements belong only to same clusters. The influence of elements amongst different clusters is derived indirectly by assessing the clusters themselves.

The questionnaire layout used for the comparisons is shown in the Table 7-6. The following example shows the pairwise comparisons and calculation of priorities with respect to the *communication* parent element. Communication, as indicated in the Table 7-5, is influenced by all the remainder R-PSFs. However, *workload* and *system design* are the only R-PSFs within their clusters, hence their influence on communication will be indirectly derived from the total elements and clusters pairwise comparisons. Therefore, seven comparisons, within three clusters, i.e. dynamic personal, organisational, personal, can be directly performed.

Table 7-6 Pairwise comparisons for the parent element *communication*

| Communication | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Communication |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | **5** | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | **4** | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | **7** | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | **6** | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |

The Table 1-7, in turn, illustrates the pairwise obtained matrix from the SME's responses for the children elements included in the *dynamic personal* factors cluster.

Table 7-7 Pairwise comparisons for elements

| Communication | Distraction | Fatigue | Perception | Priorities 1 | Priorities 2 | Priorities 3 |
|---|---|---|---|---|---|---|
| Distraction | 1 | 1 | 5 | 0.46647 | 0.46633 | 0.46647 |
| Fatigue | 1 | 1 | 4 | 0.43303 | 0.43300 | 0.43303 |
| Perception | 1/5 | 1/4 | 1 | 0.10050 | 0.10067 | 0.10050 |

For the calculation of the priorities, the matrix is first raised to arbitrary large powers, and the elements of each of the rows are summed. Finally, all the rows are normalized by dividing each row sum by the total sum of all rows (Saaty, 2008); the derived results constitute the desired priorities. The process of raising to powers stops when the difference between the components of the obtained priority vector at the $k^{th}$ power and at the $(k + 1)^{th}$ power is less than some very small value. For this example, the priorities were derived from the matrix after having raised to the $4^{th}$ power. Results are indicated in the column *Priorities 1* of Table 7-7. The entries of each priority vector add up to one and are placed in their appropriate location in the unweighted supermatrix, as shown in the next section.

However, the literature identifies two more ways to calculate the priorities. The first in Lee (2010), describes a process of three steps. In the first the values in each column of the pairwise matrix are summed. In the second step the matrix is normalised by dividing the elements in each column by the sum of their respective column. Finally, in the third step the elements in each row of the normalised matrix are summed and this sum is divided by the number of elements in the row. These final numbers give an estimate of the relative priorities of the elements being compared, as shown in the column *Priorities 2* of Table 7-7. Finally, a third approach (Paraskevopoulos, n.d.) estimates the priorities by firstly multiplying the entries in each row of the matrix and deriving then the $n^{th}$ root of each row. In turn, the $n^{th}$ roots are summed and the $\sum n^{th}$ is obtained. The priorities are finally derived by normalising the $n^{th}$ root of each row with the sum of $n^{th}$ roots. As shown in the Table 7-7, the column *Priorities 3* has identical values with Saaty's calculations. Although all approaches result in identical results, the approach of Saaty is used to calculate the priorities.

Having calculated the priorities from the pairwise matrix, the next step is to define whether the matrix, or in other words, the respondent, is consistent. Therefore, the maximum eigenvalue $\lambda_{max}$, the consistency index (CI) and subsequently the consistency ratio (CR) are estimated, as shown in the Table 7-8. To calculate the $\lambda_{max}$, the matrix of judgments is multiplied on the right by the eigenvector (*Priorities 1*), resulting in a new eigenvector.

Table 7-8 R-PSFs priorities calculation

| Communication | Distraction | Fatigue | Perception | Priorities 1 | New eigenvector | λmax |
|---|---|---|---|---|---|---|
| Distraction | 1 | 1 | 5 | 0.46647 | 1.40199 | 3.00554 |
| Fatigue | 1 | 1 | 4 | 0.43303 | 1.30149 | 3.00554 |
| Perception | 1/5 | 1/4 | 1 | 0.10050 | 0.30205 | 3.00554 |

For the first row in the matrix, that is: (1×0.4667)+(1×0.43303)+(5×0.10050)=1.40199. However, the new eigenvector expresses the term Aw in the equation $Aw = \lambda_{max} w$ (Saaty, 2008). The $\lambda_{max}$ is then calculated by dividing each component of the new eigenvector with its corresponding eigenvector (*Priorities 1*). In the case that $\lambda_{max}$ differ between elements, the $\lambda_{max}$ mean is obtained. Moreover, if any of the estimates for $\lambda_{max}$ turns out to be less than *n*, or 3 in this case, then an error in the calculations has been made, which is a useful simple check of the overall calculations. Knowing the value of $\lambda_{max}$ the consistency index and consistency ratio (RI=0.58, for n=3, Table 7-3) are calculated as:

$$CI = \frac{\lambda_{max} - n}{n-1} = \frac{3.00554-3}{2} = 0.00277 \qquad \text{and} \qquad CR = \frac{CI}{RI} = 0.00477$$

Since the CR is < 0.1, the matrix is consistent and can be used for the calculation of the final priorities. The same process was followed for all the other pairwise matrices. Finally, the unweighted supermatrix was derived as shown in the Table 7-9 (in bold are indicated the values of the example given above, while the zero values suggest that there is no influence between these elements). The responses of the SME to obtain the supermatrix are included in the Appendix XI.

Table 7-9 The *unweighted* supermatrix

| $C_{11}$ | | Dynamic personal | | | Organisational | | | Personal | | System | Task | Team | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $e_{11}$ | | D | F | P | Pr | SC | S | F | T | SD | W | C | I |
| $W_{11}$ | | | | | | | | | | | | | |
| Dynamic personal | D | 0.0000 | 1.0000 | 0.5000 | 0.0000 | 0.0000 | 0.4934 | 0.5469 | 0.0000 | 0.4286 | 0.1248 | **0.4665** | 0.3333 |
| | F | 0.8333 | 0.0000 | 0.5000 | 0.0000 | 1.0000 | 0.3108 | 0.1085 | 0.0000 | 0.4286 | 0.7248 | **0.4330** | 0.3333 |
| | P | 0.1667 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.1958 | 0.3445 | 0.0000 | 0.1429 | 0.1504 | **0.1005** | 0.3333 |
| Organisational | Pr | 0.3333 | 0.3275 | 0.4545 | 0.0000 | 0.6667 | 0.5000 | 0.6044 | 0.3255 | 0.4330 | 0.6337 | 0.4761 | 0.4330 |
| | SC | 0.3333 | 0.4126 | 0.4545 | 0.8333 | 0.0000 | 0.5000 | 0.3255 | 0.0701 | 0.4665 | 0.1744 | 0.4523 | 0.4665 |
| | S | 0.3333 | 0.2599 | 0.0909 | 0.1667 | 0.3333 | 0.0000 | 0.0701 | 0.6044 | 0.1005 | 0.1919 | 0.0716 | 0.1005 |
| Personal | F | 0.1250 | 1.0000 | 0.1429 | 0.0000 | 0.5000 | 0.5000 | 0.0000 | 1.0000 | 0.5000 | 0.8333 | 0.5000 | 0.5000 |
| | T | 0.8750 | 0.0000 | 0.8571 | 1.0000 | 0.5000 | 0.5000 | 1.0000 | 0.0000 | 0.5000 | 0.1667 | 0.5000 | 0.5000 |
| System | SD | 1.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | 1.0000 |
| Task | W | 1.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 | 0.0000 | 1.0000 | 1.0000 |
| Team | C | 0.8571 | 0.5000 | 0.8000 | 0.5000 | 0.5000 | 0.5000 | 0.1667 | 0.8750 | 0.5000 | 0.2500 | 0.0000 | 1.0000 |
| | I | 0.1429 | 0.5000 | 0.2000 | 0.5000 | 0.5000 | 0.5000 | 0.8333 | 0.1250 | 0.5000 | 0.7500 | 1.0000 | 0.0000 |

D: Distraction, F: Fatigue, P: Perception, Pr: Procedures, SC: Safety culture, S: Supervision, F: Familiarity, T: Training, SD: System design, W: Workload, C: Communication, I: Information

### 7.5.7.2 Categories pairwise comparisons – the cluster supermatrix

The relative measurement of all R-PSFs categories with respect to their influence on human performance is estimated in a similar way. In particular, SMEs were asked to assess the relative importance of one of two R-PSFs categories over the second, on influencing a third R-PSFs category with respect to the influence of R-PSFs on operators performance. Table 7-10 illustrates an example of the matrix that the experts were asked to fill in. For this example, which refers to the *Personal Factors* category, participants were asked to assess the importance of the R-PSFs by replying to the question whether "*considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the* ***personal*** *factors*?"

Table 7-10 Example of a pairwise comparison matrix for a category of factors

| Personal | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Personal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Task | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

Again, pairwise matrices for all parent elements and their children are created, the priorities for each matrix are calculated and finally the consistency ratio for each matrix is obtained. Table 7-11 illustrates the pairwise comparison matrix for the *Personal Factors* cluster.

Table 7-11 Pairwise comparisons for clusters

| Personal | Dynamic | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| Dynamic | 1 | 3 | 1 | 1 | 1/2 | 3 |
| Organisational | 1/3 | 1 | 1/2 | 1 | 1/2 | 2 |
| Personal | 1 | 2 | 1 | 2 | 1 | 2 |
| System | 1 | 1 | 1/2 | 1 | 1/2 | 2 |
| Task | 2 | 2 | 1 | 2 | 1 | 1 |
| Team | 1/3 | 1/2 | 1/2 | 1/2 | 1 | 1 |

Since the consistency ratio of the matrix is 0.06069, it is raised to its $5^{th}$ power until the desired priorities are derived (indicated in bold in the Table 7-12). Once all priorities are obtained, the cluster matrix is designed as shown in the Table 7-12. Similar to the unweighted supermatrix, the responses of the SME used to obtain the cluster supermatrix are provided in the Appendix XI.

Table 7-12 The cluster matrix

| | Dynamic Personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| **Dynamic Personal** | 0.20020 | 0.03150 | **0.20627** | 0.23450 | 0.22300 | 0.08305 |
| **Organisational** | 0.12824 | 0.29324 | **0.11616** | 0.13540 | 0.14519 | 0.27615 |
| **Personal** | 0.14675 | 0.06685 | **0.21213** | 0.29556 | 0.22031 | 0.14336 |
| **System** | 0.20505 | 0.15508 | **0.13772** | 0.00000 | 0.26229 | 0.17543 |
| **Task** | 0.11659 | 0.19052 | **0.22900** | 0.18937 | 0.00000 | 0.09322 |
| **Team** | 0.20318 | 0.26281 | **0.09872** | 0.14518 | 0.14921 | 0.22880 |

Results indicate that the *task*, *personal* and *dynamic personal* R-PSFs categories almost equally affect the *personal* factors category. In contrast, according to the SME, *team* factors are considered to have a very limited impact on *personal* factors. In addition, the direct and indirect influence of one category of factors on another are derived. For example, as shown in the Figure 7-5, for the *dynamic personal* category it can be observed that only 20% is explained by direct influence, i.e. internal dependencies, whereas the remaining 80% is justified by indirect influence. On the other hand, as no internal dependencies exist amongst the 12 R-PSFs within the *system* and *task* clusters, the direct influence of those two categories is zero.

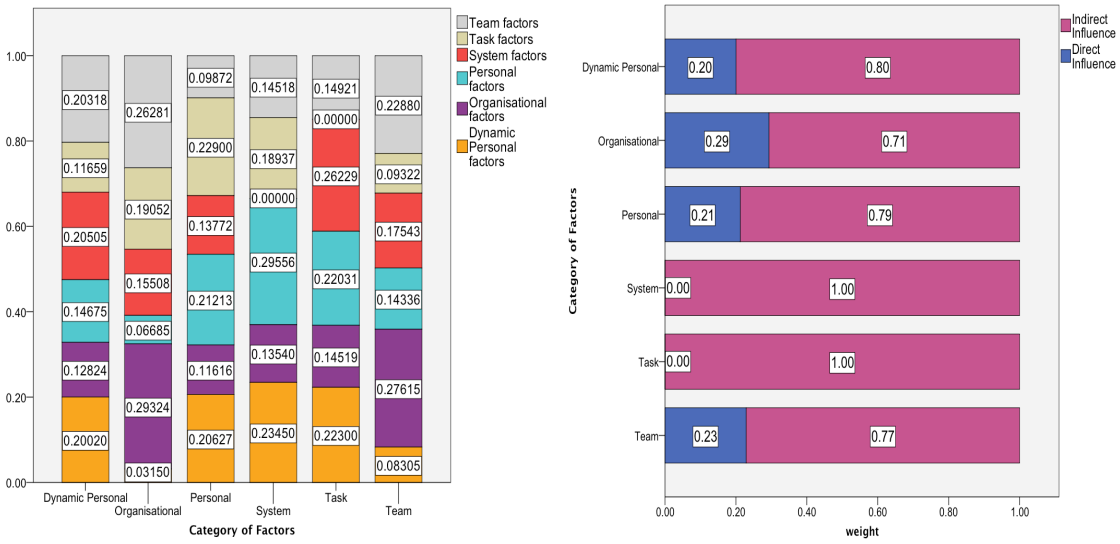



Figure 7-5 The direct and indirect influence of factors categories

### 7.5.7.3 The weighted supermatrix

Once the unweighted and cluster matrices are estimated, the weighed supermatrix is computed by multiplying each of the values in the cluster columns with the corresponding values in the unweighted supermatrix. To compute the values in the column *communication* (indicated in bold font), for example, the values in column *team* of the cluster matrix with the values in column *communication* in the unweighted matrix are multiplied. That is for the first three rows, which correspond to *dynamic personal* factors category,

$$0.08305\times0.46647 = 0.03874, \qquad 0.08305\times0.43303 = 0.03596, \qquad 0.08305\times0.10050 = 0.00835$$

Following the same rational, the weighted supermatrix is derived, as shown in Table 7-13.

Table 7-13 The weighted supermatrix

| | | Dynamic personal | | | Organisational | | | Personal | | System | Task | Team | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | D | F | P | Pr | SC | S | F | T | SD | W | C | I |
| Dynamic personal | D | 0.0000 | 0.2002 | 0.1001 | 0.0000 | 0.0000 | 0.0155 | 0.1128 | 0.0000 | 0.1005 | 0.0278 | **0.0387** | 0.0277 |
| | F | 0.1668 | 0.0000 | 0.1001 | 0.0000 | 0.0315 | 0.0098 | 0.0224 | 0.0000 | 0.1005 | 0.1616 | **0.0360** | 0.0277 |
| | P | 0.0334 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0062 | 0.0711 | 0.0000 | 0.0335 | 0.0335 | **0.0083** | 0.0277 |
| Organisational | Pr | 0.0427 | 0.0420 | 0.0583 | 0.0000 | 0.1955 | 0.1466 | 0.0702 | 0.0670 | 0.0586 | 0.0920 | **0.1315** | 0.1196 |
| | SC | 0.0427 | 0.0529 | 0.0583 | 0.3923 | 0.0000 | 0.1466 | 0.0378 | 0.0144 | 0.0632 | 0.0253 | **0.1249** | 0.1288 |
| | S | 0.0427 | 0.0333 | 0.0117 | 0.0785 | 0.0977 | 0.0000 | 0.0081 | 0.1243 | 0.0136 | 0.0279 | **0.0198** | 0.0278 |
| Personal | F | 0.0183 | 0.1467 | 0.0210 | 0.0000 | 0.0334 | 0.0334 | 0.0000 | 0.3756 | 0.1478 | 0.1836 | **0.0717** | 0.0717 |
| | T | 0.1284 | 0.0000 | 0.1258 | 0.1073 | 0.0334 | 0.0334 | 0.2121 | 0.0000 | 0.1478 | 0.0367 | **0.0717** | 0.0717 |
| System | SD | 0.2051 | 0.2051 | 0.2051 | 0.0000 | 0.1551 | 0.1551 | 0.1377 | 0.2439 | 0.0000 | 0.2623 | **0.1754** | 0.1754 |
| Task | W | 0.1166 | 0.1166 | 0.1166 | 0.0000 | 0.1905 | 0.1905 | 0.2290 | 0.0000 | 0.1894 | 0.0000 | **0.0932** | 0.0932 |
| Team | C | 0.1742 | 0.1016 | 0.1625 | 0.2110 | 0.1314 | 0.1314 | 0.0165 | 0.1530 | 0.0726 | 0.0373 | **0.0000** | 0.2288 |
| | I | 0.0290 | 0.1016 | 0.0406 | 0.2110 | 0.1314 | 0.1314 | 0.0823 | 0.0219 | 0.0726 | 0.1119 | **0.2288** | 0.0000 |

D: Distraction, F: Fatigue, P: Perception, Pr: Procedures, SC: Safety culture, S: Supervision, F: Familiarity, T: Training, SD: System design, W: Workload, C: Communication, I: Information

#### 7.5.7.4 The limit supermatrix – the R-PSFs final weightings

For the calculation of the limit priorities, the weighted matrix is raised to limiting powers until its columns become identical. Raising the weighted matrix to its $5^{th}$ power, the desired limit supermatrix, which contains the limit priorities, is obtained, as shown in the Table 7-14.

Final results indicate that the respondent considers *system design* as the most significant contributor factor to the performance of operators. In turn, he identifies *workload* as the second more important factor, while *communication* and *information* as the third and fourth respectively. These values represent the weightings in the HuPeROI equation.

Table 7-14 The limit supermatrix – final priorities

| | | Dynamic personal | | | Organisational | | | Personal | | System | Task | Team | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | D | F | P | Pr | SC | S | F | T | SD | W | C | I |
| Dynamic personal | D | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 | 0.0503 |
| | F | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 | 0.0554 |
| | P | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 | 0.0216 |
| Organisational | Pr | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 | 0.0873 |
| | SC | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 | 0.0902 |
| | S | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 | 0.0414 |
| Personal | F | 0.1029 | 0.1030 | 0.1030 | 0.1030 | 0.1029 | 0.1029 | 0.1030 | 0.1030 | 0.1030 | 0.1030 | 0.1030 | 0.1030 |
| | T | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 | 0.0855 |
| System | SD | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 | 0.1470 |
| Task | W | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 | 0.1107 |
| Team | C | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 | 0.1064 |
| | I | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 | 0.1014 |

D: Distraction, F: Fatigue, P: Perception, Pr: Procedures, SC: Safety culture, S: Supervision, F: Familiarity, T: Training, SD: System design, W: Workload, C: Communication, I: Information

### 7.5.8 The ratings of the R-PSFs

For the rating of the R-PSFs the concept of the SLIM methodology is adopted. Thus, the SMEs are asked to consider the quality of the identified factors for each of the operational erroneous actions. According to the scenario, a train driver fails to stop the train before passes a signal that displays red aspect (SPAD). In this situation, the three operational errors are related to the: (i) signal detection, (ii) signal interpretation and finally (iii) executed action, which are in absolute agreement with the concept of the railway operational errors, as described in the Section 4.1.

For each of the errors, the SMEs were asked to rate to what extent each of the R-PSFs were good or bad in the action under consideration. The participants indicated their responses in the Table 7-15, which here illustrates the answers of the SME who was also used as example for the assessment of the R-PSFs.

Table 7-15 The rating of R-PSFs for the three type of actions (potential errors)

| Rating of Railway Performance Shaping Factors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Scenario - Errors** | **R-PSFs** | | | | | | | | | | | |
| **SPAD**<br>Train driver, open line (not tunnel), day operation, good weather conditions, good visibility | **Training**<br>(incl. experience) | **Familiarity**<br>(incl. expectation and routine) | **Distraction**<br>(incl. concentration, vigilance, situational awareness loss) | **Fatigue**<br>(incl. shift pattern and fitness to work) | **Perception**<br>(incl. interpretation) | **Workload**<br>(incl. stress and time pressure) | **Communication**<br>(incl. teamwork) | **Quality and Trust in Information** | **Safety culture**<br>(incl. SMS) | **Quality of Procedures** | **Supervision** | **System design**<br>(incl. HMI) |
| **Job at time of SPAD**<br>Train driver **fails** to stop the train before passes signal at danger | 0 = no training,<br>50= some training,<br>100= very good training | 0 = no familiarity<br>50 = some level of familiarity<br>100 = high levels of familiarity | 0 = very distracted<br>50 = some distraction<br>100 = no distraction | 0 = too fatigued<br>50 = some level of fatigue<br>100 = no fatigued | 0 = no perception<br>50 = some perception<br>100 = high perception levels | 0 = extreme workload<br>50 = some level of workload<br>100 = no workload | 0 = poor quality<br>50 = average quality<br>100 = very good quality | 0 = poor quality<br>50 = average quality<br>100 = high level quality | 0= poor safety culture<br>50 = some safety culture<br>100= very good safety culture | 0 = poor quality<br>50 = average quality<br>100 = very good quality | 0 = no supervision distracted<br>50 = some supervision<br>100 = very good supervision | 0 = poor system design<br>50 = average quality<br>100 = very good system design |
| 1. **Signal detection** | 80 | 80 | 10 | 30 | 20 | 60 | 70 | 70 | 70 | 80 | 80 | 50 |
| 2. **Signal interpretation** | 80 | 60 | 20 | 40 | 10 | 60 | 70 | 70 | 70 | 80 | 80 | 50 |
| 3. **Action executed** | 70 | 40 | 10 | 30 | 20 | 50 | 60 | 40 | 70 | 50 | 50 | 50 |

The R-PSFs rating is shown in the Figure 7-6. It can be seen, that the respondent had a constant perception of the extent to which each R-PSF is optimal or suboptimal with respect to each action.

In particular, it is shown that for the tasks *signal detection* and *signal interpretation* only the *dynamic personal factors* are account for as being strongly suboptimal, while the majority of remainder factors can be claimed to be positive. On the other hand, it is clear that the SME considers that the R-PSFs tend to be much more suboptimal in the case of the *action executed correctly*. For this action, the *dynamic* R-PSFs have a negative impact on the *task*, *team* and *organisational* factors.

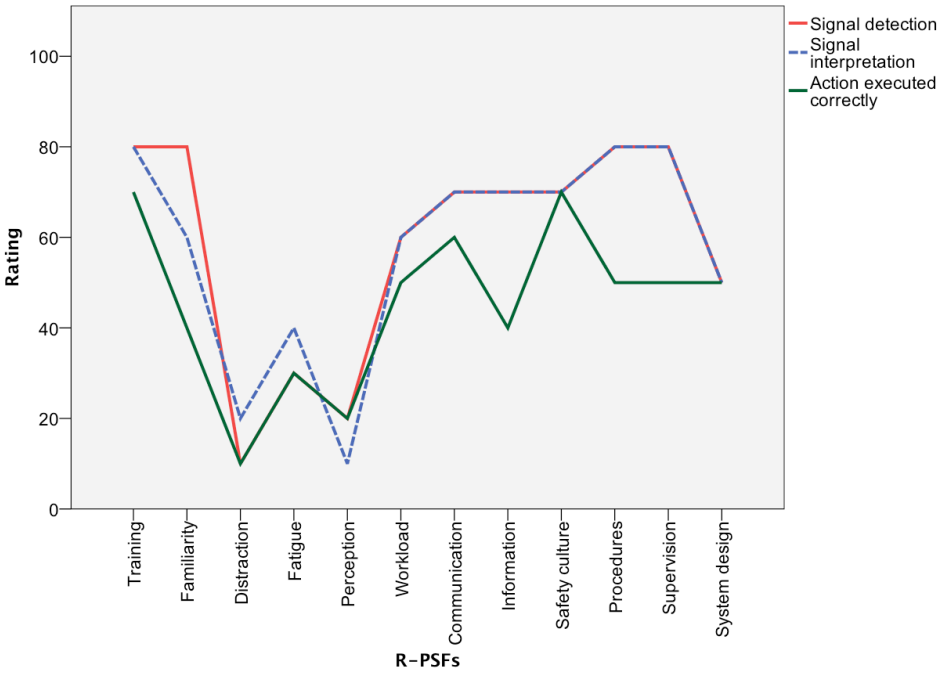



Figure 7-6 The rating of R-PSFs per type of error

### 7.5.9 The HuPeROI calculations

Having assessed (weighted) the contribution of each R-PSF to human performance and rated to what extent every factor is positive or negative with respect to an erroneous action for the defined scenario, the HuPeROI for each action is then calculated using the equation:

$$HuPeROI_j = \sum_{i=1}^{n} w_i \times r_{ij} \qquad (7\text{-}3)$$

Based on the derived responses the HuPeROI for the first action, the signal detection, is computed as:

$$HuPeROI_{signal\ detected} = \left(w_{distraction} \times r_{distraction,\ signal\ detected}\right) + \left(w_{fatigue} \times r_{fatigue,\ signal\ detected}\right) + \cdots + \left(w_{information} \times r_{information,\ signal\ detected}\right) = 62.8116$$

Similarly, the other two indexes are found

$$HuPeROI_{signal\ interpreted} = 61.5934$$

$$HuPeROI_{action\ executed\ correctly} = 48.7667$$

The HuPeROI, based on the concept of SLIM, reflects the overall belief of the SMEs, regarding the positive or negative effects of the R-PSFs on the likelihood of success for the task under consideration. Braking down the results, it can be observed that for all types of errors, distraction and fatigue are considered the most suboptimal factors; that is, they degraded the performance of the operator the most. In addition, the ratings for the three different errors can be compared. It can then be seen that from the ratings of the factors that the quality of information, quality of procedures and supervision contributed significantly to the final outcome, which indicates the *action executed correctly* as the least successful action. This is explained firstly, by the fact that all the three R-PSFs were assigned with much lower scores for this action. Secondly, by the fact that the factors' quality of information and quality of procedures were weighted with relatively high scores amongst the R-PSFs, and thus contributed to the final calculations respectively.

In summary, as described in the Section 7.4, the HuPeROIs indicate the relative likelihoods of different erroneous actions, where the greater the HuPeROI the greater the probability of success for the task (action) under consideration. Within this context, the HuPeROI should be used as an indicator to define the most prone error amongst a series of actions. It should also be used to highlight the most significant contributing factors for such an error.

Subsequently, it can be claimed that the HuPeROI results can be used prospectively in order to identify the most likely human errors, between a set of errors. In addition, the findings can be used to explore how the defined R-PSFs (i) have been assessed to contribute to the performance of the operators, and (ii) have been rated regarding their influence on the reliability of operators to perform their tasks.

### 7.5.10 The change of overall performance under certain conditions

As mentioned in the Section 7.5.4 in the last part of the questionnaire, the SMEs were also asked to indicate the influence of specific external conditions on their performance for the SPAD scenario. These include: adverse and intermediate weather conditions, night operations, degraded operations, i.e. speed restrictions, and finally operations in tunnels (only for the underground services). For this, a 0 to 100 scale was used, where 0 indicates no influence on performance, while 100 expresses the maximum of influence (either improving or reducing), as shown in the Appendix XI. Again, the purpose of those questions was to indicate whether the SMEs consider any special conditions as factors that may affect their overall performance.

### 7.5.11 Possible sources of errors and limitations of the HuPeROI

The HuPeROI framework presented in this thesis addresses particular weaknesses concerning the modelling of factors that affect human performance. In particular, compared to the traditional HPA techniques, it determines the influence of the R-PSFs on the performance of railway operators accounting for both direct and indirect dependencies amongst the relevant factors. Additionally, it allows researchers to identify the relative likelihoods of different errors for a certain task. However, the approach has limitations and possible sources of errors. The latter refer primarily to the shortcomings of the survey, while the former to the shortcomings of the methodology and its outcomes.

#### 7.5.11.1 Possible sources of errors

There are two main potential sources of errors in the survey, respondent and data pre-processing. The respondent errors may primarily occur due to the imprecision in the formulation of questions and the inconsistent interpretation of the factors to be assessed. The former was addressed by the validation of the questionnaire and the preliminary study, as described in the Section 7.5.4 and 7.5.5. The case of inconsistent interpretation was also tackled by the validation of the questionnaire, while the distribution of the definitions and examples of the factors should have mitigated the creation of such errors. Finally, the possibility of errors arising from the pre-processing of the responses was reduced by due diligence at the data completion and input stage of data for the calculations. The former was ensured by the physical presence of the author during the completion of questionnaires, while the latter with double-checking of each input by an independent, third assessor.

#### 7.5.11.2 Limitations of the HuPeROI

There are two principal limitations related to the HuPeROI framework. The first is the complex approach to weighting the R-PSFs. The second reflects the SMEs' understanding of the fundamental scale and the translation of a person's judgment into a number. Furthermore, HuPeROI does not calculate error probabilities for the different types of actions.

For the R-PSFs weightings, a large number of pairwise comparisons is required, to represent all the factor dependencies and subsequently, to determine their final contribution to human performance. For example, in the survey presented in this thesis each of the SMEs was asked to complete 18 matrices and to conduct 148 pairwise comparisons. This makes the methodology cumbersome especially for those participants with no previous experience in similar studies, e.g. train drivers.

In addition, the large number of matrices is problematic for both the respondents and researchers. In particular, for this research the author transformed 52 questionnaires, and entries from 936 matrices and in total almost 140,000 pairwise comparisons into data that could be analysed further. This is time consuming, and requires extra care at the input stage of the data for the calculations.

The second limitation is also related to the conventional ANP model and reflects on its inefficiency to capture any inherent fuzziness or uncertainty in the SMEs judgments during the process of the pairwise comparisons. Although the simplicity of the fundamental 1-9 scale makes it easily understandable by the assessors (Zhang et al., 2009), it fails to account for any potential uncertainties related to the mapping of one's perception or judgment to a number. Thus, HuPeROI is developed on the assumption that all the SMEs perceive in the same way the importance of each number compared to the other numbers in the scale.

In addition to the limitations above, a third weakness of HuPeROI is the fact that it does not provide error probabilities for the different types of erroneous actions. In other words, although HuPeROI allows researchers to identify the ranking of the likelihoods of the erroneous actions, it does not express the different likelihoods in terms of probabilities.

Despite its limitations, HuPeROI is the first complete framework tailored to the railway industry, which allows researchers to:

(i) determine the factors that affect the performance of the railway operators for a broad variety of scenarios;

(ii) assess the importance of the factors that affect human performance, accounting for all the dependencies amongst the factors;

(iii) define the quality of the factors for several erroneous actions for a task under consideration and finally;

(iv) indicate the most erroneous action among a series of actions.

HuPeROI's results can be used to provide insights on operator performance for several railway operational tasks. This should enable railway organisations to develop sound solutions to enhance human performance, firstly by improving the quality of the factors that affect railway operators and, secondly, by mitigating the impact of those factors on human performance.

## 7.6 Summary

This chapter presented the development of a new approach referred to as the Human Performance Railway Operational Index (HuPeROI) and its underlying theory. The HuPeROI aims firstly to assess the influence of each of the most significant contributing factors on the performance of railway operators. Secondly to rate how good or bad each of the factors is considered for different human errors and, thirdly, to provide insights on operators' performance by comparing the success likelihood indexes of different human actions.

The chapter started with the introduction of the different Multi Criteria Decision Making (MCDM) methodologies, which could be applied to assess the contribution of each of the R-PSFs on human performance. This was followed by the justification of the relevance of the Analytic Network Process (ANP) as the most appropriate approach to be used for the development of the HuPeROI. The ANP methodology was then detailed and the integration of the ANP and SLIM approaches demonstrated. Furthermore, the steps to weight and rate each of the R-PSFs as well as to calculate the HuPeROI were explained in detail. This was followed by the implementation of the HuPeROI based on a case study. The analysis and results of the case study are presented thoroughly in the next chapter.

# Chapter 8 ANALYSIS AND RESULTS OF THE HUPEROI CASE STUDY

The previous chapter developed the HuPeROI framework and presented the background to the case study for this thesis. This chapter applies the HuPeROI framework to the case study, and discusses the results.

The chapter starts by describing the profiles of the SMEs, followed by the presentation of the weightings for categories and elements of the Railway Performance Shaping Factors (R-PSFs). The ratings of the R-PSF elements are then specified and the indexes for the three different erroneous actions are calculated. The chapter concludes with the validation of the results and summary of the findings.

## 8.1 The sample characteristics

In was highlighted in the Section 7.5.6 that a total of 56 responses were collected throughout the survey from SMEs belonging to three UK train operator companies (Piccadilly Line, LUL, First Hull Trains, and FirstScot Rail) and one HFs (RSSB) experts group. The responses for the weighting of the factors, i.e. the assessment of factors impact on human performance, were transformed into pairwise matrices. The matrices represent the strength of dominance of one element over another element. For each matrix the consistency ratio of each determined. This resulted in 45 acceptable responses. For the remaining 11, the process described in the Section 7.3.2 was followed. Firstly, the most inconsistent (larger) assessment in the matrix was defined. Secondly, the range of values to which that assessment can be changed to improve inconsistency were determined. Finally, the respondents were asked to change their assessments to a plausible value in that range. Four sets of matrices were rejected, as the respondents refused to modify their answers. This resulted eventually in 52 questionnaires and usable responses from 35 train drivers, 14 operations managers and 3 HFs experts. The survey captured interesting information on the demographics of the SMEs, i.e. age groups, gender, and operational experience. Firstly, the results confirm the general belief and published studies, e.g. (Robison, 2012), that UK railway operations are run mainly by male professionals. In particular, only 8.6% of the train drivers are female. In addition, results show that about 60% of the train drivers have more

than 11 years of experience, while 64.3% of the operations managers have more than 20. Furthermore, about 30% of the train drivers have been involved in the occurrence of at least one SPAD. The design to capture the representative diversity of the SMEs was based on previous studies in the field, e.g. (Ryan et al., 2009). The main characteristics of the sample are presented in Table 8-1.

Table 8-1 Characteristics of the sample

| Characteristics of sample | | |
|---|---|---|
| **Type of employees** | **Frequency** | **Percent** |
| Train drivers | 35 | 67.3 |
| Operations managers | 14 | 26.9 |
| HFs experts | 3 | 5.8 |
| **Gender** | **Frequency** | **Percent** |
| Male | 47 | 90.4 |
| Female | 5 | 9.6 |
| **Age group** | **Frequency** | **Percent** |
| Under 25 | 1 | 3.8 |
| 25-34 | 7 | 13.5 |
| 35-44 | 17 | 32.69 |
| 45-56 | 22 | 40.38 |
| Over 56 | 5 | 9.6 |
| **Operational experience** | **Frequency** | **Percent** |
| Less than a year | 7 | 13.5 |
| 1 to 5 | 1 | 3.8 |
| 6 to 10 | 10 | 19.2 |
| 11 to 19 | 13 | 19.2 |
| 20 or more | 21 | 42.3 |
| **Involved to a SPAD before** | **Frequency** | **Percent** |
| Yes | 15 | 28.8 |
| No | 37 | 71.2 |

It can be seen from Table 8-1 that the frequencies within the classification subgroups of age and operational experience vary considerably. Therefore, a second classification was performed to minimise such variations and to facilitate the analysis and interpretation of results. Firstly, for the age group, it was found that the youngest SME was older than 23 years at the time of the survey. Therefore, this SME was moved to the next group, renamed *24-35 years old*. Similarly, one SME within the original group with 1-5 years of operational experience was also moved to the next group, renamed *5-10 years of experience*. This was because this SME had more than 3 years of service. In addition, the groups *less than a year* and *1-4 years of operational experience* were not merged, as no SME was identified to represent the second range of years of operational experience. Table 8-2 shows the characteristics of the sample after re-grouping.

Table 8-2 Characteristics of the sample after regrouping

| Characteristics of sample | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Train Driver | | | | Operations Manager | | | |
| | | Age group | | | | Age group | | | |
| | | 24-34 | 35-44 | 45-56 | Over 56 | 24-34 | 35-44 | 45-56 | Over 56 |
| **Gender** | Male | 6 | 7 | 15 | 4 | – | 6 | 7 | 1 |
| | Female | 1 | 2 | – | – | – | – | – | – |
| **Operational experience** | < 1 year | 5 | 1 | 1 | – | – | – | – | – |
| | 1 to 4 yrs | – | – | – | – | – | – | – | – |
| | 5 to 10 yrs | 2 | 4 | 1 | – | – | – | 2 | – |
| | 11 to 19 yrs | – | 4 | 3 | 2 | – | 2 | – | 1 |
| | ≥ 20 yrs | – | – | 10 | 2 | – | 4 | 5 | – |
| **SPAD involvement** | No | 6 | 8 | 9 | 2 | – | 3 | 6 | – |
| | Yes | 1 | 1 | 6 | 2 | – | 3 | 1 | 1 |

From the characteristics, and although there are a number of empty cells, the sample could be argued to be reasonably representative of the population. For example, it is not surprising that none of the operations managers was under 34 years of age or with less than 5 years operational experience. This is because they tend to be ex-train drivers who move to operational support roles, including training, instruction and management. On the other hand, it is a limitation that no train drivers with 1-4 years of experience are included in the sample. This should be addressed in future research. Therefore, generalisation of the substantive findings from the analysis of the resulting dataset to passenger TOCs at large requires caution, statistical weighting and the exercise of judgment. Moreover, any attempt to generalise the results should also account for substantial differences amongst the operations systems; e.g. the Piccadilly Line services operate on a manual basis, whilst the Jubilee Line services are automated. Finally, although the demographics of the HFs experts cannot be considered representative, the HFs review team consists of two females and one male, between 35-44 years old, with at least 5 years of experience in the railway industry (but more than 10 years in the field of HFs).

Having defined the groups of interest for the analysis, the next section presents the results from the pairwise comparisons for the weightings of the R-PSFs categories. The analysis explores whether the derived data are related to the different types of the SMEs, i.e. train drivers, operations managers, and HFs experts. Therefore, comparison of the responses determined by the members of each subgroup is performed to identify if the results of any subgroup vary significantly. In addition, for the train drivers, it is examined whether the responses are affected by gender, age, level and type of experience.

## 8.2 The weightings of the R-PSFs *lite* categories

This section presents the findings of the weightings for the R-PSFs categories, which result in the cluster supermatrix for each participant. The normality of data was first tested using the Shapiro-Wilk test, which compares the scores in the sample to a normally distributed set of scores with the same mean and standard deviation. If the test is significant, ($p<.05$), the distribution of the sample is significantly different from a normal distribution and vice versa (Field, 2009). The Shapiro-Wilk test was used, as it is more powerful to detect differences from normality, compared, for example, to Kolmogorov-Smirnov, Anderson-Darling and Lilliefors tests (Field, 2009, Razali and Wah, 2011). Results show that the majority of the data are not normally distributed, therfore necessitating the use of the relevant non-parametric tests, i.e. Kruskal-Wallis, Mann-Whitney and Friedman's, and ANOVA tests.

Data were afterwards tested for homogeneity, to indicate whether the responses collected from the different subgroups vary significantly. For this, Kruskal-Wallis tests were applied, as described in the Section 6.3.2. When results indicate significant differences, *post hoc* tests were used to identify which particular participants are responsible for such differences. Therefore, the Mann-Whitney with Bonferroni correction (MW/B) non-parametric test was applied. Similar to Kruskal-Wallis, MW/B is based on the ranking and grouping of data. The Mann-Whitney U statistic is calculated as:

$$U = n_1 n_2 + \frac{N_1(N_1+1)}{2} - R_1 \qquad (8\text{-}1)$$

where $n_1$ and $n_2$ are the samples sizes of groups 1 and 2 respectively, and $R_1$ is the sum of ranks for group 1. The test is applied to each pairwise combination of variables to examine the null hypothesis whereby the two samples are not significantly different. In practice, the test statistic U represents the difference between the two rank totals, where the smaller the U (considering the number of participants in each group), the less likely it is that such a difference occurred by chance. In turn, the Bonferroni correction is used to avoid accidentally rejecting the null hypothesis by using more rigid rejection criteria derived from the initial significance level divided by the number of pairs being analysed (Field, 2009). Finally, Field (2009) suggests that when the sample sizes are less than 25 per group, then the Kolmogorov-Smirnov Z (KS-Z) test should be applied. It should be noted that where the results from the KS-Z test are not presented, there are no significant differences compared to the Mann-Whitney results.

Table 8-3 Descriptive statistics for the R-PSFs *lite* categories weighting

| | Variable | N* | Mean | SD | SE | 95% confidence interval | | Min | Max | Normality** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound | | | |
| Dynamic | Dynamic | 52 | **0.3059** | 0.1459 | 0.0202 | 0.2653 | 0.3466 | 0.0506 | 0.588 | ✓ |
| | Personal | 52 | **0.1464** | 0.0930 | 0.0129 | 0.1205 | 0.1723 | 0.0415 | 0.390 | ✗ |
| | Organisational | 52 | **0.1707** | 0.0864 | 0.0119 | 0.1466 | 0.1947 | 0.0502 | 0.387 | ✗ |
| | Team | 52 | **0.0753** | 0.0518 | 0.0072 | 0.0609 | 0.0898 | 0.0225 | 0.203 | ✗ |
| | Task | 52 | **0.1937** | 0.1213 | 0.0168 | 0.1599 | 0.2275 | 0.0325 | 0.561 | ✗ |
| | System | 52 | **0.1077** | 0.0636 | 0.0088 | 0.0900 | 0.1255 | 0.0313 | 0.294 | ✗ |
| Personal | Dynamic | 52 | **0.2832** | 0.1384 | 0.0192 | 0.2446 | 0.3217 | 0.0252 | 0.534 | ✓ |
| | Personal | 52 | **0.1900** | 0.1030 | 0.0143 | 0.1613 | 0.2187 | 0.0620 | 0.506 | ✗ |
| | Organisational | 52 | **0.1349** | 0.0670 | 0.0092 | 0.1162 | 0.1535 | 0.3222 | 0.395 | ✗ |
| | Team | 52 | **0.0812** | 0.0550 | 0.0076 | 0.0658 | 0.0965 | 0.0249 | 0.277 | ✗ |
| | Task | 52 | **0.1979** | 0.0987 | 0.0137 | 0.1705 | 0.2255 | 0.0438 | 0.460 | ✓ |
| | System | 52 | **0.1126** | 0.0642 | 0.0089 | 0.0948 | 0.1305 | 0.0271 | 0.372 | ✗ |
| Organisational | Dynamic | 52 | **0.2112** | 0.1451 | 0.0129 | 0.1708 | 0.2516 | 0.0251 | 0.583 | ✗ |
| | Personal | 52 | **0.1560** | 0.0862 | 0.0119 | 0.1320 | 0.1800 | 0.0351 | 0.410 | ✗ |
| | Organisational | 52 | **0.2038** | 0.1317 | 0.0182 | 0.1671 | 0.2404 | 0.0445 | 0.597 | ✗ |
| | Team | 52 | **0.1159** | 0.0720 | 0.0099 | 0.0958 | 0.1359 | 0.0310 | 0.323 | ✗ |
| | Task | 52 | **0.1721** | 0.1109 | 0.0153 | 0.1412 | 0.2030 | 0.0317 | 0.552 | ✗ |
| | System | 52 | **0.1408** | 0.0841 | 0.0116 | 0.1174 | 0.1643 | 0.0229 | 0.362 | ✗ |
| Team | Dynamic | 52 | **0.2106** | 0.1367 | 0.0189 | 0.1725 | 0.2487 | 0.0227 | 0.519 | ✗ |
| | Personal | 52 | **0.1641** | 0.0784 | 0.1088 | 0.1422 | 0.1859 | 0.0472 | 0.384 | ✗ |
| | Organisational | 52 | **0.1952** | 0.1119 | 0.0155 | 0.1640 | 0.2263 | 0.0393 | 0.521 | ✗ |
| | Team | 52 | **0.1548** | 0.1422 | 0.0197 | 0.1152 | 0.1944 | 0.0266 | 0.618 | ✗ |
| | Task | 52 | **0.1583** | 0.1004 | 0.0139 | 0.1303 | 0.1863 | 0.0353 | 0.449 | ✗ |
| | System | 52 | **0.1168** | 0.0544 | 0.0075 | 0.1016 | 0.1319 | 0.0248 | 0.275 | ✓ |
| Task | Dynamic | 52 | **0.3255** | 0.1656 | 0.0229 | 0.2794 | 0.3716 | 0.0319 | 0.640 | ✓ |
| | Personal | 52 | **0.2022** | 0.1001 | 0.0138 | 0.1744 | 0.2301 | 0.0730 | 0.545 | ✗ |
| | Organisational | 52 | **0.2269** | 0.1375 | 0.0190 | 0.1886 | 0.2652 | 0.0471 | 0.616 | ✗ |
| | Team | 52 | **0.1010** | 0.0589 | 0.0081 | 0.0846 | 0.1174 | 0.0266 | 0.286 | ✗ |
| | System | 52 | **0.1441** | 0.0837 | 0.0116 | 0.1208 | 0.1674 | 0.0363 | 0.426 | ✗ |
| System | Dynamic | 52 | **0.2451** | 0.1595 | 0.0221 | 0.2007 | 0.2896 | 0.0282 | 0.600 | ✗ |
| | Personal | 52 | **0.1997** | 0.1068 | 0.0148 | 0.1700 | 0.2295 | 0.0463 | 0.516 | ✗ |
| | Organisational | 52 | **0.2150** | 0.1309 | 0.0181 | 0.1785 | 0.2514 | 0.0410 | 0.536 | ✗ |
| | Team | 52 | **0.1252** | 0.0827 | 0.0114 | 0.1021 | 0.1482 | 0.0292 | 0.347 | ✗ |
| | Task | 52 | **0.2148** | 0.1345 | 0.0186 | 0.1773 | 0.2522 | 0.0521 | 0.649 | ✗ |

*the mean, SE and SD when HFs responses are excluded change after the 3rd decimal
** normality at 95% level of confidence

In addition, findings were tested for outliers, i.e. data points that are not consistent with the others within a sample from the same population (Field, 2009). For this, boxplots were created for each of the variables. When outliers were identified, two datasets were created, with and without, and the statistical tests performed on both datasets. The test results showed no significant difference between the two datasets. Therefore, the following sections present the results for the complete dataset.

### 8.2.1 Influence of the R-PSFs categories on dynamic personal factors

Starting with the weightings of the R-PSFs categories with respect to the dynamic personal factors, results as shown in Table 8-4, indicate significant differences in the responses between the subgroups of the review team for two combinations of factors. The influence of the system and organisational factors on dynamic personal factors.

Table 8-4 Effect of R-PSFs categories on Dynamic Personal Factors for the SMEs subgroups

| Dynamic personal | Dynamic personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| H | 2.998 | 6.089 | 3.110 | 8.586 | 1.168 | .581 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | p>.05 | **p<.05** | p>.05 | **p<.05** | p>.05 | p>.05 |

For the former, the Mann-Whitney test suggests that the train drivers weightings were higher compared to those of operations managers and HFs experts, (U=147, r=-.309) and (U=11, r=-.364) respectively. That is, train drivers consider that the factors related to the system (e.g. system design and HMI) affect the factors related to the operator at the time of the occurrence (e.g. distraction, and loss of concentration), more compared to the judgements of the operations managers and HFs experts.

For the latter, *post hoc* tests fail to show the origin of such a difference. However, it could be assumed that the difference is justified by the judgements of trains drivers compared to those of operations managers and HFs experts, where the train drivers weight the influence of organisational factors, e.g. safety culture, on dynamic personal factors with higher scores. Subsequently, it can be claimed that the characteristics of the subgroups have an impact on the SMEs weightings. Nevertheless, in the case of the HFs experts, it should be noted that the small size of the group permits the mean value of weightings to be influenced more readily by any one judge that varies significantly from the rest of the group.

Finally, between the train drivers in particular, the results do not indicate any significant differences for any combination of factors with respect to the control groups operational experience, gender, age or previous SPAD experience. In contrast, significant difference on the weightings of organisational factors on dynamic personal factors is identified between the responses for the different types of operations ($H(2)=7.062$, $p<.05$). This may imply different levels of organisational support amongst organisations, and therefore, further research is recommended.

### 8.2.2 Influence of the R-PSFs categories on personal factors

Table 8-5 shows the results from the Kruskal-Wallis test for the influence of the R-PSFs caegories on personal factors, as derived from the responses of the different types of SMEs.

Table 8-5 Effect of R-PSFs categories on Personal Factors for the SMEs subgroups

| Personal | Dynamic personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| H | 8.230 | 4.280 | 2.290 | 1.768 | 1.097 | 3.801 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | **p<.05** | p>.05 | p>.05 | p>.05 | p>.05 | p>.05 |

Results show significant differences in the responses only for the influence of dynamic personal factors on personal factors. In particular, *post hoc* tests suggest that the weightings of the train drivers are lower compared to those of operations managers (U=138, r=-.338). In addition, significant difference amongst the responses of HFs experts and operations managers was identified for the influence of dynamic personal factors on personal factors (U=4, z=-2.016, p<.05, r=-.488).

Amongst train drivers, statistics indicate significant differences in the responses with respect to type of operation, operational experience and gender. Concerning the influence of gender, male train drivers suggest that the influence of personal factors on personal factors is higher compared to the judgments given by the female personnel (U=12, z=-2.121, p<.05, r=-.358). However, no specific explanation could be identified for this observation.

For the operational experience, significant differences were identified in the influence of organisational factors on personal factors (H(3)=9.921, p<.05). In particular, the least experienced personnel consider such an influence less important compared to personnel with 11-19 years (U=10, r=-.569) and 20 or more (U=18, r=-.465) of operational experience. This can be explained by the fact that the least experienced personnel have just completed their training program, where the importance of organisational principles and factors are highlighted. Subsequently, the operators feel confident about this factor and they do not consider them as highly important.

Finally, with respect to the type of operation, the difference (H(2)=14.787, p<.05) reflects the impact of the organisational factors on personal factors, which is similar to the one found for the dynamic personal factors and can be interpreted accordingly. Subsequently, it can be concluded that the weightings for the influence of the R-PSFs categories on personal factors is linked to the duties and characteristics of the SMEs.

### 8.2.3 Influence of the R-PSFs categories on organisational factors

The Kruskal-Wallis tests in Table 8-6 show significant differences in the responses between the SMEs for the influence of organisational factors and task factors on organisational factors.

Table 8-6 Effect of R-PSFs categories on Organisational Factors for the SMEs subgroups

| | Dynamic personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| H | 2.154 | 6.602 | .724 | 1.444 | 5.279 | .751 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | p>.05 | p<.05 | p>.05 | p>.05 | p<.05 | p>.05 |

The difference in the responses for the influence of organisational factors on organisational factors is explained due to the higher scores that the HFs experts have assigned compared to train drivers (U=11, r=-.364) and operations managers (U=4, r=-.519). On the other hand, the difference of task on organisational factors is justified by the higher scores given by the operations managers compared to train drivers (U=145, r=-.316) and HFs experts (U=14, r=-435).

For the case of train drivers, significant differences in the data were found due to the type of operations, operational experience and previous SPAD experience. In particular, drivers with previous SPAD experience suggest that the task factors affect the organisational factors significantly more than those with no SPAD history (U=57, z=-2.483, p<.05, r=-.419,). In other words such a finding may indicate that due to the task characteristics, the operator overlooked the required organisational factors, resulting in a SPAD.

In addition, with respect to operational experience, statistics show differences in the responses on the influence of dynamic personal factors (H(3)=8.841, p<.05) and system factors (H(3)=14.097, p<.05) on organisational factors. Breaking down the results for the differences on the impact of the dynamic personal factors on organisational factors, *post hoc* tests identify the two pairs involved. Firstly, the least experienced personnel have assigned higher scores compared to the employees with 5-10 years of operational experience (U=7, r=-.597). Secondly, train drivers with 5-10 years of experience have given lower scores to the impact of dynamic personal factors on organisational factors, compared to those with 11-19 years of experience (U=5, r=-.70). On the other hand, for the differences in the responses on the impact of system factors on organisational factors, three pairs involved were found. The first is between the least experienced personnel and the personnel with 5-10 years of experience, with the latter assigning higher scores compared to the former (U=2, r=-.768).

The second, between personnel with 5-10 years of experience and those with 11-19 years, with the former giving higher scores compared to the latter (U=2, r=-.834). Furthermore, the employees with 5-10 years of operational experience assigned higher scores to the impact of system factors on organisational factors compared to the most experienced personnel, (U=8, r=-.659). Finally, significant differences, were identified for the assessments of train drivers on the influence of organisational factors (H(2)=7.872, p<.05) and task factors (H(2)=6.657, p<.05) on organisational factors. Consequently, similar to all the previous R-PSFs categories, it is clear that the characteristics and duties of the SMEs affect their assessments.

### 8.2.4 Influence of the R-PSFs categories on team factors

Table 8-7 presents the Kruskal-Wallis results for the influence of the R-PSFs categories on team factors as derived from the responses of the SMEs subgroups.

Table 8-7 Effect of R-PSFs categories on Team Factors for the SMEs subgroups

| | Dynamic personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| H | 3.217 | 10.059 | 1.884 | 4.170 | 5.689 | 3.042 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | p>.05 | p<.05 | p>.05 | p>.05 | p<.05 | p>.05 |

Results show that the SMEs judgements differ significantly with respect to how they assess the influence of the organisational factors and task factors on team factors. In particular, for the first pair of factors, train drivers indicate such an influence with higher scores compared to the operations managers (U=104, r=-.446). For the second pair of factors, the judgments of train drivers differ with those of operations managers. However, this time the latter suggests that task factors are significantly more influential than team factors (U=159, r=-.272).

Amongst train drivers, the significant differences identified are explained by the gender and operational experience of the SMEs. In particular, it was found that in contrast to their male colleagues, female train drivers consider that the system factors have a higher effect on team factors (U=4, z=-2.593, r=-.438, p<.05). Regarding operational experience, significant differences were identified for the importance of dynamic personal factors on team factors (H(3)=3.034, p<.05) and team factors on themselves (H(3)=8.549, p<.05). Again, if the results are broken down, it is seen that the least experienced train drivers indicate that the dynamic personal factors have a higher effect, compared to the indications of employees

with 5-10 years of operational experience (U=8, r=-.563). In contrast, for the influence of team factors on themselves, the more experienced train drivers from the same pair, indicates a higher effect compared to the indications of the least experienced drivers (U=5, r=-.666). Finally, the drivers with 11-19 years of experience assigned higher scores to the influence of dynamic personal factors on team factors, compared to their colleagues with 5-10 years of experience (U=6, r=-.674).

### 8.2.5 Influence of the R-PSFs categories on task factors

This section presents the results of the assessments by the SMEs of the influence of the R-PSFs categories on task factors. Table 8-8 demonstrates the findings for the different types of SMEs. However, it should be reminded here that the task factors within the R-PSFs *lite* taxonomy do not indicate any inner dependencies. Therefore, the corresponding cell has zero value.

Table 8-8 Effect of R-PSFs categories on Task Factors for the SMEs subgroups

| | Dynamic personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| H | 15.364 | 7.194 | 1.114 | 4.794 | .000 | 3.795 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | p<.05 | p<.05 | p>.05 | p>.05 | 1.000 | p>.05 |

As can be seen in Table 8-8 the SMEs judgements differ significantly regarding the influence of dynamic personal and organisational factors on task factors. Specifically, for the first pair of factors, train drivers (U=72, r=-.547) and HFs experts (U=4, r=-.519) assigned lower scores than the operations managers. However, for the second pair, the operations managers rate the influence of organisational factors on task factors as significantly higher (U=128, r=-.369) compared to train drivers. For all the remaining combinations of factors, no significant differences amongst the respondents were identified.

Furthermore, for the train drivers, analysis identified differences for several pairs of factors. These are explained by the different types of operation, age, and operational experience. Firstly, it was found that the responses of train drivers are significantly different with respect to the influence of team factors on task factors (H(2)=6.383, p<.05). In particular, results indicate that the train drivers belonging to one of the three TOCs indicated significantly higher scores than the train drivers of the remainder two TOCs. Thus, further research has been recommended to the relevant TOC to investigate whether their working groups operate in an efficient manner.

On experience, differences in the assessments were found with respect to the influence of personal factors on task factors (H(3)=8.103, p<.05). However, both the Mann-Whitney and Kolmogorov-Smirnov Z tests failed to identify the origin of this difference. On the age of train drivers, the identified differences reflect the assessments of the organisational (H(3)=12.882, p<.05) and personal factors (H(3)=13.898, p<.05) on task factors. For the impact of organisational factors on task factors, *post hoc* tests showed that train drivers between 35-44 years old assessed such an impact with significant higher values compared to the drivers between 24-34 years old (U=1, r=-.807). In addition, 35-44 year old drivers assigned significantly higher scores to the influence of organisational factors on task factors compared to the drivers between 45-56 years old (U=27, r=-.493). On the other hand, for the same pairs of age groups, the opposite findings were identified for the influence of personal factors on task factors. Firstly, the younger train drivers evaluated the influence of organisational factors on task factors with lower scores compared to their elder colleagues (U=3, r=-.754), while again the younger operators assigned significantly lower scores to the impact of personal factors on task factors compared to the elder drivers (U=16, r=-.627).

### 8.2.6 Influence of the R-PSFs categories on system factors

This section illustrates the findings regarding the influence of the R-PSFs on system factors. Firstly, Table 8-9 indicates significant differences in the responses of the SMEs subgroups only with respect to the importance of task factors on system factors. *Post hoc* tests show that train drivers evaluate such an importance with significantly lower scores, compared to operations managers (U=142, r=-.325).

Table 8-9 Effect of R-PSFs categories on Task Factors for the SMEs subgroups

| | Dynamic personal | Organisational | Personal | System | Task | Team |
|---|---|---|---|---|---|---|
| H | .577 | 2.837 | .394 | .000 | 7.128 | .583 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | p>.05 | p>.05 | p>.05 | 1.000 | p<.05 | p>.05 |

For train drivers, differences were identified with regard to how drivers assess the influence of task factors on system factors for the different types of operations (H(2)=6.192, p<.05). Moreover, differences were also found on the assessments of drivers with different levels of operational experience regarding the impact of dynamic personal factors (H(3)=15.828, p<.05) and task factors (H(3)=9.993, p<.05) on system factors. Specifically, it was identified that train drivers with 5-10 years of experience consider the influence of dynamic personal factors on system factors as significantly lower compared to the least experienced personnel

(U=.00, r=-.88) and those with 11-19 years of experience (U=1, r=-.807). On the other hand, results indicate also that the least experienced personnel compared to those with 5-10 and 20 or more years of operational experience, consider the influence of task factors on system factors as significantly lower, with (U= 3, r=-.734) and (U=5, r=-.717), respectively.

Finally, the assessments of the impact of task factors on system factors are also affected by the train drivers' previous SPAD experiences. It was found that drivers with no SPAD experience assigned significantly lower values to the impact of task factors on system factors, compared to those who were involved in at least one SPAD in the past (U=67, z=-2.118, p<.05, r=-.358).

The previous sections presented the results on the weighting of the R-PSFs categories, as derived from the SMEs assessments. In general, the findings indicate limited significant differences in the responses amongst the SMEs, primarily for the dynamic personal and organisational factors. This could be explained by the different backgrounds of the SMEs, i.e. duties, responsibilities and positions, of the SMEs. Therefore, regardless of the fact that the sizes of the SME subgroups vary substantially, and such generalisations of the findings require caution, it can be seen that the perspective of the SMEs of the influence of the R-PSFs on human performance can be influenced by the characteristics of the SMEs. Moreover, for train drivers in particular, it is apparent that their responses are, to a certain extent, explained by their individual characteristics, which are influenced primarily by the type of operation and their operational experience.

The results also showed that all the SMEs acknowledge the significant influence of the dynamic personal factors on all R-PSFs categories. Furthermore, it was found that the dynamic personal factors and personal factors display similarities with respect to how they are influenced by the other R-PSFs categories. Similarities are also observed for the categories of organisational and team factors, as well as task and system factors. In addition, review can be undertaken and the similarities and differences between the responses of the SMEs with respect to their weightings on the R-PSFs elements identified.

The following section presents the final weightings for the R-PSFs elements with respect to their influence on human performance as derived from the limit supermatrices (see Section 7.5.7.3 and 7.5.7.4).

## 8.3 The final priorities of the R-PSFs elements

This section demonstrates the final results of the R-PSFs elements weightings, as obtained from the corresponding *limit* supermatrices (Section 7.5.7.4). These results present the contribution of each factor to operator performance, where the larger the weight value, the more important the R-PSF is in successfully completing a task. The normality of data was first tested for using the Shapiro-Wilk test. The results (Table 8-10) show that nine variables are not normally distributed. For these, the Kruskal-Wallis and when necessary the Mann-Whitney tests, were used to test whether the responses could be assumed to belong to the same population.

Table 8-10 Descriptive statistics for the R-PSFs *lite* elements weighting

| Variable | N* | Mean | SE | SD | 95% confidence interval | | Min | Max | Normality** |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | | |
| Distraction | 52 | **0.0864** | 0.0067 | 0.0486 | 0.0728 | 0.0999 | 0.0142 | 0.2102 | ✓ |
| Fatigue | 52 | **0.0958** | 0.0066 | 0.0477 | 0.0825 | 0.1091 | 0.0225 | 0.1929 | ✗ |
| Perception | 52 | **0.0425** | 0.0034 | 0.0247 | 0.0356 | 0.0494 | 0.0061 | 0.1165 | ✗ |
| Procedures | 52 | **0.0704** | 0.0041 | 0.0296 | 0.0621 | 0.0786 | 0.0189 | 0.1468 | ✓ |
| Safety culture | 52 | **0.0794** | 0.0053 | 0.0387 | 0.0686 | 0.0902 | 0.0205 | 0.1781 | ✗ |
| Supervision | 52 | **0.0597** | 0.0048 | 0.0349 | 0.0500 | 0.0695 | 0.0122 | 0.1949 | ✗ |
| Familiarity | 52 | **0.1124** | 0.0069 | 0.0497 | 0.0986 | 0.1263 | 0.0271 | 0.2789 | ✗ |
| Training | 52 | **0.0982** | 0.0058 | 0.0421 | 0.0865 | 0.1099 | 0.0206 | 0.2369 | ✗ |
| System design | 52 | **0.1074** | 0.0054 | 0.0388 | 0.0966 | 0.1182 | 0.0327 | 0.2102 | ✓ |
| Workload | 52 | **0.1264** | 0.0084 | 0.0607 | 0.1095 | 0.1433 | 0.0364 | 0.3044 | ✗ |
| Communication | 52 | **0.0627** | 0.0038 | 0.0277 | 0.0550 | 0.0705 | 0.0224 | 0.1257 | ✗ |
| Information | 52 | **0.0582** | 0.0041 | 0.0295 | 0.4998 | 0.0664 | 0.0140 | 0.1159 | ✗ |

*the mean, SE and SD when HFs responses are excluded change after the 3rd decimal
**normality at 95% level of confidence

For the remaining three normally distributed variables, the independent t-test and analysis of variance (ANOVA) were applied. The former is used to test the null hypothesis whereby the means collected from two different groups are equal. The t-statistics is expressed as:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \qquad (8\text{-}2)$$

where $\overline{X}_1$ and $\overline{X}_2$ are the means of the samples 1 and 2 of sizes $n_1$ and $n_2$ respectively, and $s_p$ is the pooled variance, i.e. the variance weighted by the size of sample on which it is based (Field, 2009). If the t-statistic is significant, $p<.05$, then the null hypothesis is rejected and the means of the two groups are considered significantly different. The ANOVA, on the other hand, is used to test whether the means of more than two independent groups are significantly different. ANOVA is developed on the F-statistic, which is a measure of the ratio

of the variations explained by the model and unsystematic (extraneous) factors (Field, 2009). It is computed by dividing the model mean squares ($MS_M$) by the residual mean squares ($MS_R$). Again, if the F-statistic is significant then the null hypothesis is rejected and the means of the groups are considered significantly different. Then, *post hoc* tests (using the Bonferroni correction) were applied to compare the different combinations of the groups and to identify which particular means significantly differ (Field, 2009, pp.373-375).

The findings were tested for outliers and when such cases were identified, two datasets were created. However, as no significant difference was identified, the following sections present the results for the complete dataset. For brevity, only the first component of the 12 R-PSFs elements is given. The main relevant means for the classification groups are given in the Table 8-11.

Table 8-11 The final R-PSFs weightings means per classification group

| Final weightings (means) of R-PSFs elements | | Train Drivers | Operation managers | HFs experts | Train Drivers TOC 1 | Train Drivers TOC 2 | Train Drivers TOC 3 |
|---|---|---|---|---|---|---|---|
| Dynamic Personal Factors | Distraction | .0770 | .1163 | .0569 | .0880 | .0532 | .0879 |
| | Fatigue | .0926 | .1117 | .0587 | .1023 | .0737 | .1000 |
| | Perception | .0376 | .0571 | .0420 | .0488 | .0253 | .0354 |
| Organisational Factors | Procedures | .0730 | .0617 | .0805 | .0530 | .0856 | .0872 |
| | Safety culture | .0857 | .0566 | .1119 | .0657 | .0928 | .1060 |
| | Supervision | .0617 | .0517 | .0753 | .0481 | .0601 | .0826 |
| Personal Factors | Familiarity | .1157 | .1060 | .1052 | .1475 | .1017 | .0866 |
| | Training | .1021 | .0874 | .1050 | .1268 | .0890 | .0817 |
| System Factors | System design | .1131 | .0969 | .0905 | .1097 | .1197 | .1106 |
| Task Factors | Workload | .1187 | .1479 | .1168 | .1128 | .1495 | .0932 |
| Team Factors | Communication | .0632 | .0566 | .0869 | .0503 | .0753 | .0679 |
| | Information | .0595 | .0502 | .0803 | .0470 | .0742 | .0609 |

### 8.3.1 Distraction – Loss of concentration

For distraction, as shown in the Table 8-12 a significant difference was found amongst the responses of the different types of the SMEs. In particular*, post hoc* tests provide evidence that the difference is caused by the weightings of train drivers and HFs experts compared to operations managers. The later assigned higher scores compared to the former.

For train drivers, the mean weightings were tested against the type of operation, gender, age, operational experience and previous SPAD experience. However, no significant differences were identified for any subgroups. However, it is important to note the fact that

the train drivers of TOC1 and TOC3 assessed the importance of distraction on their performance with higher values compared to those of TOC2. This finding could be used to investigate any similarities and differences of the services, their networks and operational features.

Table 8-12 Effect of classification groups on weighting distraction

| Distraction | F or t-statistic* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 4.354 | 2 | **p<.05** |
| Type of operation (train drivers only) | 2.122 | 2 | p>.05 |
| Gender (train drivers only) | -0.400 | 33 | p>.05 |
| Age (train drivers only) | 0.424 | 3 | p>.05 |
| Operational experience (train drivers only) | 1.007 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 0.983 | 33 | p>.05 |

*t-statistic for binary variables

### 8.3.2 Fatigue

For fatigue, the Kruskal-Wallis tests in Table 8-13, did not indicate any significant difference between the mean weightings determined from the responses of the SMEs. However, at first glance the values as shown in the Table 8-11 may imply such a difference. Thus, Mann-Whitney and Kolmogorov-Smirnov Z (KS-Z) tests were performed for each pair of SMEs. While the Mann-Whitney test confirms the initial findings, the KS-Z indicates a significant difference between the weightings of the operations managers and HFs experts (z=-1.572, p<.05), with the former assigning higher scores to the contribution of fatigue.

Table 8-13 Effect of classification groups on weighting fatigue

| Fatigue | H or U* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 4.561 | 2 | p>.05 |
| Type of operation (train drivers only) | 2.130 | 2 | p>.05 |
| Gender (train drivers only) | 36 (z=-.707) | - | p>.05 |
| Age (train drivers only) | 0.448 | 3 | p>.05 |
| Operational experience (train drivers only) | 8.989 | 3 | **p<.05** |
| Involvement in SPAD (train drivers only) | 122 (z=-110) | - | p>.05 |

*U statistic for binary variables

Comparison of the weightings derived from the train drivers was also performed. Firstly, the responses were analysed against the different types of operations. However, the results did not provide evidence of significant differences. Furthermore, it was found that neither gender, age, nor previous SPAD experience has a significant effect on the assessments of train drivers. On the other hand, a significant difference was revealed with respect to the train drivers' experience. In particular, *post hoc* tests indicate that train drivers with 11-19

years of experience consider that fatigue affects the performance of the operators, more than the operators with 5-10 of experience (U=6, z=-2.699, p<.05, r=-674) and those with more than 20 years of experience (U=25, z=-2.061, p<.05, r=-.449).

### 8.3.3 Perception

In terms of perception results in Table 8-14 indicate a significant difference only between the weightings as derived from the different types of SMEs. Specifically, operations managers weighted the impact of perception on the performance with higher scores compared to train drivers (U=111, r=-.424).

In contrast, no significant differences were identified for any of the remaining control groups. Furthermore, it is also clear that general perception, as shown in Table 8-10, is not considered from any of the SMEs groups to have a high impact on the performance of the operators.

Table 8-14 Effect of classification groups on weighting perception

| Perception | H or U* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 9.234 | 2 | **p<.05** |
| Type of operation (train drivers only) | 5.543 | 2 | p>.05 |
| Gender (train drivers only) | 45 (z=-177) | - | p>.05 |
| Age (train drivers only) | 2.326 | 3 | p>.05 |
| Operational experience (train drivers only) | 2.301 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 91 (z=-1.242) | - | p>.05 |

*U statistic for binary variables

### 8.3.4 Quality of procedures

Regarding the weightings of procedure quality results in Table 8-15 do not reveal a significant difference between the responses of the different types of the SMEs. In contrast, a significant difference was identified between the responses of train drivers, when compared against the type of operation. It should be noted that for the organisational R-PSFs, the type of operation, reflects the railway organisation. Subsequently, the difference in weightings represents the differences between the TOCs. In particular, results show that the train drivers of two TOCs assess the quality of procedures with higher scores compared to the third TOC. Such a finding may imply that those train drivers have faced SPAD occurrences in the past, whilst the necessary written or oral procedures were not clear or available. However, it may also imply that those train drivers are more aware of the important role of the procedures in safe and undisturbed operations. Therefore, further

investigation is recommended on the role of procedures on operator performance and how the latter perceive such a role.

Table 8-15 Effect of classification groups on weighting of quality procedures

| Quality of procedures | F or t-statistic* | df | Sig. |
|---|---|---|---|
| Type of employees | 0.913 | 2 | p>.05 |
| Type of operation (train drivers only) | 6.417 | 2 | **p<.05** |
| Gender (train drivers only) | 0.333 | 33 | p>.05 |
| Age (train drivers only) | 1.234 | 3 | p>.05 |
| Operational experience (train drivers only) | 1.165 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | -0.238 | 33 | p>.05 |

*t-statistic for binary variables

### 8.3.5 Safety culture

With respect to safety culture, findings, as shown in Table 8-16, show significant differences amongst the responses of the different types of SMEs. In particular, Mann-Whitney tests revealed that train drivers (U=122, r=-.388) and HFs experts (U=4, r=-.519) consider the contribution of the safety culture on the performance of operators as more important, compared to the judgments derived from the operations managers.

Table 8-16 Effect of classification groups on weighting safety culture

| Safety culture | H or U* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 10.001 | 2 | **p<.05** |
| Type of operation (train drivers only) | 8.393 | 2 | **p<.05** |
| Gender (train drivers only) | 35 (z=-.766) | | p>.05 |
| Age (train drivers only) | 3.863 | 3 | p>.05 |
| Operational experience (train drivers only) | 4.038 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 95 (z=-1.095) | | p>.05 |

*U statistic for binary variables

The survey revealed two main reasons that explain these results. The first, justifies the judgements of the HFs experts, which due to their experience and background support the importance of safety culture on all organisations activities. The substantial role of safety culture on the operations was discussed in chapter 5, where it was shown that organisations with positive maturity scores have less occurrences that lead to severe consequences; while it is also widely addressed in the literature, e.g. (Guldenmund, 2000, Sorensen, 2002, Choudhry et al., 2007, Luther and Johnson, 2008). On the other hand, it was found that the operations managers do not consider safety culture as major contributor factor to SPAD occurrences. Although they acknowledge its importance on the safety of services, they highlight instead the significant role of personal (including the dynamic) factors.

For train drivers, significant differences in the data were identified only with respect to the different types of operations. Operators of TOC1 assessed the influence of safety culture on their performance with lower scores compared to those of TOC2 (U=33, r=-.482) and TOC3 (U=29, r=-.49). Comparing the results, discussions within the organisations are currently being held to identify whether such responses are explained by any previous occurrences, where the safety culture of organisations significantly contributed to the occurrence.

### 8.3.6 Supervision

Regarding supervision, results in Table 8-17 do not indicate any significant differences in the responses either between the types of SMEs or the types of train drivers. Considering its relatively low mean scores (Table 8-10) it could be claimed that the SMEs do not consider supervision as important performance factor within the context of the described scenario.

Table 8-17 Effect of classification groups on weighting supervision

| Supervision | H or U* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 2.892 | 2 | p>.05 |
| Type of operation (train drivers only) | 5.453 | 2 | p>.05 |
| Gender (train drivers only) | 35 (z=-.766) | | p>.05 |
| Age (train drivers only) | 4.933 | 3 | p>.05 |
| Operational experience (train drivers only) | 2.661 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 95 (z=-1.095) | | p>.05 |

*U statistic for binary variables

### 8.3.7 Familiarity

Considering familiarity, it was found, as shown in Table 8-18, that the characteristics of the SMEs do not significantly affect their judgements. Moreover, given the high scores of familiarity it is reasonable to claim that all the experts consider familiarity as a substantial factor for SPAD occurrences. The impact of familiarity on the performance of the operators can be expressed in two ways. Firstly, high levels of familiarity may reduce the vigilance of the operators and subsequently increase the likelihood of a SPAD. Secondly, low levels of familiarity may indicate employees with poor knowledge of a route. Hence, the likelihood of a passed signal could increase, mainly within complex layouts where operators may not be aware of the exact signals' location.

In addition, for the train drivers in particular, differences in the weightings were found for the classification groups' type of operation and operational experience. For the former, however, *post hoc* tests failed to indicate the origin of such differences. On the other hand, the results

for operational experience showed that the least experienced personnel assessed the influence of familiarity with higher scores compared to the more experienced personnel, i.e. 11-19 (U=12, r=-.516) and 20 or more (U=17, r=.-512) years of experience.

Table 8-18 Effect of classification groups on weighting familiarity

| Familiarity | H or U* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 0.124 | 2 | p>.05 |
| Type of operation (train drivers only) | 6.197 | 2 | **p<.05** |
| Gender (train drivers only) | 28 (z=-1.179) | | p>.05 |
| Age (train drivers only) | 5.660 | 3 | p>.05 |
| Operational experience (train drivers only) | 6.272 | 3 | **p<.05** |
| Involvement in SPAD (train drivers only) | 77 (z=-1.753) | | p>.05 |

*U statistic for binary variables

### 8.3.8 Training

Similar to familiarity, results in Table 8-19 for training provide evidence that the types of the SMEs do not have an impact on their responses. Moreover, results for the train drivers indicate that their responses differ significantly when compared with the types of operation (which reflects also the TOC) and operational experience. However, for the former *post hoc* tests again fail to indicate the origin of the difference. For the latter, it was found that the least experienced personnel consider the impact of training on human performance more important compared to the more experienced personnel, i.e.11-19 years of experience (U=8, r=-.622).

Table 8-19 Effect of classification groups on weighting training

| Training | H or U* | df | Sig. |
|---|---|---|---|
| Type of SMEs | 0.124 | 2 | p>.05 |
| Type of operation (train drivers only) | 6.233 | 2 | p<.05 |
| Gender (train drivers only) | 25 (z=-1.355) | | p>.05 |
| Age (train drivers only) | 5.700 | 3 | p>.05 |
| Operational experience (train drivers only) | 9.580 | 3 | **p<.05** |
| Involvement in SPAD (train drivers only) | 85 (z=-1.461) | | p>.05 |

*U statistic for binary variables

The similarities in the results for familiarity and training, as expressed and underpinned by the level of experience of the operators, confirm the strong dependency between the two variables.

### 8.3.9 System design

Similar to the majority of the previous R-PSFs, the characteristics of the SMEs were not found to affect the weightings of system design, as results in Table 8-20 indicate. In addition, due to the relatively high scores, it is apparent that all the SMEs acknowledge the important role of system design. This is primarily expressed by the location of the signals, on the prevention of SPADs.

With respect to train drivers, it was identified that the type of operation does not affect their assessments. However, a significant difference was found between the female and male train drivers. In particular, female employees weighted the impact of system design with higher scores compared to their male colleagues. Finally, a significant difference was also identified for the different levels of operational experience. Results show that train drivers with 11-19 years of experience assessed the importance of system design considerably less than all the remaining employees.

Table 8-20 Effect of classification groups on weighting system design

| System design | F or t-statistic | df | Sig. |
|---|---|---|---|
| Type of employees | 1.177 | 2 | p>.05 |
| Type of operation (train drivers only) | 0.196 | 2 | p>.05 |
| Gender (train drivers only) | -2.328 | 33 | p<.05 |
| Age (train drivers only) | 0.285 | 3 | p>.05 |
| Operational experience (train drivers only) | 3.065 | 3 | p<.05 |
| Involvement in SPAD (train drivers only) | -0.174 | 33 | p>.05 |

*t-statistic for binary variables

### 8.3.10 Workload

With respect to workload, statistics in Table 8-21 show that the profession of the SMEs does not have a significant effect on their responses. It is notable that for all the SMEs, workload is considered as the most significant contributing factor for the SPAD scenario. Moreover, train drivers with at least one SPAD experience account for the impact of workload on their performance considerably higher compared to those with no SPAD records. Finally, it was also found that the responses between train drivers from the different TOCs are not significantly different. Nevertheless, it should be noted that the operators of the TOC2 assigned higher workload scores, which probably reflects the type of service.

Table 8-21 Effect of classification groups on weighting workload

| Workload | H or U* | df | Sig. |
|---|---|---|---|
| Type of employees | 4.345 | 2 | p>.05 |
| Type of operation (train drivers only) | 3.704 | 2 | p>.05 |
| Gender (train drivers only) | 41 (z=-.412) | | p>.05 |
| Age (train drivers only) | 3.865 | 3 | p>.05 |
| Operational experience (train drivers only) | 5.313 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 61 (z=-2.337) | | **p<.05** |

*U statistic for binary variables

### 8.3.11 Communication between employees

For the influence of communication on the performance of operators, results in Table 8-22 show strong evidence that the responses do not significantly differ for any classification group.

Table 8-22 Effect of classification groups on weighting communication

| Communication | H or U* | df | Sig. |
|---|---|---|---|
| Type of employees | 2.829 | 2 | p>.05 |
| Type of operation (train drivers only) | 2.976 | 2 | p>.05 |
| Gender (train drivers only) | 37 (z=-.648) | | p>.05 |
| Age (train drivers only) | 1.775 | 3 | p>.05 |
| Operational experience (train drivers only) | 5.199 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 115 (z=-.365) | | p>.05 |

*U statistic for binary variables

### 8.3.12 Quality of information

Finally, regarding the weightings for information quality, results in Table 8-23 indicate no significant difference for any classification group.

Table 8-23 Effect of classification groups on weighting quality of information

| Information | H or U* | df | Sig. |
|---|---|---|---|
| Type of employees | 2.550 | 2 | p>.05 |
| Type of operation (train drivers only) | 4.163 | 2 | p>.05 |
| Gender (train drivers only) | 24 (z=-1.414) | | p>.05 |
| Age (train drivers only) | 0.937 | 3 | p>.05 |
| Operational experience (train drivers only) | 4.301 | 3 | p>.05 |
| Involvement in SPAD (train drivers only) | 122 (z=-.110) | | p>.05 |

*U statistic for binary variables

### 8.3.13 Review of the R-PSFs elements weightings

In the previous subsections the results for the weighting of the R-PSFs elements were presented. The final findings indicate that the duties of the SMEs do not significantly affect their judgements for the majority of the R-PSFs. However, as illustrated in the Figure 8-1, this is not the case for the R-PSFs distraction, perception and safety culture. This is in agreement with the weightings of the R-PSFs categories, where significant differences were found between the SMEs responses for the dynamic personal and organisational factors. Whereas such a difference could be ignored for perception as its overall assigned scores are relatively low, this is not recommended for the remaining factors.

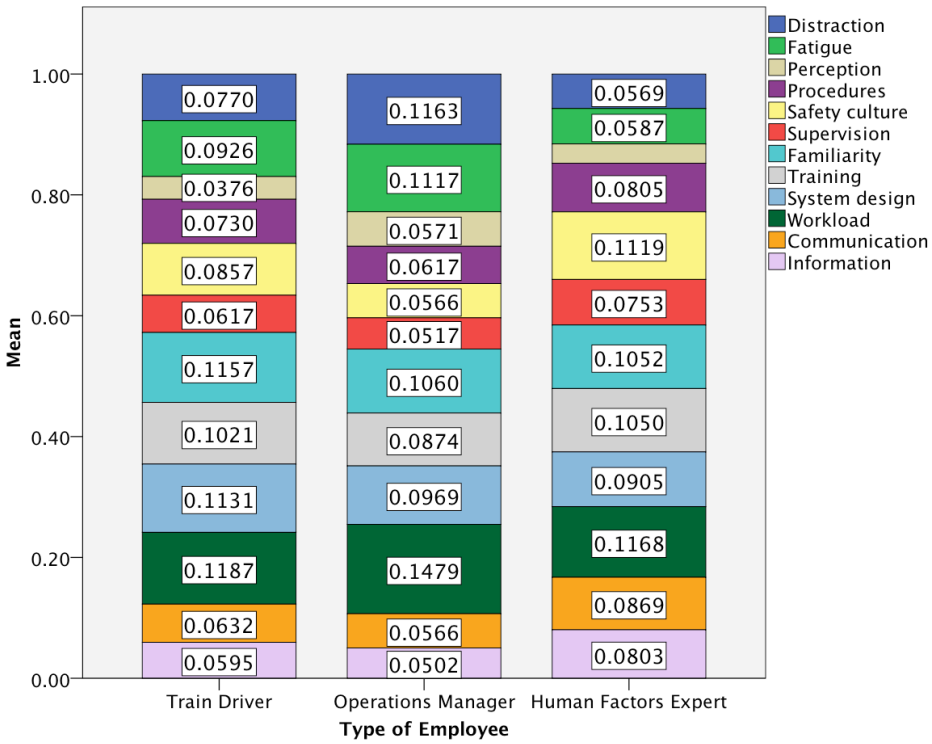



Figure 8-1 Comparison of R-PSFs elements weighting by SMEs

It was also found that workload is considered the most important R-PSF with respect to its influence on human performance for all the types of SMEs. In addition, familiarity was assessed a substantial contributor to operators performance. As mentioned in the Section 8.3.7 the less experienced personnel assigned higher scores to familiarity. This implies that operators unfamiliar with a route are likely to miss a signal, as they may not be aware of its exact location or function. However, this contradicts the opinion of those who believe that a high level of familiarity increases an operator's level of expectation, in particular with respect to the signal aspects ahead. Subsequently, if a signal does not indicate the usual aspect, it is often the case that the train drivers do not have sufficient time to stop the train. In addition, system design is considered a substantial factor, which primarily reflects on the location of

the signals. This is expressed by the ability of train drivers to identify and acknowledge their aspects.

In turn, with respect to the dynamic personal factors, fatigue is deemed to be the most influential factor on human performance. In addition, distraction is only considered to be highly important by the operations managers. This could be explained by the fact that train drivers may be reluctant to admit that they have been distracted and therefore passed a signal at danger. In addition, it is worth commenting on the assessments of HFs experts with respect to distraction as well as the dynamic personal factors. It can be seen that the HFs experts account the impact of the three dynamic personal factors on human performance relatively low compared to the other two types of SMEs. This is also in agreement with the assessment of the SMEs for the R-PSFs categories. Such an observation is related to the main belief of the HFs experts that the organisational factors are those that mostly affect the performance of the operators. Therefore, although the dynamic personal factors are not underestimated, the organisational factors are considered to be more important. It is also assumed, that if the organisational factors were improved, then the consequences of the factors related directly to frontline personnel would be mitigated. On the other hand, operations managers assessed organisational factors with relatively low weightings. This may suggest that they believe that their institutions provide all the necessary support to their personnel to avoid SPAD occurrences.

Finally, it should be also stated that for the analysis of the drivers' assessments, any substantial differences among the operational system should not be neglected. For instance, it cannot not be ignored that while the London Underground services are equipped with the ATP protection system, the mainline operations are fitted with the TPWS. The results in this study do not provide any evidence that such differences affected the way that the train drivers have weighted the different R-PSFs. However, for any future studies researchers should be aware and take into consideration any systemic differences in the interpretation of their results.

These results were used to identify whether the SMEs assessments are in agreement with the factors that in reality have led to past SPAD occurrences. In other words, it was investigated whether the perception of the SMEs regarding the impact of the factors that affect the performance of the operators corresponds with the findings of the investigation reports. To do so, the SMEs judgments were compared with data extracted from 58 UK SPAD investigations reports, which were included in the 479 reports dataset described in chapter 6. The frequency of the identified R-PSFs is shown in the Figure 8-2. The

comparison illustrates only the similarities and differences for the different types of the SMEs.

For example, it can be seen that although all the SMEs have acknowledged workload to be the most important contributing factor with respect to the described scenario, this is not congruent with the information extracted from the reports. It should also be noted that although the train drivers and operations managers assess fatigue as a highly important factor, the investigation reports do not confirm this. Nevertheless, the comparison shows that for familiarity, training and system design, the SMEs judgements and the reports' findings are generally in agreement.

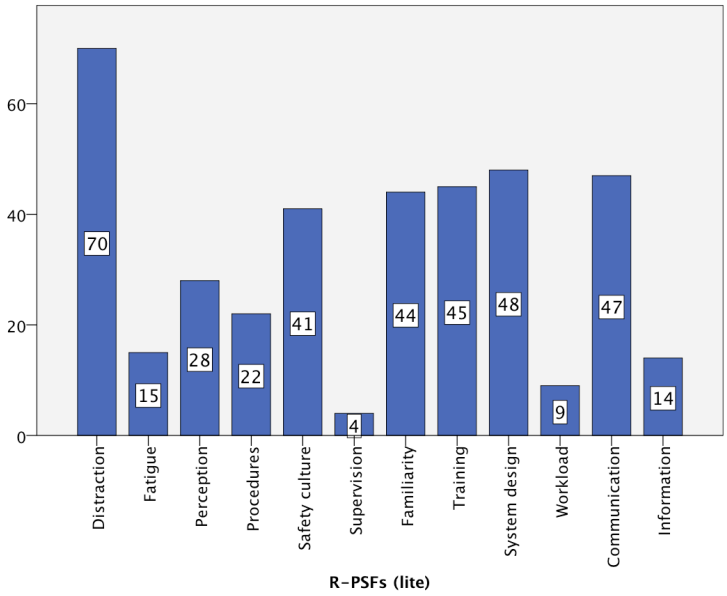



Figure 8-2 Frequency of the R-PSFs elements in the UK SPAD occurrences

Thus, the identified differences indicate that the SMEs have an established perspective with respect to the influence of particulars factors on the performance of the operators. However, such a perspective does not always reflect the actual contribution of the factors. Therefore, a series of workshops is recommended with the different railway stakeholders, primarily the frontline and managerial personnel. Such workshops will aim to highlight the differences and to further inform operators for the actual role that the R-PSFs play in the occurrence of SPADs, as well as other types of occurrences.

To this point, the weightings of the R-PSFs categories and elements with respect to their contribution to the performance of railway operators were thoroughly presented. It is recommended that the results should be interpreted and used only for the specific scenario. Nevertheless, it can be claimed that both the framework and methodology could be broadly applied to several operational scenarios. Hence, support is provided to railway organisations

to gain a more comprehensive perspective on the way that their employees perceive the impact of the R-PSFs on human performance. Subsequently, resources could be used more efficiently towards the development of sound solutions to mitigate both the appearance of such factors and their consequences.

As described in the previous chapter, further to the weighting of the R-PSFS, the HuPeROI aims also to indicate the relative likelihoods of different erroneous actions. The next section presents the ratings of such factors for three human errors that could lead to a SPAD.

## 8.4 The ratings of the R-PSFs elements

This section reviews the results from the elicitation process of the R-PSFs ratings. As a reminder, the rating represents how good or bad the R-PSFs are considered for the action under assessment, where good or bad mean that the factors either enhance or degrade the ability of the operator to conduct the action.

### 8.4.1 The SMEs ratings for "signal detection"

The first operational failure is the inability of the operator to appropriately perform the action of signal detection. The ratings of the R-PSFs are non-parametric data. Therefore, Kruskal-Wallis and Mann-Whitney tests are applied to indicate whether the SMEs judgments were affected by the characteristics of the classification groups. As can be seen in Table 8-24 no significant differences were identified between the ratings of the different SMEs. Therefore, it could be stated that the characteristics of the SMEs do not have an impact on the ratings of the R-PSFs.

Table 8-24 Effect of classification groups on the R-PSFs rating for "signal not detected"

| R-PSFs | Type of SMEs | | | Type of operations (train drivers) | | | Gender (train drivers) | | |
|---|---|---|---|---|---|---|---|---|---|
| | H | df | Sig. | H | df | Sig. | U | r | Sig. |
| Distraction | 1.911 | 2 | p>.05 | 0.081 | 2 | p>.05 | 44.5 | -0.035 | p>.05 |
| Fatigue | 0.381 | 2 | p>.05 | 0.004 | 2 | p>.05 | 42.0 | -0.061 | p>.05 |
| Perception | 5.855 | 2 | p>.05 | 0.837 | 2 | p>.05 | 43.5 | -0.045 | p>.05 |
| Procedures | 0.044 | 2 | p>.05 | 0.432 | 2 | p>.05 | 25.5 | -0.045 | p>.05 |
| Safety culture | 1.579 | 2 | p>.05 | 1.948 | 2 | p>.05 | 41.0 | -0.226 | p>.05 |
| Supervision | 0.843 | 2 | p>.05 | 0.651 | 2 | p>.05 | 35.5 | -0.070 | p>.05 |
| Familiarity | 0.236 | 2 | p>.05 | 5.368 | 2 | p>.05 | 47.0 | -0.126 | p>.05 |
| Training | 1.608 | 2 | p>.05 | 0.145 | 2 | p>.05 | 44.5 | -0.035 | p>.05 |
| System design | 4.519 | 2 | p>.05 | 2.357 | 2 | p>.05 | 35.5 | -0.190 | p>.05 |
| Workload | 0.102 | 2 | p>.05 | 4.059 | 2 | p>.05 | 42.5 | -0.126 | p>.05 |
| Communication | 1.894 | 2 | p>.05 | 0.227 | 2 | p>.05 | 44.5 | -0.055 | p>.05 |
| Information | 1.609 | 2 | p>.05 | 0.074 | 2 | p>.05 | 39.5 | -0.035 | p>.05 |

Table 8-24 (cont.) Effect of classification groups on the R-PSFs rating for "signal not detected"

| R-PSFs | Age (train drivers) | | | Ops. Experience (train drivers) | | | Involvement in SPAD (train drivers) | | |
|---|---|---|---|---|---|---|---|---|---|
| | H | df | df | H | df | Sig. | U | r | Sig. |
| Distraction | 4.997 | 3 | p>.05 | 4.037 | 3 | p>.05 | 122.5 | -0.016 | p>.05 |
| Fatigue | 5.997 | 3 | p>.05 | 3.110 | 3 | p>.05 | 125.0 | 0.000 | p>.05 |
| Perception | 6.333 | 3 | p>.05 | 0.936 | 3 | p>.05 | 123.5 | -0.009 | p>.05 |
| Procedures | 1.411 | 3 | p>.05 | 2.896 | 3 | p>.05 | 115.5 | -0.059 | p>.05 |
| Safety culture | 2.277 | 3 | p>.05 | 3.542 | 3 | p>.05 | 120.5 | -0.028 | p>.05 |
| Supervision | 1.897 | 3 | p>.05 | 0.443 | 3 | p>.05 | 87.0 | -0.237 | p>.05 |
| Familiarity | 10.132 | 3 | **p<.05** | 8.818 | 3 | **p<.05** | 85.0 | -0.250 | p>.05 |
| Training | 2.357 | 3 | p>.05 | 1.895 | 3 | p>.05 | 92.5 | -0.205 | p>.05 |
| System design | 1.998 | 3 | p>.05 | 0.161 | 3 | p>.05 | 97.0 | -0.175 | p>.05 |
| Workload | 3.352 | 3 | p>.05 | 0.979 | 3 | p>.05 | 87.5 | -0.233 | p>.05 |
| Communication | 5.287 | 3 | p>.05 | 4.221 | 3 | p>.05 | 121.5 | -0.022 | p>.05 |
| Information | 6.592 | 3 | p>.05 | 5.988 | 3 | p>.05 | 117 | -0.050 | p>.05 |

A comparison of the ratings, as obtained from the train drivers, was also performed. Results indicate significant differences only for familiarity. This difference is established for the classification groups, age and operational experience. For the former, *post hoc* tests show that train drivers between 24-34 years old assigned to familiarity lower scores compared to the drivers between 45-56 years old ($U=9$, $z=-3.091$, $p<.05$, $r=-.659$). For operational experience, the difference is originated from the responses of the least experienced train drivers and those with more than 20 years of experience. The latter rated familiarity with higher scores ($U=12$, $z=-2.527$, $p<.05$, $r=-.579$).

However, since ratings of the majority of classification groups are not significantly different, data are combined. Figure 8-3 (left) illustrates the final results. In addition, Figure 8-3 (right) displays the ratings as given by the three different SMEs subgroups. At first glance, it can be seen that all the SMEs consider distraction, fatigue and perception as the most suboptimal factors. They also consider that the training of the SMEs was above average. Moreover, the majority of the R-SPFs have been rated poor, with the scores to lie in the 40-50 range. Subsequently, total ratings are also low, i.e. do not exceed the value of 550.

Finally, it can be seen that the HFs experts, compared to train drivers and operations managers, rated the factor system design as highly suboptimal for the specific error. This is explained by the fact that the HFs experts have analysed many SPADs where the location or sighting of the signal was a leading contributor to the event.

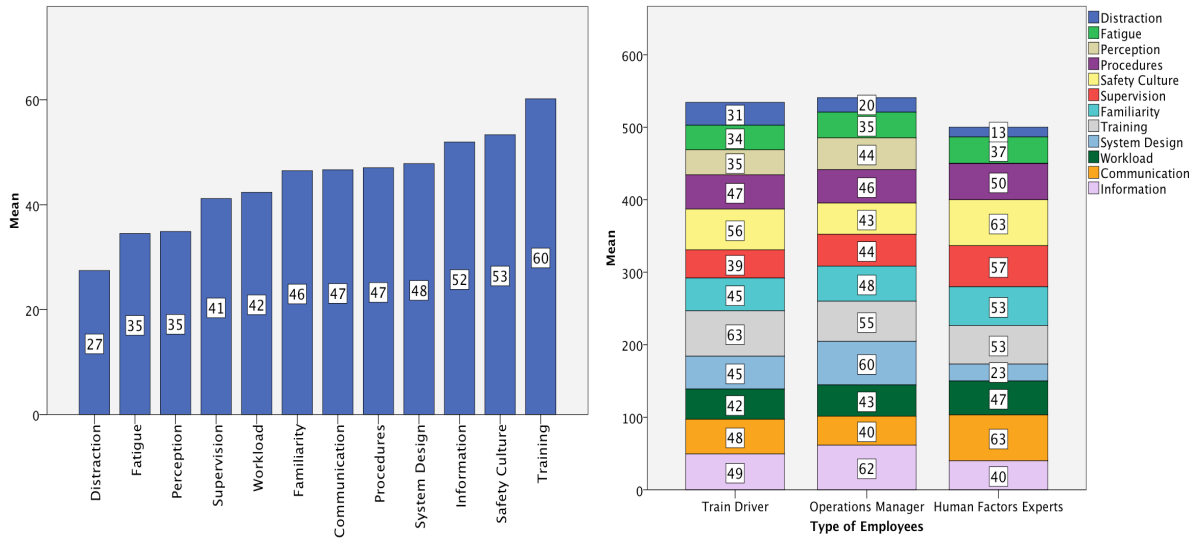



Figure 8-3 R-PSFs ratings for signal not detected

### 8.4.2 The SMEs ratings for "signal interpretation"

The second identified human error is the operators' failure to interpret the aspect of the signal.. Results (Table 8-25), indicate no significant differences for the ratings of most of the R-PSFs. In fact, the only significant difference was identified for the contributing factor system design. In particular, it was found that male train drivers assigned higher values compared to the female drivers. Nevertheless, no profound reason to explain this result was identified.

Table 8-25 Effect of classification groups on the R-PSFs rating for "signal wrongly interpreted"

| R-PSFs | Type of SMEs | | | Type of operations (train drivers) | | | Gender (train drivers) | | |
|---|---|---|---|---|---|---|---|---|---|
| | H | df | Sig. | H | df | Sig. | U | r | Sig. |
| Distraction | 0.124 | 2 | p>.05 | 0.214 | 2 | p>.05 | 36.5 | -0.115 | p>.05 |
| Fatigue | 2.033 | 2 | p>.05 | 1.842 | 2 | p>.05 | 47.5 | -0.005 | p>.05 |
| Perception | 3.234 | 2 | p>.05 | 5.702 | 2 | p>.05 | 30.0 | -0.184 | p>.05 |
| Procedures | 0.171 | 2 | p>.05 | 0.490 | 2 | p>.05 | 43.0 | -0.050 | p>.05 |
| Safety culture | 0.434 | 2 | p>.05 | 1.194 | 2 | p>.05 | 43.0 | -0.050 | p>.05 |
| Supervision | 0.768 | 2 | p>.05 | 2.109 | 2 | p>.05 | 46.0 | -0.020 | p>.05 |
| Familiarity | 5.745 | 2 | p>.05 | 5.051 | 2 | p>.05 | 34.5 | -0.136 | p>.05 |
| Training | 1.611 | 2 | p>.05 | 4.452 | 2 | p>.05 | 40.0 | -0.081 | p>.05 |
| System design | 5.016 | 2 | p>.05 | 5.812 | 2 | p>.05 | 9.5 | -0.391 | **p<.05** |
| Workload | 1.910 | 2 | p>.05 | 0.854 | 2 | p>.05 | 43.5 | -0.045 | p>.05 |
| Communication | 1.297 | 2 | p>.05 | 0.376 | 2 | p>.05 | 37.0 | -0.110 | p>.05 |
| Information | 1.985 | 2 | p>.05 | 0.709 | 2 | p>.05 | 31.0 | -0.173 | p>.05 |

Table 8-25 (cont.) Effect of classification groups on the R-PSFs rating for "signal wrongly interpreted"

| R-PSFs | Age (train drivers) | | | Ops. Experience (train drivers) | | | Involvement in SPAD (train drivers) | | |
|---|---|---|---|---|---|---|---|---|---|
| | H | df | df | H | df | Sig. | U | r | Sig. |
| Distraction | 5.011 | 3 | p>.05 | 2.021 | 3 | p>.05 | 120.0 | -0.031 | p>.05 |
| Fatigue | 3.642 | 3 | p>.05 | 3.380 | 3 | p>.05 | 98.0 | -0.169 | p>.05 |
| Perception | 1.855 | 3 | p>.05 | 1.796 | 3 | p>.05 | 101.0 | -0.152 | p>.05 |
| Procedures | 4.700 | 3 | p>.05 | 2.288 | 3 | p>.05 | 121.5 | -0.022 | p>.05 |
| Safety culture | 2.080 | 3 | p>.05 | 2.299 | 3 | p>.05 | 115.0 | -0.062 | p>.05 |
| Supervision | 4.342 | 3 | p>.05 | 0.217 | 3 | p>.05 | 92.0 | -0.207 | p>.05 |
| Familiarity | 0.389 | 3 | p>.05 | 2.777 | 3 | p>.05 | 118.0 | -0.044 | p>.05 |
| Training | 6.640 | 3 | p>.05 | 6.091 | 3 | p>.05 | 98.5 | -0.165 | p>.05 |
| System design | 0.931 | 3 | p>.05 | 1.012 | 3 | p>.05 | 104.0 | -0.132 | p>.05 |
| Workload | 5.370 | 3 | p>.05 | 1.309 | 3 | p>.05 | 93.0 | -0.199 | p>.05 |
| Communication | 3.629 | 3 | p>.05 | 3.379 | 3 | p>.05 | 121.0 | -0.025 | p>.05 |
| Information | 4.621 | 3 | p>.05 | 1.159 | 3 | p>.05 | 102.0 | -0.145 | p>.05 |

Again, as the ratings for most of the classification groups are not significantly different, data are combined. Figure 8-4 (left) displays the final results, while Figure 8-4 (right) shows the ratings as given by the three different SMEs subgroups. Distraction was characterized as the most suboptimal factor, although it is not indicated as such from all the subgroups. In addition, it can be seen that the total R-PSFs ratings are higher, compared to those for the first human error.

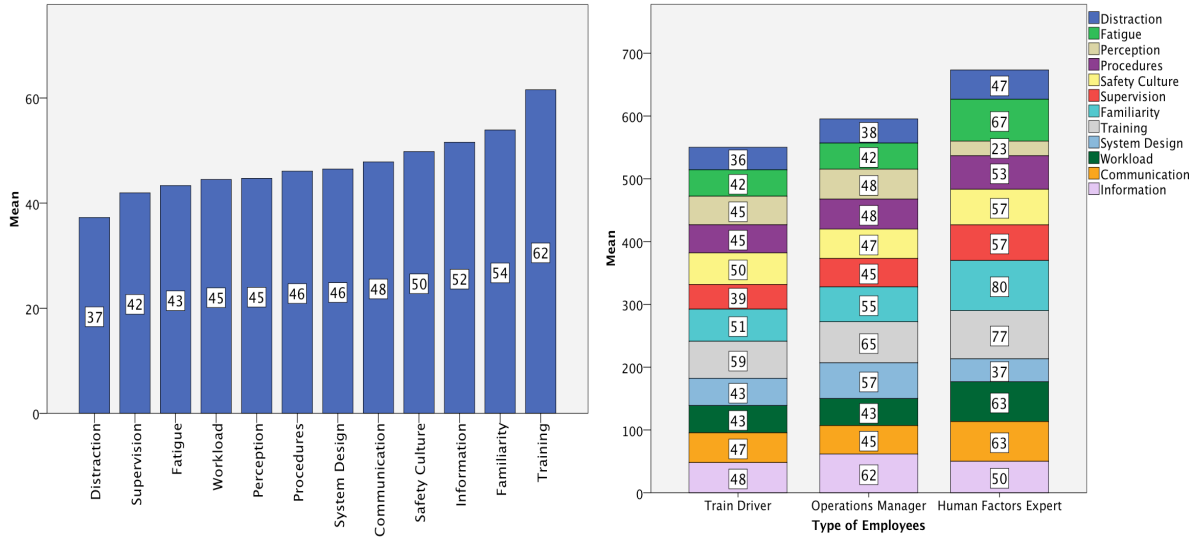



Figure 8-4 R-PSFs ratings for signal wrongly interpreted

### 8.4.3 The SMEs ratings for "executed action"

Finally, the third erroneous action is the failure of the operator to correctly execute the required action, e.g. apply braking strategy. Results (Table 8-26) show that no significant differences were found for the scores given for most of the R-PSFs.

However, it was found, that male train drivers consider less suboptimal the design of the system for that specific action, compared to the female train drivers. In addition, a difference was also identified for distraction, where younger (35-44) train drivers assigned lower scores compared to older (45-56) drivers.

Data are combined, and final results are displayed in Figure 8-5 (left). Figure 8-5 (right), in addition, displays the ratings as given by the three different SMEs subgroups. Three important findings can be observed in the two figures. Firstly, distraction is again indicated as the most suboptimal factor amongst the factors. Secondly, five factors have been assigned with scores less than 40. Thirdly, the total given ratings for the train drivers are the lowest amongst the three erroneous actions.

Due to the differences in the responses amongst the types of errors, the next section examines whether the characteristics of the actions have affected the ratings of the SMEs.

Table 8-26 Effect of classification groups on the R-PSFs rating for "action not executed correctly"

| R-PSFs | Type of SMEs | | | Type of operations (train drivers) | | | Gender (train drivers) | | |
|---|---|---|---|---|---|---|---|---|---|
| | H | df | Sig. | H | df | Sig. | U | r | Sig. |
| Distraction | 1.403 | 2 | p>.05 | 0.542 | 2 | p>.05 | 44.0 | -0.041 | p>.05 |
| Fatigue | 1.885 | 2 | p>.05 | 0.618 | 2 | p>.05 | 33.0 | -0.151 | p>.05 |
| Perception | 1.643 | 2 | p>.05 | 2.079 | 2 | p>.05 | 27.0 | -0.216 | p>.05 |
| Procedures | 1.839 | 2 | p>.05 | 1.091 | 2 | p>.05 | 41.0 | -0.071 | p>.05 |
| Safety culture | 0.839 | 2 | p>.05 | 2.095 | 2 | p>.05 | 40.5 | -0.076 | p>.05 |
| Supervision | 2.264 | 2 | p>.05 | 3.758 | 2 | p>.05 | 33.5 | -0.147 | p>.05 |
| Familiarity | 0.626 | 2 | p>.05 | 5.118 | 2 | p>.05 | 38.5 | -0.096 | p>.05 |
| Training | 1.010 | 2 | p>.05 | 1.448 | 2 | p>.05 | 33.5 | -0.146 | p>.05 |
| System design | 4.966 | 2 | p>.05 | 2.146 | 2 | p>.05 | 7.0 | -0.423 | **p<.05** |
| Workload | 2.852 | 2 | p>.05 | 2.307 | 2 | p>.05 | 27.0 | -0.211 | p>.05 |
| Communication | 2.219 | 2 | p>.05 | 0.024 | 2 | p>.05 | 38.5 | -0.096 | p>.05 |
| Information | 2.360 | 2 | p>.05 | 0.429 | 2 | p>.05 | 33.5 | -0.147 | p>.05 |
| **R-PSFs** | **Age (train drivers)** | | | **Ops. Experience (train drivers)** | | | **Involvement in SPAD (train drivers)** | | |
| | **H** | **df** | **df** | **H** | **df** | **Sig.** | **U** | **r** | **Sig.** |
| Distraction | 10.974 | 3 | **p<.05** | 3.284 | 3 | p>.05 | 121.5 | -0.022 | p>.05 |
| Fatigue | 7.215 | 3 | p>.05 | 2.690 | 3 | p>.05 | 124.0 | -0.006 | p>.05 |
| Perception | 2.219 | 3 | p>.05 | 0.423 | 3 | p>.05 | 109.0 | -0.102 | p>.05 |
| Procedures | 3.259 | 3 | p>.05 | 3.664 | 3 | p>.05 | 91.0 | -0.214 | p>.05 |
| Safety culture | 1.169 | 3 | p>.05 | 1.100 | 3 | p>.05 | 112.0 | -0.081 | p>.05 |
| Supervision | 5.191 | 3 | p>.05 | 1.569 | 3 | p>.05 | 112.5 | -0.078 | p>.05 |
| Familiarity | 0.661 | 3 | p>.05 | 5.622 | 3 | p>.05 | 89.0 | -0.225 | p>.05 |
| Training | 1.023 | 3 | p>.05 | 1.714 | 3 | p>.05 | 118.0 | -0.043 | p>.05 |
| System design | 2.282 | 3 | p>.05 | 0.596 | 3 | p>.05 | 121.5 | -0.022 | p>.05 |
| Workload | 3.268 | 3 | p>.05 | 0.490 | 3 | p>.05 | 89.5 | -0.222 | p>.05 |
| Communication | 6.462 | 3 | p>.05 | 4.358 | 3 | p>.05 | 116.5 | -0.053 | p>.05 |
| Information | 4.614 | 3 | p>.05 | 3.131 | 3 | p>.05 | 91.5 | -0.210 | p>.05 |

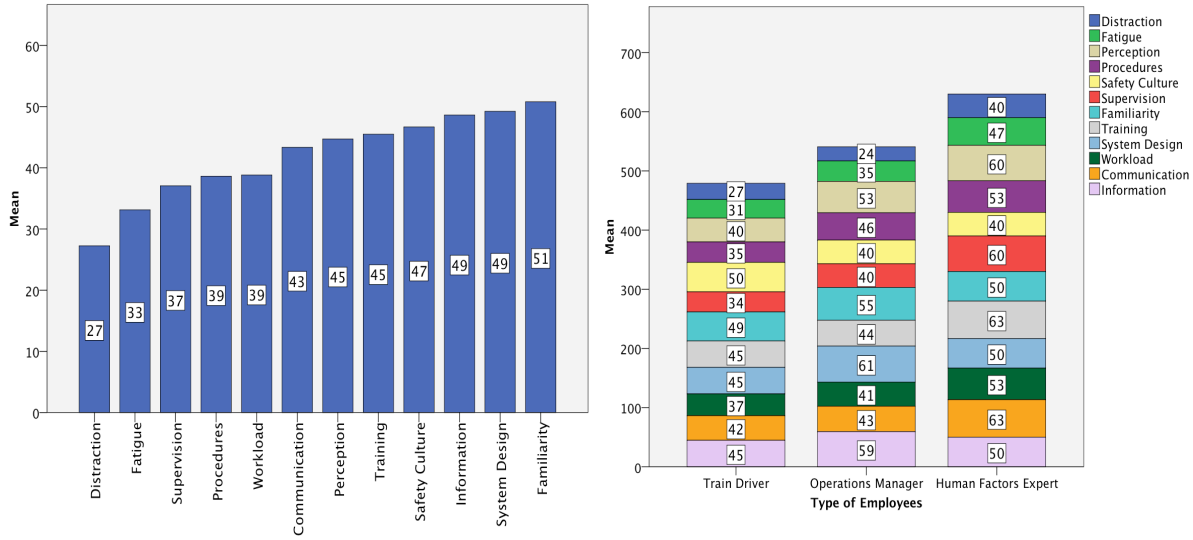



Figure 8-5 R-PSFs ratings for action not executed correctly

### 8.4.4 Effect of actions on the SMEs ratings

In order to indicate if operational actions have an impact on the SMEs ratings the Friedman's ANOVA test is used. This test examines differences between more than two conditions when same participants have been used in all conditions. In this study, the same participants were employed to indicate the quality of the contributing factors for the three actions. Friedman's test, is based on ranked data, while its test statistic, $F_r$, is calculated as:

$$F_r = \left[ \frac{12}{Nk(k+1)} \sum_{i=1}^{k} R_i^2 \right] - 3N(k+1) \qquad (8\text{-}3)$$

where $R_i$ is the sum of ranks for each group, $N$ is the total sample size and $k$ the number of conditions. For more than ten participants the test statistic has a chi-square distribution; again when the test is significant, ($p<.05$), it is assumed that the conditions are significantly different. Hence, it is accepted that the characteristics of the actions have an impact on the rating of factors. For those cases *post hoc* tests, i.e. the Wilcoxon signed-rank test with the Bonferroni correction, are performed to identify the origins of the differences (Field, 2009). The Wilcoxon signed-rank test is based on the differences between the scores in the two compared conditions. Once the differences are calculated, they are ranked and the sign of the difference (+ or -) is assigned to the rank. The test statistic is obtained by the smaller of the two values between the sum of the positive and negative ranks. The statistics for the R-PSFs ratings are illustrated in the Table 8-27.

Table 8-27 Friedman's ANOVA statistics for the R-PSFs ratings

| | Distraction | Fatigue | Perception | Procedures | Safety Culture | Supervision |
|---|---|---|---|---|---|---|
| $F_r$ | 8.218 | 9.750 | 3.871 | 12.500 | 3.622 | 2.725 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | **<.05** | **<.01** | >.05 | **<.01** | >.05 | >.05 |
| | **Familiarity** | **Training** | **System Design** | **Workload** | **Communication** | **Information** |
| $F_r$ | 1.450 | 24.476 | 0.441 | 6.117 | 6.756 | 0.970 |
| df | 2 | 2 | 2 | 2 | 2 | 2 |
| Sig. | >.05 | **<.01** | >.05 | **<.05** | **<.05** | >.05 |

As can be seen significant differences amongst the SMEs responses were identified for six factors. For these factors *post hoc* tests indicate that:

(i) although the very low ratings of distraction, as shown in the previous sections, the assigned values for signal interpretation were significantly higher compared to those for signal detection, z=-2.620, p<.01, r= -.273;

(ii) the assigned ratings of fatigue with respect to signal interpretation were significantly higher compared to signal detection, z=-2.757, p<.01, r= -.259 and action executed, z=-2.735, p<.01, r= -.270. In other words, the SMEs deem that the train operators are probably more tired when SPADs occur either due to signal not detection or actions not executed correctly;

(iii) the origin of the significant differences for the ratings of procedures could not be established;

(iv) the values for training were significantly lower when the SMEs rated the third action, i.e. action executed correctly, compared to signal detection, z=-3.395, p<.01, r=- .336 and signal interpretation, z=-4.190 p<.001, r=- .415;

(v) the significant difference in the ratings of workload is established due to the higher values assigned to signal interpretation compared to action executed correctly, z=-2.501, p<.05, r=- .247;

(vi) the ratings of communication indicate significant difference between the operational actions signal interpretation and action executed correctly, with the former to be assigned with higher values, z=-2.480, p<.05, r=- .245.

The above results indicate that the characteristics of the operational erroneous actions have an impact on the SMEs ratings. In particular, the SMEs have generally assigned lower values, to the majority of the factors when SPADs are caused due to operators failure to act as required.

Having presented the analysis of the weightings and ratings of the R-PSFs, the next section discusses the calculations of the HuPeROIs for the three operational actions.

## 8.5 The HuPeROI calculations – the relative likelihoods of errors

The last part of the thesis presents the HuPeROI calculations. As a reminder, the HuPeROI indicates the ordering of the likelihoods of the different actions under investigation. Hence, once it is determined, it could be then related to the success probability of an action that would be observed in the long run in the situation of interest.

The index for each of the actions was calculated, as shown in Section 7.5.9. Results were again tested for outliers. One outlier was found amongst the results and excluded from the analysis. Moreover, the normality of data was examined, using the Shapiro-Wilk test. All tests were found non-significant, ($p<.05$), thus confidence is gained that data is normally distributed. The HuPeROI results for each operational action are illustrated in the Figure 8-6.

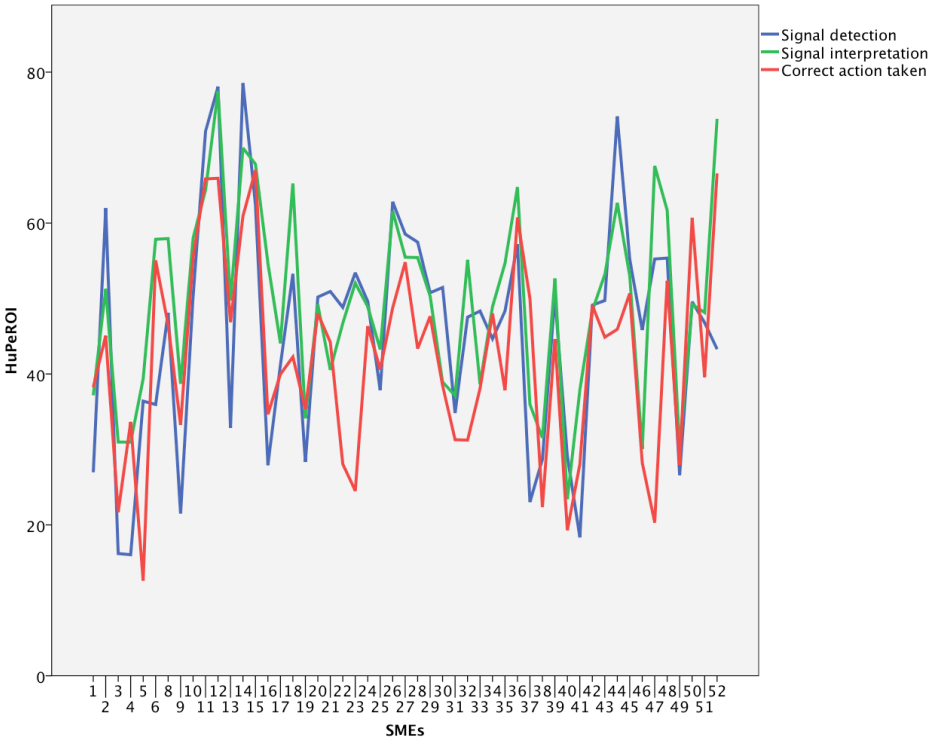



Figure 8-6 The HuPeROI calculations per type of action

Findings in the Figure 8-6 suggest that overall signal interpretation was assigned with the highest HuPeROIs amongst the three actions. Subsequently, it can be considered the least likely erroneous action. In addition, comparing the HuPeROIs per action and SME, it can be

seen that the majority of the SMEs were consistent on the way that they have assigned their scores.

To compare the results, ANOVA was applied to test whether the means of the different types of the SMEs, for each action, are significantly different. Results, as shown in the Table 8-28, indicate no significant differences between the groups of respondents.

Table 8-28 Comparison of HuPeROIs means between different SMEs groups

| Operational action | F-statistic | df | Sig. |
|---|---|---|---|
| Signal detection | 0.163 | 2 | p>.05 |
| Signal interpretation | 0.730 | 2 | p>.05 |
| Action executed | 0.855 | 2 | p>.05 |

Therefore, it is assumed that data belong to the same population. The final results are summarised in Table 8-29. Based on them, it is extracted that the most likely error is the failure of operators to execute the required action correctly in order to stop the train in time. On the other hand, the least likely error is the signal's wrong interpretation.

Table 8-29 The HuPeROI results per type of SME and operational action

| | | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Min. | Max. |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound | | |
| HuPeROI signal detection | Train Driver | 34 | 46.6603 | 17.0003 | 2.9155 | 40.7286 | 52.5920 | 16.0508 | 78.5344 |
| | Operations Manager | 14 | 43.8982 | 12.2273 | 3.2679 | 36.8384 | 50.9580 | 23.0325 | 62.4113 |
| | HFs expert | 3 | 46.7718 | 3.4724 | 2.0048 | 38.1458 | 55.3978 | 43.2878 | 50.2326 |
| | **Total** | **51** | **45.9086** | **15.2204** | **2.1313** | **41.6278** | **50.1894** | **16.0508** | **78.5344** |
| HuPeROI signal interpretation | Train Driver | 34 | 49.0766 | 13.0443 | 2.2371 | 44.5252 | 53.6280 | 23.4137 | 77.5124 |
| | Operations Manager | 14 | 49.1045 | 11.5242 | 3.0800 | 42.4506 | 55.7584 | 31.5390 | 67.8134 |
| | HFs expert | 3 | 58.2026 | 13.7110 | 7.9161 | 24.1425 | 92.2626 | 48.1374 | 73.8187 |
| | **Total** | **51** | **49.6211** | **12.6113** | **1.7659** | **46.0741** | **53.1681** | **23.4137** | **77.5124** |
| HuPeROI action executed | Train Driver | 34 | 40.9721 | 13.4339 | 2.3039 | 36.2848 | 45.6594 | 12.6227 | 65.9363 |
| | Operations Manager | 14 | 44.2281 | 12.6978 | 3.3936 | 36.8967 | 51.5596 | 22.3827 | 67.0106 |
| | HFs expert | 3 | 50.2642 | 14.3512 | 8.2857 | 14.6139 | 85.9145 | 39.5895 | 66.5786 |
| | **Total** | **51** | **42.4125** | **13.2401** | **1.8540** | **38.6887** | **46.1364** | **12.6227** | **67.0106** |

Results in Table 8-29 may also imply a significant difference in the responses of train drivers. Since the weightings of the factors are constant, such a difference is explained by the differences in the ratings of the R-PSFs. Finally, it can be seen that the standard error in the responses of the HFs experts for the last two actions can be considered large, compared to the sample mean. This is explained by the size of the group. Its small size permits the indices to be influenced more by any judge who varies significantly from the rest of the group, as shown in Figure 8-7.

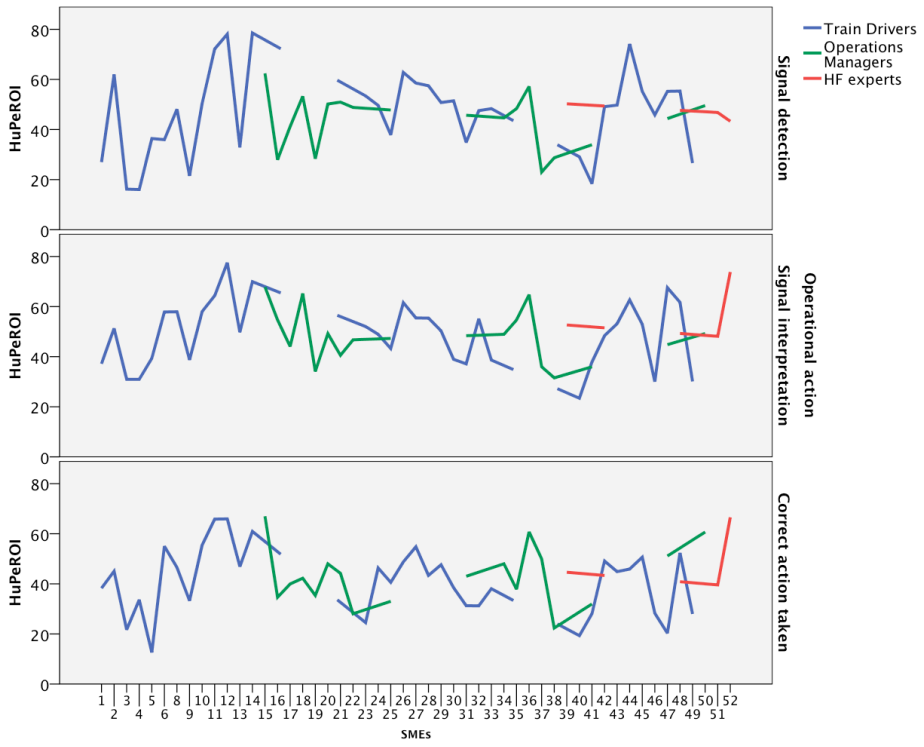



Figure 8-7 The HuPeROI calculations per type of SME and type of action

Finally, results have been tested also for significant differences amongst the different types of train drivers. Table 8-30 shows that there is no significant difference amongst the respondents.

Table 8-30 Comparison of HuPeROIs means between train drivers from different TOCs

| Operational action | F-statistic | df | Sig. |
|---|---|---|---|
| Signal detection | 0.371 | 2 | p>.05 |
| Signal interpretation | 0.293 | 2 | p>.05 |
| Action executed | 1.021 | 2 | p>.05 |

Subsequently, it can be assumed that neither the types of operations, nor the characteristics of the railway organisation have a significant impact on the perception of train drivers with respect to the most likely error that could lead to a SPAD. The Figure 8-8 illustrates the responses of the train drivers for the three TOCs.

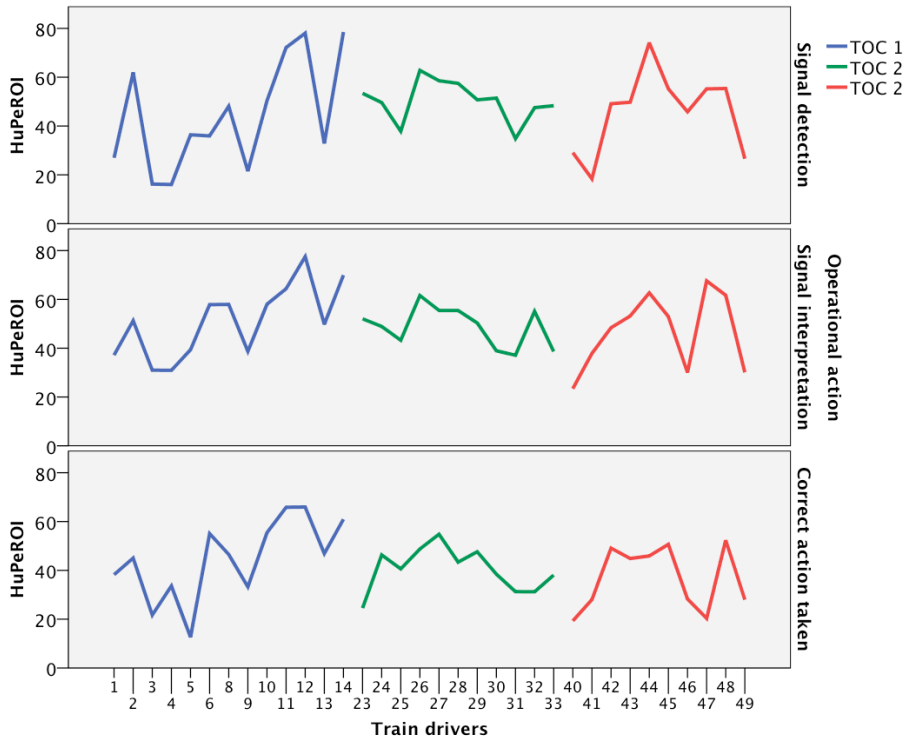



Figure 8-8 The HuPeROI calculations per train driver, TOC and type of action

Taking all the responses into account, it can be seen that all the SMEs, regardless their duties or type of operation, indicate that the least likely error to occur is the failure of the train drivers to interpret the signal. In addition, it can be also seen that the most likely error is considered the failure of the operator to execute the required action correctly. In addition, only for the train drivers, differences amongst the classification groups were also examined. Results indicate that the responses between the personnel with different operational experience significantly vary. In particular, it was found that for signal detection the least experienced personnel (< 1 year) have assigned lower scores compared to all the remainder groups of employees, as shown in the Table 8-31.

Table 8-31 The HuPeROI of signal detection for the classification group operational experience

| | Operational experience | N | Mean | Std. Deviation | Std. Error Mean | t-statistic | Sig. |
|---|---|---|---|---|---|---|---|
| HuPeROI not signal detection | < 1 year | 6 | 30.8598 | 11.5886 | 4.7310 | -2.807 | p<.05 |
| | 6-10 yrs | 6 | 55.8616 | 18.4839 | 7.5460 | | |
| HuPeROI not signal detection | < 1 year | 6 | 30.8598 | 11.5886 | 4.7310 | -2.027 | p≈.05 |
| | 11-19 yrs | 9 | 46.1959 | 15.8362 | 5.2787 | | |
| HuPeROI not signal detection | < 1 year | 6 | 30.8598 | 11.5886 | 4.7310 | -3.580 | p<.01 |
| | ≥ 20 yrs | 12 | 52.8589 | 12.5963 | 3.6362 | | |

## 8.6 Validation of the HuPeROI findings

Having defined the HuPeROI values the next step in the process is to validate the findings. Ideally, such a validation would be performed by comparing the calculations with (i) results from previous studies, and (ii) simulation observations where train drivers would be asked to operate according to the scenario guidelines. However, due to the lack of similar studies, there are no data available to directly compare the HuPeROI findings. In addition, due to resource constraints, simulation observations could not be performed within the scope of this research.

Therefore, results are compared against information obtained from relevant investigation reports and information collected from the questionnaire survey. The validation comprises two phases, the internal and external. In the internal phase, the relevant UK SPAD reports from the set of the 479 reports are reviewed. In addition, the responses of the SMEs concerning their previous SPAD experiences are also examined. In the external phase, additional SPAD reports are collected and reviewed. The findings from both phases are then compared to the HuPeROI results.

### 8.6.1 Internal validation

For the internal validation, 58 UK SPADs (see Section 8.3.13) were reviewed and the immediate causes defined. The results are summarized in the Figure 8-9.

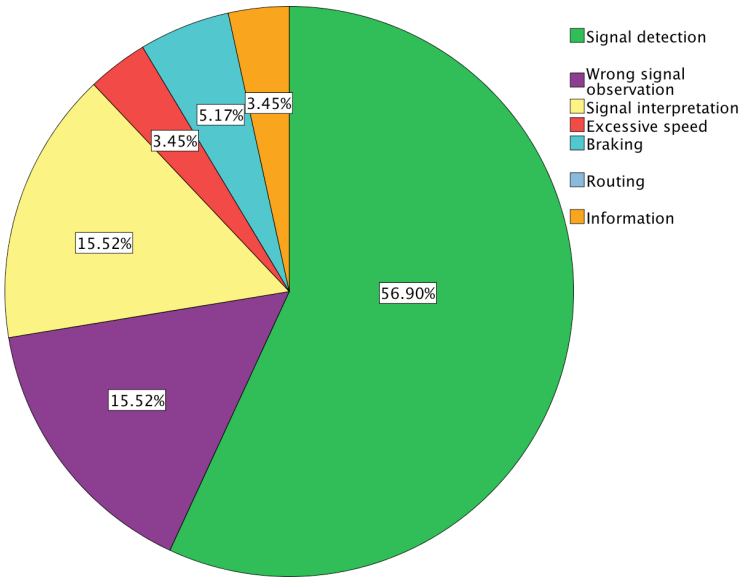



Figure 8-9 SPADs immediate causes based on UK reports

According to the data, train drivers who failed to detect a signal caused the majority of the occurrences. In addition, SPADs due to wrong signal interpretation do not account for more than 15% of total SPAD events. Finally, data suggest that SPADs caused by train drivers who fail to execute an appropriate action as required, account for less than 9% of the occurrences.

It was mentioned in the previous section that the HuPeROI values indicate the ordering of the likelihoods of different erroneous actions. Thus, the calculated values do not reflect the percentage differences found in real data. However, comparing the extracted data with the HuPeROI results, it is clear that the former can only partially confirm the latter.

In particular, there is agreement between the HuPeROI likelihoods ordering for the SPADs caused due to wrong signal interpretation or failure on signal detection and real data. Such agreement confirms that the most likely error, between the two potential errors, is the failure of train drivers to detect the signal. In contrast, disagreement is found for the SPAD occurrences caused by train drivers who did not execute the action correctly. Although, the HuPeROI calculations indicate this type of error as the most likely, real data suggest that only few occurrences were caused by an action that has not been executed as required. That means, that the performance of the operators, as expressed by the R-PSFs rating and weighing, does not necessarily reflects the likelihood of an error.

In addition, the HuPeROI results were compared to the SMEs responses, as collected from the survey (Question 1.10 in the Appendix XI). In total 15 SMEs, including 10 train drivers and 5 operations managers, stated that they had been involved to at least one SPAD in the past. Eight respondents claimed that such event occurred due to their failure to detect the signal. In addition, two respondents said that they did not interpret the signal correctly, while only 1 stated that he did not applied braking as required. Again, findings reveal a gap between the determined likelihood ordering of *task execution* and the proportion of this error in SPAD occurrences.

### 8.6.2 External validation

For the external validation, 43 additional SPAD events were collected, which refer to events that occurred on the UK mainline after 2011. The immediate causes were again extracted and the results are shown in the Figure 8-10.

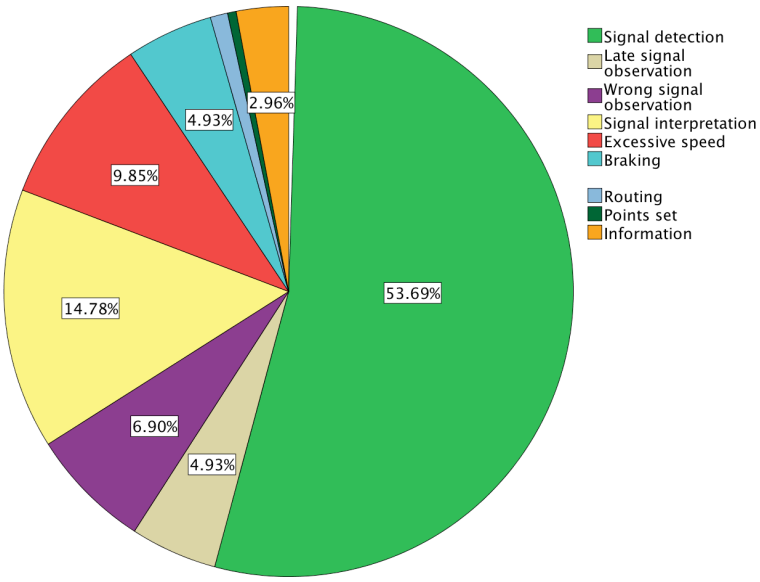



Figure 8-10 SPADs immediate causes based on UK initial reports post 2011

Data show that the failure of train drivers to detect a signal remains the prime reason for SPAD occurrences. Additionally, it is confirmed that train drivers' inability to interpret a signal lead to less SPADs compared to those caused by non-signal detection. Finally, the number of occurrences due to wrongly executed actions was found to be almost two times higher compared to the initial reports. Nevertheless, this is again lower than the events that occur due to non-signal detection. Subsequently, the HuPeROI do not reflect the likelihood of the error.

Whilst it is acknowledged that the performed validation is not optimal, it can be claimed that the HuPeROI can be used by the railway industry in order to indicate the ordering of likelihoods for different erroneous actions. Results could be also used to assess the performance of the operators, where the higher the HuPeROI value the better the performance of the operator. On the other hand, however, it was also found that the HuPeROI values do not necessarily reflect the proportion of causes. In other words, it is not always the case that the lowest assessed performance, i.e. the lowest HuPeROI value, would describe the majority of occurrences.

## 8.7 The influence of external conditions on performance

As described in the Section 7.5.10, in the last part of the questionnaire the SMEs were asked to indicate whether their performance may change when operating under certain conditions. In particular, they were asked to assess the change in their performance, when operating under (i) intermediate weather conditions, (ii) adverse weather conditions, (iii)

degraded operations, (iv) night, and (v) in a tunnel (for the underground operators only). The Table 8-32 illustrates the main findings.

Table 8-32 Change in performance under certain conditions

| | Change in performance | N | Mean | Std. Deviation | Std. Error Mean | F-statistic | Sig. |
|---|---|---|---|---|---|---|---|
| Intermediate weather conditions | Improved | 12 | 67.92 | 24.630 | 7.110 | 0.118 | p<.05 |
| | Reduced | 24 | 47.08 | 27.104 | 5.532 | 3.741 | |
| Adverse weather conditions | Improved | 18 | 71.11 | 27.842 | 6.562 | 0.775 | p<.05 |
| | Reduced | 25 | 50.00 | 30.822 | 6.164 | 0.932 | |
| Night operation | Improved | 12 | 49.17 | 30.833 | 8.915 | 1.781 | p<.05 |
| | Reduced | 11 | 38.18 | 32.502 | 9.800 | 1.902 | |
| Degraded operations | Improved | 19 | 61.05 | 33.976 | 7.795 | 2.037 | p<.05 |
| | Reduced | 14 | 48.57 | 35.270 | 9.426 | 0.576 | |
| Tunnel operation | Improved | 9 | 35.56 | 36.094 | 12.031 | 3.909 | p<.05 |
| | Reduced | 3 | 30.00 | 34.641 | 20.000 | 2.000 | |

As can be seen from this table, for all but adverse weather conditions, at least one third of the SMEs state that such conditions would not affect the operator performance. However, it is also found that when operating under degraded or night operations, slightly more SMEs indicate that the performance of the operators would be improved. This is explained by the fact that the SMEs declare that when operating under those conditions, they are more alert and vigilant.

For the impact of weather conditions on the operator performance, it can be seen that most of the SMEs consider such an impact negative. However, the SMEs who indicate that adverse or intermediate weather conditions would improve their performance have given higher scores to such change. Again, the levels of vigilance and alertness due to the difficult external conditions is claimed to positively affect the performance of the operators. On the other hand, such conditions may increase fatigue and stress levels. Overall, however, weather conditions negatively affect performance.

## 8.8 Summary

This chapter presented the analysis and results of the case study for the implementation of the HuPeROI. Although generalising the substantive findings from analysis at large will require caution, there are a number of significant achievements and findings. Firstly, it was shown how the methodology could be implemented by the railway industry to quantify the factors that affect the performance of the operators. For example, for the case study used, it

was found that the *dynamic personal factors* category has the highest impact on operator performance. Moreover, with respect to the R-PSFs elements, it was shown that all the SMEs consider workload as the most important contributing factor to human performance.

In addition, the HuPeROI framework allows analysts to extract information on the quality of the factors at the time of the occurrence for different operational errors. This is achieved by the rating of factors. Moreover, the approach provides a useful tool to determine similarities and differences amongst the responses of the different types of employees. In this case study the majority of responses are not significantly different. Nevertheless, it was clearly stated that the types of the SMEs may consider the impact of the R-PSFs on human performance differently. Thus, the collection of additional data is recommended to confirm and consolidate the initial results, as well as to provide generic statements for the railway industry.

Finally, the findings indicate that HuPeROI can be efficiently used to assess the performance of operators as function of the quality of the relevant R-PSFs. However, initial results failed to associate the performance of the operators, as expressed by the ordering of relative error likelihoods, with the results as observed by the investigation reports. Thus, the integration of the HuPeROI framework with simulation observations is recommended. The SMEs would first assess and rate the factors that affect their performance, as described in this thesis. They would then operate under the conditions of the described scenario and their behavior would be recorded. Subsequently, the industry would be able to compare the theoretical responses with real observations and finally associate the performance with types of actions and relative likelihoods.

Having presented the results of the analysis, and showing the substantial role that the HuPeROI framework can play in the enhancement of the safety on railway operation, the next chapter discusses the main findings of this research and suggests avenues for future work.

# Chapter 9 CONCLUSIONS

This chapter presents the main findings of the research on performance shaping factors and human performance in railway operations and suggests avenues for further work. The approach taken for the former is to address each of the research objectives formulated in Chapter 1 (repeated below for ease of reference) and to present the corresponding findings and main contributions of this thesis. The chapter concludes with a series of research ideas for future work and a list of publications from this thesis.

## 9.1 Revisiting the research objectives

The aim and objectives of this thesis are re-stated to facilitate the full understanding and impact of the main findings of this thesis. The aim of this thesis was to develop a novel approach to assess human performance and Chapter 1 defined a set of four objectives for this thesis. These are:

Chapter 1 defined a set of research objectives for this thesis. These are to:

1. Identify the modern railway system, categorise and describe its components and finally define its operational architecture.
2. Understand the operational concept of human errors and review the models to define their origins. In turn, identify the state-of-the-art for the existing Human Performance Analysis (HPA) techniques, review and capture their limitations.
3. Determine the contextual factors that affect railway operators performance and based on a holistic framework suggest a novel taxonomy, referred to as the Railway Performance Shaping Factors taxonomy, previously lacking in the railway domain.
4. Develop a Human Performance Railway Operational Index (HuPeROI) to:
   (i) assess the influence of each identified R-PSF on the performance of railway operators accounting for all the dependencies between the R-PSFs, and in turn to
   (ii) assess the performance of operators for several railway operational scenarios, by comparing the likelihoods of different human errors.

## 9.2 Conclusions

This section outlines the major achievements of this thesis with respect to the objectives.

### 9.2.1 The modern railway system and its operational architecture

To study and enhance the performance of railway operators and to identify the contextual factors that have an impact on their performance, first the railway system and its operational framework should be defined. This was undertaken in Chapter 2 which:

- Reviews the various definitions of the railway system, addresses their differences and similarities and introduces a new detailed definition. In particular, this definitions accounts for those parties of the system that do not belong to railway organisations, i.e. passengers, trespassers. Based on that definition, a comprehensive description of the constituents of the system in provided.
- Details the operational architecture of the railway system, referred to as Railway Operational System Architecture, which illustrates the dependencies and interactions amongst operators, equipment and infrastructure. Experts of different types of railways services and countries have validated ROSA. Thus, it can be broadly used from the railway industry from academics and practitioners for gaining understanding of the operational concept and the potential failures.

Furthermore, in Chapter 3, the following issues were analysed:

- Alternative definitions of accident and incident from major organisations in railway safety, including the definitions given by the ERA, FRA ATSB, CTSB and UIC. The differences and similarities of those definitions explained and it showed why the ERA definitions are considered the most comprehensive,
- Historic accident trends, were analysed and results apart from the continuous improvements of railway safety, revealed that humans are those who mostly affect the system in terms of safety.
- In particular data showed that within the controlled components of the railway system the great majority of serious accidents occur due to human error.

### 9.2.2 Human Error – Human Performance Analysis Techniques

The understanding of human error, the origins of human error, and their relations to human performance were substantial parts of this research in order to conceptualise operational errors. Chapter 4 contributed to this objective by:

- Providing a consolidated literature review of human error definitions, human error classifications and the existing models to analyse and investigate the origins of human errors.
- Providing critical review of the state-of-the-art literature in the filed of Human Performance Analysis and the associated techniques that examine and assess human performance.

### 9.2.3 The Railway Performance Shaping Factors taxonomy

In order to assess the performance of railway operators and to identify the causes of their operational errors, the factors that affect their performance should be defined. With respect to this, Chapter 5:

- Introduces a novel and comprehensive taxonomy for railway operations, referred to as the Railway Performance Shaping Factors (R-PSFs) taxonomy. The R-PSFs is the first taxonomy developed on a holistic approach including: a detailed literature review, on-site observations, analysis of worldwide accidents and incidents report and SMEs validation;
- Provides detailed and clear definitions for all the included factors within the R-PSFs taxonomy, so it could be easily implemented by researchers and analysts regardless of their level of expertise in the field of human factors;
- Accounts for the dependencies amongst the factors and thus supports researchers to gain a better understanding of the influence of one factor on another.

In addition within the R-PSFs concept, in Chapter 6:

- A smaller list of factors is introduced that account for the majority of the occurrences regardless of their severity;
- Such a list of factors can be used by railway stakeholders as part of the current investigation schemes in order to consistently record data regarding occurrences due to human failure.

To date the R-PSFs taxonomy is being implemented by London Underground Limited, as part of their operational training module (Appendix I). In addition, it has been reviewed and

judged applicable to capture the impact of the ERTMS implementation on human performance (Section 9.4.3)

### 9.2.4 The Safety Maturity Model

To consolidate the taxonomy, safety data collected from railway organisations were used. To assess the extent to which data were obtained from reliable sources, this thesis (Chapters 5&6) developed a novel approach referred to as Safety Maturity Model (SMM). The SMM:

- Unlike similar models, addresses not only the behavioural and/or attitudinal culture of organisations, but also technical, operational and methodological elements and actual achievements in terms of safety outcomes. In addition it quantifies the level of safety maturity.
- Although initially applied to metro services, clearly the SMM is transferable to main line operations as its main focus is on an organisation's safety standards, and not specific types of operations.

### 9.2.5 The Human Performance Railway Operational Index (HuPeROI)

The framework to assess the performance of railway operators as function of the most influential R-PSFs was developed (Chapter 7 and 8).

- The HuPeROI allows researchers to quantify the impact of each of the R-PSFs on the performance of the operators, considering all the dependencies amongst those factors;
- The framework, for first time fully integrates the SLIM with the ANP methodologies;
- The framework was tested and validates by three UK train operators and the findings indicate that it can be broadly applied in order not only to assess the performance of the operators but also to indicate the differences between the different types of networks, organisations and employees.
- Finally, the HuPeROI can be also used to identify the relationship between the assessed performance of the operators and the quality of the R-PSFs for particular errors.

## 9.3 Future work and recommendations

The research presented in this thesis introduces a framework to evaluate the factors that influence the performance of railway operators and to assess operators' performance for several operational scenarios. Such an assessment is expressed in terms of the relative likelihood of errors. However, the findings also suggest a number of directions for further research. These include:

- The transformation of the HuPeROI values into a range of probabilities. Although the comparison of the indices indicates which type of error is more likely to happen, it does not clearly define how much more likely this is.
- The conventional ANP-based decision model seems to be ineffective in dealing with the inherent fuzziness or uncertainty in judgment during the pairwise comparison process. Although the use of the discrete scale of 1-9 to represent the verbal judgment in pairwise comparisons has the advantage of simplicity, it does not take into account the uncertainty associated with the mapping of one's perception or judgment to a number. In real-life decision-making situations, the decision makers or stakeholders may be uncertain about their own level of preference, Therefore, it is worth exploring the applicability of fuzzy logic as an optimisation tool, as also indicated in the literature e.g. (Promentilla et al., 2008, Zhou, 2012).
- The ANP approach allows researcher to quantify the different R-PSFs by account for dependences between both the categories and individual elements. However, the technique assumes that no hidden or unused factors within a category exists, thus the residual influence of those factors not included in the elements within a category is not estimated. The indirect influence on performance that caused by the remaining 8 factors not included in the analysis, is yet to be identified.
- The measurement of the dynamic personnel factors. No real assessment of such factors was conducted due to resources constraints, i.e. simulators.
- Future research should explore the trade-offs required in mitigating the 12 R-PSFs with respect to safety, financial and social (reputation) consequences for railway industry. For instance, a change on the design of the system may not only be significantly more expensive than a change in the roster or shift pattern but also result in less improvement in the number of occurrences.
- As railways move towards 24/7 operations, apart from operators' performance the maintenance of railway networks becomes important. Therefore, there is a need to research the factors that affect maintenance. The R-PSFs taxonomy can provide the basis for such research.

## 9.4 Publications – conference presentations relating to this work

The following publications and conference presentations have been produced to support this research. In particular, publications are consist of book chapters, journal publications and conference proceedings.

### 9.4.1 Publication format: book chapters

Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2012). A human performance operational railway index to estimate operator's error probability. In: Stanton, Neville A. (ed.) *Advances in Human Aspects of Road and Rail Transportation*. CRC Press, pp.832-841.

Kyriakidis, M. (2011). Human Reliability Analysis. In: Kröger, Wolfgang and Zio, Enrico (eds.) *Vulnerable Systems*. Springer, pp.157-187.

### 9.4.2 Publication format: journals – published

Kyriakidis, M., Majumdar, A., Grote, G. and Ochieng, W. Y. (2012). Development and Assessment of Performance Shaping Factors Taxonomy for Railway Operations, *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2289, pp.145-153.

Kyriakidis, M., Hirsch, R. and Majumdar, A. (2012). Metro railway safety: An analysis of accident precursors, *Safety Science*, vol. 50, pp.1535-1548,.

Kyriakidis, M., Hirsch, R. and Majumdar, A. (2012). A global safety analysis and best practice for metro railways, *Proceedings of the ICE - Transport*, doi 10.1680/tran.11.00005.

### 9.4.3 Publication format: journals – in press

Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2014). A data-based framework to identify the most significant Performance Shaping Factors in railway operations, *Transportation Research Record: Journal of the Transportation Research Board*, (reccommended for publication).

Smith, P., Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2013). The Impact of ERTMS on Human Performance in Railway Operations - European Findings, *Transportation Research Record: Journal of the Transportation Research Board* (Best Paper Award for the

2013 Young Member Scholarship of the TRB Railroad Operational Safety Committee AR070).

### 9.4.4 Publication format: conference proceedings

Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2013). *Developing a Human Performance Railway Operational Index to enhance safety of railway operations* in the 10th World Congress on Railway Research, Sydney, Australia (accepted for publication).

Kyriakidis, M., Smith, P., Majumdar, A. and Ochieng, W. Y. (2012), *The influence of human performance on railways' technological changes* in the 1st RRUKA Annual Conference. London, U.K., http://rruka.org.uk/rruka-annual-conference-2012-conference-proceedings/.

Kyriakidis, M. A., Hirsch, R. and Majumdar, A. (2010), *Improving Railway Safety: Global Metro Railways' Precursor and Safety Maturity Performance* in the 20th International Railway Safety Conference, Hong Kong, China.

### 9.4.5 Conference presentations

Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2013), *A Human Performance Railway Operational Index (HuPeROI) to improve railway safety* in the the 4th International Human Factors Conference, London, U.K.

Smith, P., Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2013), *The Impact of ERTMS on Human Performance in Railway Operations - European Findings* in the Transportation Research Board 92nd Annual Meeting, Washington D.C., USA.

Kyriakidis, M., Majumdar, A. and Ochieng, W. Y. (2012), *A Human Performance Operational Railway Index to estimate operators' error probability, 3rd European Award in the* Safety and Security pillar of the Young European Arena of Research 2012, Athens, Greece.

Kyriakidis, M., Majumdar, A., Grote, G. and Ochieng, W. Y. (2012), *Development and Assessment of Performance Shaping Factors Taxonomy for Railway Operations* in Transportation Research Board *91st Annual Meeting*, Washington D.C, USA.

Kyriakidis, M. A., Hirsch, R. and Majumdar, A. (2010), *Improving Railway Safety: Global Metro Railways' Precursor and Safety Maturity Performance* in the 20th International Railway Safety Conference, Hong Kong, China.

# REFERENCES


Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. U.S.A.: John Wiley & Sons. Second ed.

Andersen, T. 1999. Human Reliability and Railway Safety. 16th ESReDA seminar, safety and reliability in transport. Published.

Association of American Railroads. (2013). *Highway-Rail Grade Crossing Safety*. Association of American Railroads. From: https://http://www.aar.org/safety/Pages/Grade-Crossing-Safety.aspx - .UaH-iuvyWhM.

Atoc. (2013). *Train Companies* [Online]. Association of Train Operating Companies. From: http://www.atoc.org/train-companies/ Accessed: 2013 09 February.

Australian Infrastructure and Transport Administration. (2009). *Transport Safety Investigation Act 2003, Act No.18 of 2003 as Amended*. Canberra, Australia: Office of Legislative Drafting and Publishing.

Australian Transport Safety Bureau. (2005). *Memorandum of Understanding between the Australian Transport Safety Bureau and the Rail Safery Regulators*. Australian Government of Transport and Regional Services.

Australian Transport Safety Bureau. (2012). *Australian Rail Safety Occurence Data - 1 July 2002 to 30 June 2012*. Australian Transport Safety Bureau.

Backbaus, J. (1980). The Pareto Principle, *Analyse & Kritik*, **2**, 146-171.

Baysari, M. T., Caponecchia, C., Mcintosh, A. S. & Wilson, J. R. (2009). Classification of errors contributing to rail incidents and accidents: A comparison of two human error identification techniques, *Safety Science*, **47**, 948-957. http://dx.doi.org/10.1016/j.ssci.2008.09.012.

Bearfield, G., Short, R. & Griffin, D. (2012). *Hazard analysis and risk assessment for rail projects Phase 1 - Final report*. London: Rail Safety and Standards Board.

Bell, J. & Holroyd, J. (2009). *Review of human reliability assessment methods*. Health & Safety Laboratory. Norwich.

Belton, V. & Stewart, T. J. (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*. Massachusetts, USA: Kluwer Academic Publishers.

Benayoun, R., Roy, B. & Sussman, N. (1966). Manual de Reference du Programme Electre. *Note de Synthese et Formaton*, Paris, France, Direction Scientifique SEMA.

Berghaus, E. (1964). *The History of Railways*. London: Barrie & Rockliff.

Blanchard, H. & Lowe, E. (2009). Using a Model of Train Driver Performance to Investigate the Contribution of Human Factors to Rail Incidents. *3rd International Conference on Rail Human Factors*, Lille, France.

Bonnett, C. F. (2005). *Practical Railway Engineering*. Imperial College Press. 2nd ed.

Boring, R. L. (2005). Human Reliability Analysis Methods for Space Safety. *NASA Annual Risk Management Conference 2005*. Orlando, FL.

Boring, R. L. (2010a). Bridging Resilience Engineering and Human Reliability Analysis *10th International Probabilistic Safety Assessment & Management Conference*. Seattle, WA.

Boring, R. L. (2010b). How Many Performance Shaping Factors are Necessary for Human Reliability Analysis. *10th International Probabilistic Safety Assessment & Management Conference*. Seattle, WA.

Boring, R. L., Gertman, D. I., Joe, J. C. & Marble, J. L. (2005). Human Reliability Analysis in the U.S. Nuclear Power Industry: A Comparison of Atomistic and Holistic Methods. *Human Factors and Ergonomics Society 49th Annual Meeting*. Orlando, FL.

Boring, R. L., Griffith, C. D. & Joe, J. C. (2007). The Measure of Human Error: Direct and Indirect Performance Shaping Factors. *joint 8th IEEE Conference on Human Factors and Power Plants / 13th Conference on Human Performance, Root Cause and Trending (IEEE HFPP & HPRCT)*, Idaho National Laboratory.

Bottero, M., Mondini, G. & Valle, M. (2007). The use of the Analytic Network Process for the sustainability assessment of an urban transformation project. *International Conference on Whole Life Urban Sustainability and its Assessment*. Glasgow. http://download.sue-mot.org/Conference-2007/Papers/Bottero.pdf.

Brackstone, G. (1999). Managing Data Quality in a Statistical Agency, *Statistics Canada, Survey Methodology*, **25**, pp. 139-149.

Brans, J. P., Vincke, P. & Mareschal, B. (1986). How to select and how to rank projects: The Promethee method, *European Journal of Operational Research*, **24**, 228-238. http://dx.doi.org/10.1016/0377-2217(86)90044-5.

Briem, V., De Lima, S. & Siotis, C. (2004). *Train Drivers and Fatal Accidents on the Rails: Psychological Aspects and Safety*. In: John Wilson, Beverley Norris, Theresa Clarke & Ann Mills (eds.) *People and Rail Systems: Human Factors at the Heart of the Railway*. London: Ashgate Publishing. 103-113.

Bubb, H. (2005). Human reliability: A key to improved quality in manufacturing, *Human Factors and Ergonomics in Manufacturing & Service Industries*, **15**, 353-368. 10.1002/hfm.20032.

Buchholz, T., Rametsteiner, E., Volk, T. A. & Luzadis, V. A. (2009). Multi Criteria Analysis for bioenergy systems assessments, *Energy Policy*, **37**, 484-495. http://dx.doi.org/10.1016/j.enpol.2008.09.054.

Cacciabue, P. C. (2004a). Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training, *Reliability Engineering & System Safety*, **83**, 229-240.

Cacciabue, P. C. (2004b). Human error risk management methodology for safety audit of a large railway organisation, *Applied Ergonomics*, **36**, 709-718.

Canadian Centre for Occupational Health and Safety. (2006). *Fit to Work* [Online]. Canadian Centre for Occupational Health and Safety,. From: Canadian Centre for Occupational Health and Safety, Accessed: 26 April 2011.

Cenelec. (1999). *Railway applications – The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS)*. EN50126:1999.

Cenelec. (2001). *Railway applications – Communications, signalling and processing systems – Software for railway control and protection systems*. EN50128:2001.

Cenelec. (2003). *Railway applications – Communication, signalling and processing systems – Safety related electronic systems for signalling*. EN50129:2003.

Chen, Z., Li, H. & Wong, C. (2005). EnvironalPlanning: Analytic Network Process Model for Environmentally Conscious Construction Planning, *Journal of Construction Engineering and Management*, **131**, 92-101. doi:10.1061/(ASCE)0733-9364(2005)131:1(92).

Choudhry, R. M., Fang, D. & Mohamed, S. (2007). The nature of safety culture: A survey of the state-of-the-art, *Safety Science*, **45**, 993-1012. 10.1016/j.ssci.2006.09.003.

Churchman, W. C. (1971). *The Design of Inquiring Systems: Basic Concepts of Systems and Organization*. New York: Basic Books.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New Jersey, USA: Lawrence Erlbaum Associates, Inc. 2nd ed.

Collins English Dictionary. (2009). Glasgow: HarperCollins Publishers. 10th ed.

Cooper, M. D. (2000). Towards a model of safety culture, *Safety Science*, **36**, 111-136. http://dx.doi.org/10.1016/S0925-7535(00)00035-7.

Cooper, S. E., Ramey-Smith, A. M., Wreathall, J., Parry, G. W., Bley, D. C., Luckas, W. J., Taylor, J. H. & Barriere, M. T. (1996). *A Technique for Human Error Analysis (ATHEANA) Technical Basis and Methodology Description*. Washington, DC: U.S. Nuclear Regulatory Commission. NUREG/CR-6350.

Costa, G. (2003). Factors influencing health of workers and tolerance to shift work, *Theoretical Issues in Ergonomics Science*, **4**, 263-288. 10.1080/14639220210158880.

Cullen, W. D. (2001). *The Ladbroke Grove Rail Inquiry Part 1 Report*. Health & Safety Executive.

Dadashi, N., Scott, A., Wilson, J. R. & Mills, A. (eds.). (2013). *Rail Human Factors: Supporting Reliability, Safety and Cost Reduction*, London, UK: CRC Press, Taylor & Francis Group.

De Ambroggi, M. (2010). *The use of cognitive simulation to support dynamic risk modelling.* PhD thesis. Department of Management, Economics and Industrial Engineering. Politecnico di Milano. Milano.

De Ambroggi, M. & Trucco, P. (2011). Modelling and assessment of dependent performance shaping factors through Analytic Network Process, *Reliability Engineering & System Safety*, **96**, 849-860. http://dx.doi.org/10.1016/j.ress.2011.03.004.

De Fontgalland, B. (1984). *The World Railway System*. Cambridge University Press.

Dhillon, B. S. (2007). *Human Reliability and Error in Transportation Systems*. Springer.

Di Pasquale, V., Iannone, R., Miranda, S. & Riemma, S. (2013). *An Overview of Human Reliability Analysis Techniques in Manufacturing Operations*. In: Massimiliano Schiraldi (ed.) *Operations Management*. InTech. pp. 221-240. From: http://www.intechopen.com/books/operations-management/an-overview-of-human-reliability-analysis-techniques-in-manufacturing-operations

Dimattia, D. G. (2004). *Human Error Probability Index for Offshore Platform Musters.* Faculty of Engineering, Chemical Engineering. Dalhousie University. Halifax, Nova Scotia.

Dupuy, M. D. (2011). *Framework for the analysis of separation-related incidents in aviation.* PhD thesis. Centre for Transport Studies, Department of Civil and Environmental Engineering. Imperial College London. London.

Embrey, D. E. (1983). *The Use of Performance Shaping Factors and Quantified Expert Judgment in the Evaluation of Human Reliability: An Initial Appraisal*. Brookhaven National Laboratory. NUREG/ CR-2986.

Embrey, D. E., Humphreys, P. C., Rosa, E. A., Kirwan, B. & Rea, K. (1984). *SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgement*. United States Nuclear Regulatory Commission. Washington DC.

Esveld, C. (2001). *Modern Railway Track*. Zaltbommel, The Netherlands: MRT-Productions. Second ed.

Eub. (2011). *Untersuchungsbericht-Zugkollision Überleitstelle Hordorf / Strecke Mag- deburg Hbf - Halberstadt am 29.01.2011*. Bonn, Germany: Eisenbahn-Unfalluntersuchungsstelle Des Bundes. From: http://www.eisenbahn-unfalluntersuchung.de/cln_033/nn_316888/SharedDocs/Publikationen/EUB/DE/Untersuchungsberichte/2011/023__Hordorf.html.

Eurocontrol. (2003). *The Human Error in ATM Technique (HERA-JANUS)*. Brussels: Eurocontrol.

Eurocontrol. (2007). *Overview of HRA Methods - Farandole Project*. Eurocontrol.

European Commission. (1996). *96/48/EC Directive of 23 July 1996 on the interoperability of the trans-European high-speed rail system.* OJ L235. From: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996L0048:en:HTML Accessed: 19 December 2012.

European Commission. (2000). *HUSARE Human Safe Rail in Europe*. European Commission. RA-97-RS-2094.

European Commission. (2001a). *2001/16/EC Directive of the European Parliament and of the Council of 19 March 2001 on the interoperability of the trans-European conventional rail system.* OJ L110/1.

European Commission. (2001b). *2001/14/EC Directive of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification.* OJ L075/29. From: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32001L0014:en:HTML Accessed: 19 December 2012.

European Commission. (2003). *COMMISSION REGULATION (EC) No 1192/2003 of 3 July 2003 amending Regulation (EC) No 91/2003 of the European Parliament and of the Council on rail transport statistics.* L 167/13.

European Commission. (2004). *2004/49/EC Directive of the European Parliament and of the Council of 29 April 2004 - Railway Safety Directive.* OJ L220/16.

European Commission. (2007). *2007/59/EC Directive of the European Parliament and of the Councli on the certification of train drivers operating locomotives and trains on the railway system in the Community.* OJ L315/51.

European Commission. (2008). *2008/164/EC Decision of 21 December 2007 concerning the technical specification of interoperability relating to 'persons with reduced mobility' in the trans-European conventional and high-speed rail system.* OJ L64/72.

European Commission. (2009). *2009/149/EC Directive of 27 November 2009 amending Directive 2004/49/EC of the European Parliament and of the Council as regards Common Safety Indicators and common methods to calculate accident costs.* OJ L313/65.

European Commission. (2013). *Road fatalities in the EU since 2011* [Online]. From: http://ec.europa.eu/transport/road_safety/specialist/statistics/ Accessed: 11 June 2013.

European Railway Agency. (2008). *Railway safety performance in the European Union*. Lille, France: European Railway Agency.

European Railway Agency. (2009). *Glossary of Railway Terms*. European Railway Agency. From: http://www.era.europa.eu/Document-Register/Pages/Glossary-of-railway-terms.aspx Accessed: 15 December 2012.

European Railway Agency. (2010). *Application guide for the design and implementation of a Railway Safety Management System*. Valenciennes - France: European Railway Agency.

European Railway Agency. (2011). *Railway safety performance in the European Union*. Lille, France: European Railway Agency.

European Railway Agency. (2012a). *Common Safety Indicators Data for 2011*. From: http://www.era.europa.eu/Document-Register/Pages/Common-Safety-Indicators-data-for-2011.aspx Accessed: 10 May 2013.

European Railway Agency. (2012b). *Level crossing safety in the European Union*. European Railway Agency. From: http://www.era.europa.eu/Document-Register/Documents/Level_crossing_safety_EU_2012.pdf.

European Railway Agency. (2012c). *Railway safety performance in the European Union*. Lille, France: European Railway Agency.

European Railway Agency. (2013a). *Implementation Guidance for CSIs, Annex 1 of Directive 2004/49/Ec as Amended by Directive 2009/149/Ec*. Lille, France: European Railway Agency.

European Railway Agency. (2013b). *Investigation Reports Advanced Search* [Online]. European Railway Agency. From: https://pdb.era.europa.eu/safety_docs/naib/search.aspx Accessed: 25 February 2013.

European Railway Agency Database of Interoperability and Safety. (2012). *Accident Investigation Notification and Reports Database* [Online]. From: https://pdb.era.europa.eu/safety_docs/naib/default.aspx.

Eurostat, International Transport Forum & United Nations Economic Commission for Europe. (2009). *Glossary for Transport Statistics*.

Evans, A. W. (2011). Fatal train accidents on Europe's railways: 1980-2009, *Accident Analysis and Prevention*, **43**, 391-401. 10.1016/j.aap.2010.09.009.

Federal Railroad Administration. (2003). *FRA Guide for Preparing Accident/Incident Reports*. U.S. Department of Transportation.

Federal Railroad Administration. (2007). *Collision Hazard Analysis Guide: Commuter and Intercity Passenger Rail Service*. Washington, DC: U.S. Department of Transportation.

Federal Railroad Administration. (2008). *Rail Trespasser Fatalities Developing Demographic Profiles*.

Federal Railroad Administration. (2011). *FRA Guide for Preparing Accident/Incident Reports*. U.S. Department of Transportation.

Federal Railroad Administration. (2012). *Ten year accident / incident overview by railroad* [Online]. Federal Railroad Administration. From: http://safetydata.fra.dot.gov/officeofsafety/publicsite/Query/tenyr1a.aspx.

Federal Railroad Administration. (2013). *Train Accident Cause codes* [Online]. Federal Railroad Administration,. From: http://safetydata.fra.dot.gov/officeofsafety/publicsite/downloads/appendixC-TrainaccidentCauseCodes.aspx?State=0 Accessed: 25 February 2013.

Ferguson, I. & Anderson, N. (2010). The ORR Railway Management Maturity Model and its Use in Benchmarking Safety and Securing Continued Improvement. *20th International Railway Safety Conference*, Hong Kong, China.

Field, A. (2009). *Discovering Statistics Using SPSS*. SAGE Publications Ltd. 3rd ed.

Fleming, M. (2000). *Safety Culture Maturity Model*. The Keil Centre. Hse.

Forester, J., Kolaczkowski, A., Cooper, S., Bley, D. & Lois, E. (2007). *ATHEANA user's guide*. Washington, DC: United States Nuclear Regulatory Commission.

Forester, J., Kolaczkowski, A., Lois, E. & Kelly, D. (2006). *Evaluation of human reliability analysis methods against good practices*. United States Nuclear Regualtory Commission.

Forester, J., Ramey-Smith, A., Bley, D., Kolaczkowski, A. & Cooper, S. (1998). *Discussion of comments from a peer review of a technique for human event analysis (ATHEANA)*. Seattle, Washington: United States Nuclear Regulatory Commission.

Forsdike, N., Turner, C., Bellerby, F., Nelson, S. & Mcguire, P. (2007). *Improving the content and Placement of Anti-Trespass Signs*. In: John R. Wilson, Beverley Norris, Theresa Clarke & Ann Mills (eds.) *People and Rail Systems Human Factors at the Heart of the Railway*. Ashgate. pp. 37-47.

French, P. W. (2008). *U.S. Railroad Safety Statistics and Trends*. Association of American Railroads.

Garner, T., Newman, M., Lowe, C. & Hamilton, I. W. (2007). *Prediction of Signaller Workload*. In: John R. Wilson, Beverley Norris, Theresa Clarke & Ann Mills (eds.) *People and Rail Systems Human Factors at the Heart of the Railway*. Hampshire, England: Ashgate. 237-245.

Gerdes, V. (1995). HRA techniques; A selection matrix, *Microelectronics Reliability*, **35**, 1215-1231. http://dx.doi.org/10.1016/0026-2714(95)99373-Q.

Gertman, D. & Blackman, H. (1993). *Human Reliability and Safety Analysis Data Handbook*. Wiley.

Gertman, D., Blackman, H., Marble, J., Byers, J. & Smith, C. (2004). *The SPAR-H human reliability analysis method*. Idaho National Laboratory. Washington, DC: United States Nuclear Regulatory Commission.

Gervásio, H. & Simões Da Silva, L. (2012). A probabilistic decision-making approach for the sustainable assessment of infrastructures, *Expert Systems with Applications*, **39**, 7121-7131. http://dx.doi.org/10.1016/j.eswa.2012.01.032.

Gibson, H. W. & Kirwan, B. (2008a). Application of the CARA HRA Tool to Air Traffic Management Safety Cases. *9th International Conference on Probabilistic Safety Assessment and Management (PSAM9)*, Hong Kong, China.

Gibson, H. W., Mills, A., Smith, S. & Kirwan, B. K. (2013a). *Railway Action Reliability Assessment A Railway - Specific Approach to Human Error Quantification*. In: Nastaran Dadashi, Anita Scott, John R. Wilson & Ann Mills (eds.) *Rail Human Factors Supporting Reliability, Safety and Cost Reduction*. London, UK: Taylor & Francis Group. 671-676.

Gibson, H. W., Smith, S., Lowe, E., Mills, A., Morse, G. & Carpenter, S. (2013b). *Incident Factor Classification System*. In: Nastaran Dadashi, Anita Scott, John R. Wilson & Ann Mills (eds.) *Rail Human Factors Supporting Reliability, Safety and Cost Reduction*. London, UK: Taylor & Francis Group. 683-691.

Gibson, H. W., Walker, G., Stanton, N. A. & Baber, C. (2004). *Event Analysis of Systemic Teamwork Methodology for Railway Data: Protection Scenario*. Human Factors Integration Defence Technology Centre.

Gibson, W. H. & Kirwan, B. (2008b). Current Trends in Human Reliability Assessment. In: Taylor & Francis, ed. *International Conference on Contemporary Ergonomics (CE2008)*. Nottingham, UK. http://www.crcnetbase.com/doi/abs/10.1201/9780203883259.ch25, 158-163.

Greening, L. A. & Bernow, S. (2004). Design of coordinated energy and environmental policies: use of multi-criteria decision-making, *Energy Policy*, **32**, 721-735. http://dx.doi.org/10.1016/j.enpol.2003.08.017.

Groth, K. M. (2009). *A data-informed model of performance shaping factors for use in human reliability analysis.* PhD thesis. Department of Mechanical Engineering. University of Maryland. Maryland.

Guldenmund, F. W. (2000). The nature of safety culture: a review of theory and research, *Safety Science*, **34**, 215-257.

Gunderson, S. (2005). A Review of Organizational Factors and Maturity Measures for System Safety Analysis, *Systems Enginnering*, **8**, 10.1002/sys.20033.

Guy, A. & Rees, J. (2011). *Early Railways 1569-1830*. Great Britain: Shire Publications Ltd.

Hall, S. (2005). *Modern Signalling Handbook*. Surrey, UK: Ian Allan Publishing. Third ed.

Hallbert, B., Whaley, A., Boring, R., Mccabe, P. & Chang, Y. (2007). *Human Event Repository and Analysis (HERA): The HERA Coding Manual and Quality Assurance*. Idaho National Laboratory. Office of Nuclear Regulatory Research U.S. Nuclear Regulatory Commission.

Hamilton, I. W. & Clarke, T. (2005). Driver performance modelling and its practical application to railway safety, *Applied Ergonomics*, **36**, 661-670.

Hammerl, M., De Filippis, M., Steinhäuser, I., Torens, C., Gantz, O., Meyer Zu Hörste, M. & Lemmer, K. (2009). From a Testing Laboratory for Railway Technical Components to a Human Factors Simulation Environment. *3rd International Rail Human Factors Conference*, Lille, France.

Hammerl, M. & Vanderhaegen, F. (2009). Human Factors in Railway System Safety Analysis Process. *3rd International Human Factors Conference* Lille, France.

Härmä, M., Sallinen, M., Ranta, R., Mutanen, P. & Müller, K. (2002). The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers, *Journal of Sleep Research*, **11**, 141-151.

Health and Safety Executive. (1999). *Reducing Error and Influencing Behaviour*. London: Health and Safety Executive. 2nd ed.

Health and Safety Executive. (2000). *The train collision at Ladbroke Grove 5 October 1999*. Health and Safety Executive Books.

Heinrich, H. W. (1931). *Industrial Accident Prevention: A Scientific Approach*. New York: McGraw-Hill.

Hidden, A. (1989). *Investigation into the Clapham Junction Railway Accident*. London: Her Majesty's Stationery Office.

Hinkle, D. E., Wiersma, W. & Jurs, S. G. (1979). *Applied Statistics for the Behavioral Sciences*. Chicago, USA: Rand McNally College Publishing Company.

Hirsch, R. (2006). Reducing Risk by Probabilistic Assessment, Defence in Depth and Precursor Monitoring. *China International Railway and Metro Safety Conference*, Imperial College, CoMET.

Hirschberg, S. (2004). *Human Reliability Analysis in Probabilistic Safety Assessment for Nuclear Power Plants*. Organisation for Economic Co-Operation and Development Nuclear Energy Agency.

Hoffrichter, A. (2009). *Aspects of Railway Electrification in Britain.* School of Engineering and Applied Science. Aston University. Birmingham.

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method*. Elsevier.

Hollnagel, E. (2004). *Barriers and Accident Prevention*. Aldershot, UK: Ashgate.

Hollnagel, E. (2012). *FRAM: The Functional Resonance Analysis Method - Modelling Complex Socio-Technical Systems*. Ashgate.

Hollnagel, E., Woods, D. D. & Leveson, N. (eds.). (2006). *Resilience Engineering: Concepts and Precepts*: Ashgate Publishing Company.

House, R. J. & Shamir, B. (1993). *Toward the integration of transformational, charismatic, and visionary theories.* In: Martin M. Chemers & Roya Ayman (eds.) *Leadership theory and research: Perspectives and directions*. San Diego, CA, US: Academic Press. pp. 81-107.

Huang, J. (2009). *Aviation Safety through the Rule of Law*. Aviation Law and Policy Series Series.: Kluwer Law International.

Humphrey, W. S. (1989). *Managing the Software Process*. Addison-Wesley Professional.

Hwang, C.-L. & Yoon, K. (1981). *Multiple attribute decision making*. Springer.

Independent Transport Safety Regulator. (2011). *Managing signals passed at danger*. Sydney, Australia.

International Organization for Standardization. (2005). *ISO 9000:2005 - Quality management systems - Fundamentals and vocabulary*. International Organization for Standardization.

International Union of Railways. (2007). *Safety Database Activity Report 2007 - Significant Accidents 2006 - Benchmarking and Appendix*. Paris, France.

International Union of Railways. (2008). *Safety Database Activity Report 2008 - Significant Accidents 2007 - Benchmarking and Appendix*. Paris, France.

International Union of Railways. (2009). *Safety Database Activity Report 2009 - Significant Accidents 2008 - Benchmarking and Appendix*. Paris, France.

International Union of Railways. (2010). *Safety Database Activity Report - Significant Accidents 2009 - Benchmarking and Appendix*. Paris, France: Raffaele Iannone.

International Union of Railways. 2011a. Railway Statistics 2010 Synopsis. Published: Paris, France.

International Union of Railways. (2011b). *Safety Database Activity Report - Significant Accidents 2010 - Benchmarking and Appendix*. Paris, France.

International Union of Railways. (2012). *Safety Database Activity Report 2012 - Significant Accidents 2011*. Paris, France.

Kariuki, S. G. (2007). *Integrating human factors into chemical process quantitative risk analysis* Plant and Safety Technology. TU Berlin. Berlin.

Karsak, E. E., Sozer, S. & Alptekin, S. E. (2003). Product planning in quality function deployment using a combined analytic network process and goal programming approach, *Computers & Industrial Engineering*, **44**, 171-190. http://dx.doi.org/10.1016/S0360-8352(02)00191-2.

Keeney, R. L. & Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. New York, NY, USA: Wiley.

Khan, F. I., Amyotte, P. R. & Dimattia, D. G. (2006). HEPI: A new tool for human error probability calculation for offshore operation, *Safety Science*, **44**, 313-334. 10.1016/j.ssci.2005.10.008.

Kim, J., Jung, W., Jang, S.-C. & Wang, J.-B. (2006). A Case Study for the Selection of a Railway Human Reliability Analysis Method. *16th International Railway Safety Conference*, Dublin, Ireland.

Kim, J. W. & Jung, W. (2003). A taxonomy of performance influencing factors for human reliability analysis of emergency tasks, *Journal of Loss Prevention in the Process Industries*, **16**, 479-495. 10.1016/S0950-4230(03)00075-5.

Kirwan, B. (1994). *A Guide to Practical Human Reliability Assessment*. Taylor & Francis.

Kirwan, B. (1997). The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI, *Reliability Engineering & System Safety*, **56**, 107-133. http://dx.doi.org/10.1016/S0951-8320(97)00006-9.

Kirwan, B. & Gibson, H. W. (2007). CARA: A Human Reliability Assessment Tool for Air Traffic Safety Management - Technical Basis and Preliminary Architecture. *15th Safety-Critical Systems Symposium*. Bristol, UK. Springer, pp. 197-214.

Kolaczkowski, A., Forester, J., Lois, E. & Cooper, S. (2005). *Good Practices for Implementing Human Reliability Analysis*. Washington, DC: Office of Nuclear Regulatory Research U.S. Nuclear Regulatory Commission.

Kumamoto, H. & Henley, E. J. (1996). *Probabilistic Risk Assessment and Management for Engineers and Scientists*. New York: IEEE Press. 2nd ed.

Kyriakidis, M. (2009). *A scoping method for human performance integrity and reliability assessment in process industries.* Laboratory for Safety Analysis, Institute for Energy Technology, D-MAVT. ETH Zurich. Zurich.

Kyriakidis, M. (2011). *Human Reliability Analysis*. In: Wolfgang Kröger & Enrico Zio (eds.) *Vulnerable Systems*. Springer. pp. 157-187.

Kyriakidis, M., Majumdar, A., Grote, G. & Ochieng, W. Y. (2012). Development and Assessment of Performance Shaping Factors Taxonomy for Railway Operations, *Transportation Research Record: Journal of the Transportation Research Board*, **2289 / 2012**, 145-153. 10.3141/2289-19.

Lee, M.-C. (2010). *The Analytic Hierarchy and the Network Process in Multicriteria Decision Making: Performance Evaluation and Selecting Key Performance Indicators Based on ANP Model*. In: Marius Crisan (ed.) *Convergence and Hybrid Information Technologies*. InTech. From: http://www.intechopen.com/books/convergence-and-hybrid-information-technologies/the-analytic-hierarchy-and-the-network-process-in-multicriteria-decision-making-performance-evaluati

Leiden, K., Keller, J. & French, J. (2001). *Context of Human Error in commercial Aviation*. Ames Research Center. Boulder, CO: National Aeronautics and Space Administration System-Wide Accident Prevention Program.

Leveson, N. (2004). A new accident model for engineering safer systems, *Safety Science*, **42**, 237-270. http://dx.doi.org/10.1016/S0925-7535(03)00047-X.

Løken, E. (2007). Use of multicriteria decision analysis methods for energy planning problems, *Renewable and Sustainable Energy Reviews*, **11**, 1584-1595. http://dx.doi.org/10.1016/j.rser.2005.11.005.

Lombardi, P. L., Lami, I. M., Bottero, M. & Grasso, C. (2007). Application of the Analytic Network Process and the Multi-modal framework to an urban upgrading case study. *International Conference on Whole Life Urban Sustainability and its Assessment*. Glasgow. http://download.sue-mot.org/Conference-2007/Papers/Lombardi1.pdf.

Lunt, J. & Hartley, R. (2004). *Literature Review of Post Traumatic Stress Disorder amongst Rail Workers*. Sheffield, UK: Health and Safety Laboratory. HSL/2004/16.

Luther, R. E. & Johnson, C. E. (2008). Culture Management in the U.K. Rail Industry. *3rd International Conference on System Safety*, Birmingham, UK.

Lyons, M., Adams, S., Woloshynowych, M. & Vincent, C. (2004). Human reliability analysis in healthcare: A review of techniques, *International Journal of Risk & Safety in Medicine*, **16**, 223-237.

Matthews, G., Davies, R. D., Westerman, S. J. & Stammers, R. S. (2000). *Human Performance: Cognition, Stress and Individual Differences*. Psychology Press.

Maurino, D. (1998). ICAO supports proactive approach to managing human factors issues related to advanced technology, *ICAO Journal*.

May, S. & Mumby, D. K. (eds.). (2005). *Engaging Organizational Communication Theory and Research - Multiple Perspectives*: SAGE Publications, Inc.

Mccrum-Gardner, E. (2008). Which is the correct statistical test to use?, *The British journal of oral & maxillofacial surgery*, **46**, 38-41.

Mcleod, R. W., Walker, G. H. & Moray, N. (2005). Analysing and modelling train driver performance, *Applied Ergonomics*, **36**, 671-680. 10.1016/j.apergo.2005.05.006.

Meade, L. M. & Presley, A. (2002). R&D project selection using the analytic network process, *Engineering Management, IEEE Transactions on*, **49**, 59-66. 10.1109/17.985748.

Meister, D. (1964). Methods of Predicting Human Reliability in Man–Machine Systems, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, **6**, 621-646. 10.1177/001872086400600605.

Meister, D. (1989). *Conceptual aspects of human factors*. Baltimore, MD: Johns Hopkins University Press.

Mendoza, G. A. & Martins, H. (2006). Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms, *Forest Ecology and Management*, **230**, 1-22. http://dx.doi.org/10.1016/j.foreco.2006.03.023.

Murrell, K. F. H. (1965). *Human performance in industry*. New York: Reinhold Publishing.

Murtagh, V. (2011). *Railway systems and safety - a personal view*. CTS Seminars. Imperial College London.

National Rail Enquiries. (2012). *Train Operating Companies* [Online]. From: http://www.nationalrail.co.uk/tocs_maps/tocs/TrainOperators.aspx Accessed: 09 September 2012.

National Transportation Safety Board. (2010). *Railroad Accident Report: Collision of Metrolink Train 111 With Union Pacific Train LOF65-12 Chatsworth, California September 12, 2008*. Washington, D.C., USA. NTSB/RAR-10/01.

Nelsen, R. B. (n.d.). *Kendall tau metric. Encyclopedia of Mathematics.* [Online]. Springer. From: http://www.encyclopediaofmath.org/index.php/Kendall_tau_metric Accessed: 15 October 2012.

Network Rail. (2008a). *National operations centre controller job description*. National Rail.

Network Rail. (2008b). *National operations signaller job description*. Network Rail.

Network Rail. (2008c). *National operations train driver job description*. Network Rail.

Network Rail. (2012a). *Move my freight by rail* [Online]. From: http://www.networkrail.co.uk/aspx/10442.aspx Accessed: 09 September 2012.

Network Rail. (2012b). *Trespassing and vandalism* [Online]. London, UK: Network Rail. From: http://www.networkrail.co.uk/aspx/1037.aspx Accessed: 15 January 2013.

Nock, O. S. (1975). *Railways Then and Now. A World History*. London: Paul Elek Ltd.

Office of Rail Regulation. (2011a). *Railway Management Maturity Model (RM3)*. Office of Rail Regulation.

Office of Rail Regulation. (2011b). *Safety certificates and authorisation* [Online]. London, UK: Office of Rail Regulation. From: http://www.rail-reg.gov.uk/server/show/ConWebDoc.8735 Accessed: 14 September 2011.

Office of Rail Regulation. (2011c). *Signals passed at danger (SPADs)* [Online]. From: http://www.rail-reg.gov.uk/server/show/nav.1137 Accessed: 28 October 2012.

ÖZtüRk, Z. K. L. (2006). A review of multi criteria decision making with dependency between criteria. *18th International Conference on Multiple Criteria Decision Making*, Chania, Greece.

Paraskevopoulos, K. (n.d.). *The Analytic Hierarchy Process (AHP)*. International Hellenic University. From: http://rad.ihu.edu.gr/fileadmin/labsfiles/decision_support_systems/lessons/ahp/AHP_Lesson_1.pdf.

Park, K. S. & Lee, J. I. (2008). A new method for estimating human error probabilities: AHP-SLIM, *Reliability Engineering & System Safety*, **93**, 578-587. 10.1016/j.ress.2007.02.003.

Pasquini, A., Rizzo, A. & Save, L. (2004). A methodology for the analysis of SPAD, *Safety Science*, 437-455.

Paulk, M. C., Weber, C. V., Curtis, B. & Chrissis, M. B. (1995). *The Capability Maturity Model: Guidelines for Improving the Software Process*. Carnegie Mellon University

Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.

Pirdashti, M., Tavana, M., Hassim, H. M., Behzadian, M. & Karimi, I. A. (2011). A taxonomy and review of the multiple criteria decision-making literature in chemical engineering, *International Journal of Multicriteria Decision Making*, **1**, 407-467. 10.1504/IJMCDM.2011.043556.

Priestley, K. & Lee, G. (2008). Human Factors in Railway Operations. *International Conference on Challenges for Railway Transportation in Information Age*, Hong Kong, China.

Promentilla, M. a. B., Furuichi, T., Ishii, K. & Tanikawa, N. (2008). A fuzzy analytic network process for multi-criteria evaluation of contaminated site remedial countermeasures, *Journal of Environmental Management*, **88**, 479-495. http://dx.doi.org/10.1016/j.jenvman.2007.03.013.

Pyy, P. (2000). *Human Reliability Analysis Methods for Probabilistic Safety Assessment*. Technical Research Centre of Finland. Finland: Vtt Publications.

Qureshi, Z. H. (2007). A Review of Accident Modelling Approaches for Complex Socio-Technical Systems. In: Tony Cant, ed. *12th Australian Conference on Safety-Related Programmable Systems (SCS 2007)*. Adelaide, Australia. http://crpit.com/confpapers/CRPITV86Qureshi.pdf: ACS, 47-59.

Rail Accident Investigation Board. (2003). *Signals Passed at Danger (SPADs)*. GO/RT3252.

Rail Accident Investigation Board. (2009). *Near-miss at Hanger Lane junction*. From: http://www.raib.gov.uk/publications/investigation_reports/reports_2010/report052010.cfm.

Rail Safety and Standards Board. (2004a). *Common factors in SPADs*. London, UK. HEL/RSSB/03984a/RT2.

Rail Safety and Standards Board. (2004b). *Rail-Specific Human Reliability Assessment Technique for Driving Tasks*. London, UK.

Rail Safety and Standards Board. (2004c). *Road Vehicle Level Crossings*. London, UK.

Rail Safety and Standards Board. (2004d). *Train accidents and their precursors - annual safety performance report 2004*. London, UK.

Rail Safety and Standards Board. (2004e). *Understanding human factors: a guide for the railway industry*. London, UK.

Rail Safety and Standards Board. (2006). *Human Factors Study of Fatigue and Shift Work*. London, UK.

Rail Safety and Standards Board. (2009). *Driver Advisory Information for Energy Management and Regulation - Stage 1*. London, UK.

Rail Safety and Standards Board. (2011). *Investigation Guidance Part 3 Practical Support for Accident Investigators*. London, UK.

Rail Safety and Standards Board. (2012a). *Non-technical skills for rail: A list of skills and behavioural markers for drivers with guidance notes*. London, UK.

Rail Safety and Standards Board. (2012b). *Railway Action Reliability Assessment user manual - A technique for the quantification of human error in the rail industry*. London, UK. From: http://www.rssb.co.uk/sitecollectiondocuments/pdf/reports/research/T270-guide-final4.pdf.

Rail Safety and Standards Board. (2013). *Annual Safety Performance Report 2012/13*. London, UK. From: http://www.rssb.co.uk/SPR/REPORTS/Documents/ASPR_2012-13_FullReport.pdf.

Railway Association of Canada. (2011). *Public Safety* [Online]. From: http://www.railcan.ca/safety/public_safety Accessed: 18 November 2012.

Railway Technical Web Pages. (2012a). *Automatic Train Operation* [Online]. From: http://www.railway-technical.com/sigtxt4.shtml Accessed: 15 March 2013.

Railway Technical Web Pages. (2012b). *Infrastructure* [Online]. From: http://www.railway-technical.com/track.shtml.

Ransom, P. J. G. (1981). *The Archaeology of Railways*. Surrey: World's Work Ltd.

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations, *Journal of Occupational Accidents*, **4**, 311-333. 10.1016/0376-6349(82)90041-4.

Razali, N. M. & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests, *Journal of Statistical Modeling and Analytics*, **2**, pp. 21-33.

Reason, J. (1990). *Human Error*. Cambridge University Press.

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate Publishing Ltd.

Reason, J. (2000). Human error: models and management, *BMJ*, **320**, 768-770. 10.1136/bmj.320.7237.768.

Reer, B. (2008). Review of advances in human reliability analysis of errors of commission, Part1: EOC identification, *Reliability Engineering & System Safety*, **93**, 1091-1104.

Reinach, S. & Viale, A. (2006). Application of a human error framework to conduct train accident/incident investigations, *Accident Analysis and Prevention*, **38**, 396-406. 10.1016/j.aap.2005.10.013.

Reinach, S., Viale, A. & Donald, G. (2007). *Human Error Investigation Software Tool (HEIST)*. Washington, DC, U.S.A.: Federal Railroad Administration.

Revill, G. (2012). *Railway*. Object Series. London: Reaktion Books Ltd.

Robison, M. (2012). *On track with diversity*. London, UK: Aslef & the Institute of Employment Rights.

Roth, E. M. & Malsch, N. (1999). *Understanding How Train Dispatchers Manage and Control Trains Results of a Preliminary Cognitive Task Analysis*. Federal Railroad Administration. Cambridge, MA., USA. DOT/FRA/ORD-99/XX.

Roy, B. & Bouyssou, D. (1986). Comparison of two decision-aid models applied to a nuclear plant siting example, *European Journal of Operational Research*, **25**, 200-215.

Rumsey, A. (2009). *IRSE seminar on communications based train control*. London, UK: Institution of Railway Signal Engineers.

Ryall, J. (2012). *310 mph 'floating' trains unveiled in Japan* [Online]. Tokyo: The Telegraph. From: http://www.telegraph.co.uk/news/worldnews/asia/japan/9702521/310mph-floating-trains-unveiled-in-Japan.html Accessed: 26 November 2012.

Ryan, B., Wilson, J. R., Sharples, S., Morrisroe, G. & Clarke, T. (2009). Developing a Rail Ergonomics Questionnaire (REQUEST), *Applied Ergonomics*, **40**, 216-229.

Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used, *Mathematical Modelling*, **9**, 161-176. http://dx.doi.org/10.1016/0270-0255(87)90473-8.

Saaty, R. W. (2003). *Decision Making in Complex Environments The Analytic Hierarchy Process (AHP) for Decision Making and The Analytic Network Process (ANP) for Decision Making with Dependence and Feedback*. Pittsburgh, PA, USA: Creative Decisions Foundation. From: http://www.croce.ggf.br/dados/Tutorial superdecisons.pdf.

Saaty, T. (2004). Fundamentals of the analytic network process — Dependence and feedback in decision-making with a single network, *Journal of Systems Science and Systems Engineering*, **13**, 129-157. 10.1007/s11518-006-0158-y.

Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York, NU, USA: McGraw-Hill.

Saaty, T. L. (1996). *Decision Making with Dependence and Feedback: The Analytic Network Process*. RWS. 2001 ed.

Saaty, T. L. (2008). The Analytic Hierarchy and Analytic Network Measurement Processes: Applications to Decisions under Risk, *European Journal of Pure and Applied Mathematics*, **1**, pp. 122-196.

Saaty, T. L. & Vargas, L. G. (2006). *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks*. International Series in Operations Research & Management Science Series.: Springer.

Salant, P. & Dillman, D. A. (1994). *How to conduct your own survey*. New York, USA: John Wiley and Sons.

Salmon, P. M., Stanton, N. A., Lenné, M., Jenkins, D. P., Rafferty, L. & Walker, G. H. (2011). *Human Factors Methods and Accident Analysis Practical Guidance and Case Study Applications*. Ashgate.

Sasou, K. & Reason, J. (1999). Team errors: definition and taxonomy, *Reliability Engineering & System Safety*, **65**.

Shappell, S. A. & Wiegmann, D. A. (2000). *The human factors analysis and classification system-HFACS*. Federal Aviation Administration Office of Aviation Medicine. DOT/FAA/AM-00/7.

Shaver, E. (2009). *The Many Definitions of Human Factors and Ergonomics* [Online]. From: http://www.thehumanfactorblog.com/2009/03/16/the-many-definitions-of-human-factors-and-ergonomics/ Accessed: 8 April 2012.

Sheskin, D. J. (2007). *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman & Hall/CRC.

Shorrock, S. T. & Kirwan, B. (2002). Development and application of a human error identification tool for air traffic control, *Applied Ergonomics*, **33**, 319-336.

Smith, P., Kyriakidis, M., Majumdar, A. & Ochieng, W. Y. (2013). The Impact of ERTMS on Human Performance in Railway Operations - European Findings, *Transportation Research Record: Journal of the Transportation Research Board - in press*.

Sorensen, J. N. (2002). Safety culture: a survey of the state-of-the-art, *Reliability Engineering & System Safety*, **76**, 189-204.

Stanton, N., Hedge, A., Brookhuis, K., Salas, E. & Hendrick, H. (eds.). (2005a). *Handbook of Human Factors and Ergonomics Methods*: CRC Press LLC.

Stanton, N. A. (ed.) (2012). *Advances in Human Aspects of Road and Rail Transportation*: CRC Press.

Stanton, N. A., Salmon, P. M., Walker, G. H., Baber, C. & Jenkins, D. P. (2005b). *Human Factors Methods: A Practical Guide for Engineering and Design*. Ashgate Publishing, Limited.

Sträter, O. (2000). *Evaluation of human reliability on the basis of operational experience.* PhD Thesis. Faculty of Economics and Social Sciences. Munich Technical University. Munich, Germany.

Subotic, B. (2007). *Framework for the analysis of controller recovery from equipment failures in air traffic control.* Department of Civil and Environmental Enginnering. Imperial College London. London.

Svedung, I. & Rasmussen, J. (2002). Graphic representation of accident scenarios: mapping system structure and the causation of accidents, *Safety Science*, **40**, 397-417. http://dx.doi.org/10.1016/S0925-7535(00)00036-9.

Swain, A. D. (1990). Human Reliability Analysis: Need, Status, Trends and Limitations, *Reliability Engineering & System Safety*, **29**, 301-313.

Swain, A. D. & Guttmann, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. United States Nuclear Regulatory Commission.

Swiss Federal Railways. (2010). *Fragebogen Ereignisanalyse Lokführer*. Swiss Federal Railways.

Swissinfo. (2013). *Safety questions raised following train crash* [Online]. From: http://www.swissinfo.ch/eng/swiss_news/Safety_questions_raised_following_train_crash.html?cid=36602564 Accessed: 07 August 2013.

Taslicali, A. K. & Ercan, S. (2006). THE ANALYTIC HIERARCHY & THE ANALYTIC NETWORK PROCESSES IN MULTICRITERIA DECISION MAKING: A COMPARATIVE STUDY, *Journal of Aeronautics and Space Technologies*, **2**, pp. 55-65.

Topham, G. (2013). *Railway accidents happen because someone makes a mistake* [Online]. The Guardian. From: http://www.theguardian.com/uk-news/2013/jul/25/railway-accidents-human-error-warning-systems Accessed: 27 July 2013.

Transportation Safety Board of Canada. (2012). *Statistical Summary Railway Occurences 2011*. Quebec, Canada: Transportation Safety Board of Canada.

Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A Comparative Study*. Kluwer Academic Publishers.

Valonen, K. (2000). *Investigation of rail accidents - International comparison.* Mechanical Enginnering. Helsinki University of Technology. Helsinki.

Vanderhaegen, F. (2001). A non probabilistic prospective and retrospective HRA method – application to railway system, *Reliability Engineering & System Safety*, **71**, 1-13.

Wang, R. Y. & Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers, *J. Manage. Inf. Syst.*, **12**, 5-33.

Weisberg, H. F., Krosnick, J. A. & Bowen, B. D. (1996). *An Introduction to survey research, polling, and data analysis* Thousand Oaks, CA: Sage. 3rd ed.

Wickens, C. D. (1992). *Engineering psychology and human performance*. New York: Harper Collins. 2nd ed.

Williams, J. C. (1986). HEART - A proposed method for assessing and reducing human error. *9th Advances in Reliability Technology Symposium*, University of Bradford.

Williams, J. C. (1988). A Data-Based Method for Assessing and Reducing Human Error to Improve Operational Procedures. *4th IEEE Conference: Human Factors and Power Plants*, Monterey, CA , USA, IEEE.

Wilson, J. R., Farrington-Darby, T., Cox, G. & Hockey, G. R. J. (2007a). The railway as a socio-technical system: human factors at the heart of successful rail engineering, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **221**, 101-115. 10.1243/09544097JRRT78.

Wilson, J. R., Norris, B., Clarke, T. & Mills, A. (eds.). (2005). *Rail Human Factors Supporting the Integrated Railway*: Ashgate.

Wilson, J. R., Norris, B., Clarke, T. & Mills, A. (eds.). (2007b). *People and Rail Systems Human Factors at the Heart of the Railway*: Ashgate.

Wilson, J. R. & Norris, B. J. (2005). Rail human factors: past, present and future, *Applied Ergonomics*, **36**, 649-660. 10.1016/j.apergo.2005.07.001.

Wilson, J. R. & Norris, B. J. (2006). Human factors in support of a successful railway: a review, *Cognition, Technology & Work*, 10.1007/s10111-005-0016-6.

Wolmar, C. (2007). *Fire & Steam. A New History of the Railways in Britain*. London: Atlantic Books.

Wreathall, J., Bley, D., Roth, E., Multer, J. & Raslear, T. (2004). Using an integrated process of data and modelling in HRA, *Reliability Engineering & System Safety*, **83**, pp. 221-228.

Wreathall, J., Roth, E., Bley, D. & Multer, J. (2003). *Human Reliability Analysis in Support of Risk Assessment for Positive Train Control*. Cambridge, MA, U.S.A.: Federal Railroad Administration.

Wright, K., Embrey, D. & Anderson, M. (n.d.). *Getting at the Underlying Systemic Causes of SPADS: A New Approach*. Wigan, UK: Human Relaibility Associates Ltd. From: http://www.humanreliability.com/articles/Getting at the underlying causes of SPADs.pdf.

Xu, L. & Yang, J.-B. (2001). *Introduction to multi-criteria decision making and the evidential reasoning approach*. Manchester School of Management.

Zhang, Z., Liu, X. & Yang, S. (2009). *A Note on the 1-9 Scale and Index Scale In AHP*. In: Yong Shi, Shouyang Wang, Yi Peng, Jianping Li, et al. (eds.) *Cutting-Edge Research Topics on Multiple Criteria Decision Making*. 35. Springer Berlin Heidelberg. 630-634. From: http://dx.doi.org/10.1007/978-3-642-02298-2_92

Zhou, X. (2012). *Fuzzy Analytical Network Process Implementation with Matlab*. MATLAB - A Fundamental Tool for Scientific Computing and Engineering Applications - Volume 3 Series. 39357.

# APPENDICES

Appendices .....**Error! Bookmark not defined.**
Appendix I Implementation of R-PSFs taxonomy .....334
Appendix II Overview of contextual factors categories .....335
Appendix III Overview of individual contextual factors.....339
Appendix IV Definitions of proposed R-PSFs .....345
Appendix V The Safety Maturity Model features .....360
Appendix VI Report on near-miss at Hanger Lane junction.....364
Appendix VII The Swiss questionnaire (English version) .....380
Appendix VIII Crosstabulations of reference database.....388
Appendix IX Loglinear analysis tests .....404
Appendix X Material to support SMEs common understanding of R-PSFs.....408
Appendix XI The R-PSFs “weighting” and “ranking” questionnaire .....411
Appendix XII The ratings of the R-PSFs per human error and classification group .....422

# Appendix I Implementation of the R-PSFs taxonomy



Network Operational Learning

Network Operational Learning
London Underground Limited
210-212 Baker Street
London NW1 5RT.

Miltos Kyriakidis
Skempton Bldg
Imperial College
London SW7 2AZ.

3rd July 2013

Dear Mr Kyriakidis,

I am writing to confirm that we are using your research findings for training operational managers investigating train incidents.

Firstly, it is used to demonstrate that research is ongoing, and is currently being carried out on the Underground.

The R-PSF taxonomy and definitions of factors are used to help delegates understand the range of potential root causes for incidents where signals have been passed at danger.

We also emphasize that the findings suggest that 90% of incidents are accounted for by a limited number of factors.

I believe it has been very beneficial to implement this research into our training material. I look forward to seeing the outcome of your research on London Underground.

Yours sincerely,

Tristram Pitt
Operational Trainer
Design Team.

MAYOR OF LONDON

# Appendix II Overview of contextual factors categories

## APPENDIX II OVERVIEW OF CONTEXTUAL FACTORS CATEGORIES

| Factor category | TRACE-RR | RARA | HFACS-RR | HERMES | HFs in Railways | R-PSFs |
|---|---|---|---|---|---|---|
| **1** | Infrastructure and traffic | Environment | | | | System |
| **2** | Communication | | | Communication within organisation | | Team |
| **3** | Procedures and documentation | Procedures | Organisational factors | Regulations/ rules | Organisational | Organisational |
| **4** | Information | | | | | Team |
| **5** | Training, knowledge and experience | Competence management | Personal factors | Training methods and simulators | Personal dependent | Personal, Organisational |
| **6** | Workplace design, HMI | Interface design | | Technological interfaces, Communication means | Design of HMI | System |
| **7** | In-cab environment | Environment | Environmental factors | Comfort of work context | Working conditions | System |
| **8** | Personal | Person | Conditions of operators | Comfort of work context | Personal individual | Personal, Dynamic personal |
| **9** | Social and team | | | | Organisational employee | Team |
| **10** | | | | Shifts planning | Organisational employee | Organisational |
| **11** | | Task design, Complexity | | | Organisational standard | Task |
| **12** | | Workload | | | | Task |
| **13** | | Staffing | | Roster | | Organisational |
| **14** | | Distraction | | | | Dynamic personal |

## APPENDIX II OVERVIEW OF CONTEXTUAL FACTORS CATEGORIES

| Factor category | THERP | HERA | ATHEANA | CARA | HERA-JANUS | R-PSFs |
|---|---|---|---|---|---|---|
| **1** | | | | Environment | Traffic | System |
| **2** | Situational characteristics | Communication | Communication | | Communication | Team |
| **3** | Work and task instructions | Procedures and reference documents, work processes | Procedures | Procedures | Documentation and procedures, Organisational | Organisational |
| **4** | | | | | | |
| **5** | Organismic factors | Experience and training | Training and experience | Competence management | Training and experience | Personal, Organisational |
| **6** | Task and equipment characteristics | Ergonomics and HMI | Ergonomics, HMI, Instrumentation | Interface design | Workplace design & HMI | System |
| **7** | Situational characteristics | Environment | Environment and working conditions | Environment | Weather, Environment | Environmental, System |
| **8** | Organismic factors | | Personal fitness level | Person | Personal | Personal, Dynamic personal |
| **9** | Situational characteristics | Team dynamics and characteristics | Dynamics and characteristics of team | | Team | |
| **10** | Work and task instructions | Available time | Time pressure | | | Task |
| **11** | Work and task instructions | Complexity | Complexity | Task design | | Task |
| **12** | Physiological & Psychological stressors | Stress & stressors | Workload and stress | | | Task, Dynamic personal |
| **13** | | | Staffing | | | Organisational |
| **14** | | | | | | |

## APPENDIX II OVERVIEW OF CONTEXTUAL FACTORS CATEGORIES

| Factor category | CREAM | HRMS | SPAR-H | SLIM | HEPI | R-PSFs |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | | | | Communication | | Team |
| 3 | Availability of procedures/ plans, Adequacy of organisation | Procedures | Work processes | Procedures and administrative controls | Procedures | Organisational |
| 4 | | Quality of information and interface | | | | Team |
| 5 | Adequacy of training and experience | Training/expertise/ experience/ competence | Training and experience | Training and experience | Training and experience | Personal, Organisational |
| 6 | Adequacy of MMI | | Ergonomics – HMI | Instrumentation, HMI | | System |
| 7 | Working conditions | | | Environmental factors | Environmental factors | Environmental |
| 8 | | | Fitness for duty | Fitness level | | Personal |
| 9 | Crew collaboration | | | Team dynamics and characteristics | Teamwork | Team |
| 10 | Available time, Time of day | Time | Available time | Time available | Time pressure | Task |
| 11 | Number of simultaneous goal | Task complexity, Task organisation | Task complexity | Complexity | Complexity | Task |
| 12 | | | Stress and stressors | Workload and stress | Stress | Task, Dynamic personal |
| 13 | Time of day | | Staffing | | | Organisational |
| 14 | | | | | Distraction | Dynamic personal |

# Appendix III Overview of individual contextual factors

| Individual Factors | TRACE-RR | RARA | HERMES | HFs in Railways |
|---|---|---|---|---|
| **1** | Unfamiliarity | Unfamiliarity | | |
| **2** | Time availability | Time availability | | |
| **3** | Ability to detect and perceive | | | |
| **4** | Features over-ride allowed | | | |
| **5** | Positioning and layout | | | Positioning and layout of HMI |
| **6** | High workload | Workload | | |
| **7** | Technique learning | Technique learning | | |
| **8** | Procedures clarity | Procedures | Regulations and rules | Standards, Rules and guidelines |
| **9** | Risk perception | Mismatch between perceived and real risk | | Safety awareness |
| **10** | Poor feedback | Interface | | Quality of feedback |
| **11** | Delayed/ Incomplete feedback | | | |
| **12** | Driver experience | Operator inexperience | | Experience |
| **13** | Information quality & availability | Quality of information, Interface | Communication | |
| **14** | Objectives conflict | Objectives conflict | | |
| **15** | No diversity | | | |
| **16** | Education and training | Operator inexperience | Training methods and simulators | Training provided by organisation, Skills |
| **17** | Trust in information | | | |
| **18** | Job clarity | Unclear allocation and responsibility | | |
| **19** | Physical capabilities | Physical capabilities | | |
| **20** | Emotional stress | Emotional stress | | Emotional tension |
| **21** | Health | Fitness for duty | | Health |
| **22** | Morale | Morale | | Motivation |
| **23** | Consistency of displays | | | |
| **24** | In-cab environment | Environment | Comfort of working contexts | Working conditions |
| **25** | Concentration | | | |
| **26** | Fatigue | Fatigue | | |
| **27** | Distraction / 3rd parties | Distraction | | |
| **28** | Team relations | | | Social aspects |
| **29** | No progress tracking | Checking | | |

| Individual Factors | TRACE-RR | RARA | HERMES | HFs in Railways |
|---|---|---|---|---|
| **30** | | Under load task | | |
| **31** | | Task complexity | | Task design |
| **32** | | | Communication means | |
| **33** | | | Technological interfaces | |
| **34** | | | Roster/shifts planning | |
| **35** | | | | Individual characteristics |
| **36** | | | | Roster planning, Tiredness |
| **37** | | | | Layout of work environment |
| **38** | | | | Leadership |
| **39** | | | | |
| **40** | | | | |
| **41** | | | | |
| **42** | | | | |

## APPENDIX III OVERVIEW OF INDIVIDUAL CONTEXTUAL FACTORS CATEGORIES

| Individual Factors | THERP | HEART | HERA | ATHEANA | CARA |
|---|---|---|---|---|---|
| **1** | Frequency and repetitiveness | Unfamiliarity | | | Unfamiliarity |
| **2** | Task speed | Time availability | Available time | Time pressure | |
| **3** | Perceptual requirements | | | | |
| **4** | | | | | |
| **5** | | | | | |
| **6** | | Low workload | | Workload | Cognitive overload |
| **7** | State of current practice or skills | Technique learning | | | Technique learning |
| **8** | Procedures | Procedures | Procedures and documents | Procedures | Procedures |
| **9** | High jeopardy risk | Mismatch between perceived and real risk | | | Risk taking |
| **10** | Feedback, HMI | Interface, Instrumentation | Ergonomics and HMI | Ergonomics, HMI, Instrumentation | Poor feedback, HMI, Instrumentation |
| **11** | Feedback | | | | |
| **12** | | Operator inexperience | Experience and training | Training and experience | Training and experience, On-job training |
| **13** | Communication, Complexity | Quality of information, Interface | Complexity, Communication | Communication | Trust in system, Hand-over practices |
| **14** | | Objectives conflict | | | |
| **15** | | | | | |
| **16** | Training and experience | Operator inexperience | Experience and training | Training and experience | Training and experience, On job training |
| **17** | | | | | |
| **18** | | Unclear allocation and responsibility | | | Team coordination |
| **19** | | Physical capabilities | | Personal fitness level | |
| **20** | Stress | Emotional stress | Stress & stressors | Stress | Emotional stress |
| **21** | Fitness for duty | Fitness for duty | | | |
| **22** | Rewards, Threats | Morale | | | Morale |
| **23** | Control-display relationship | Consistency of displays | | | |
| **24** | Working environment | Environment | Environment | Environment and working conditions | Working environment, Weather |
| **25** | | | | | |
| **26** | Fatigue | Fatigue | | | |
| **27** | Distraction | Distraction | | | |

| Individual Factors | THERP | HEART | HERA | ATHEANA | CARA |
|---|---|---|---|---|---|
| **28** | Team Relations, Team structure and communication | | Team dynamics and characteristics | Dynamics and characteristics of team | Communication |
| **29** | | Checking | | | Checking |
| **30** | Task load, Task criticality, Monotony | Under load task | | | |
| **31** | Task instructions | | Task complexity | Complexity | |
| **32** | | | | | |
| **33** | Availability/ adequacy of special equipment, tools and supplies | | | | |
| **34** | | | | | |
| **35** | Individual characteristics | | | | |
| **36** | Work hours/ work breaks, Shift Rotation, Staffing | | | Staffing | |
| **37** | | | | | |
| **38** | | | | | |
| **39** | Interpretation | | | | |
| **40** | Decision making | | | | |
| **41** | Vigilance | | | | Vigilance |
| **42** | | Safety culture | | | |

| Individual Factors | HERA-JANUS | CREAM | HRMS | SPAR-H | SLIM | HEPI |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | | Available time, Time of day | Time | Available time | Time available | Time pressure |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | Documentation and procedures | Availability of procedures/ plans, adequacy of organisation | Procedures | | Procedures and administrative controls | Procedures |
| 9 | | | | | | |
| 10 | Workplace design & HMI | Adequacy of MMI | | Ergonomics – HMI | Instrumentation, HMI | |
| 11 | | | | | | |
| 12 | Training and experience | | Training/ expertise/ experience/ competence | Training and experience | Training and experience | Training and experience |
| 13 | Communication | | Quality of information and interface | | Communication | |
| 14 | | | | | | |
| 15 | | | | | | |
| 16 | Training and experience | Adequacy of training and experience | Training/ expertise/ experience/ competence | Training and experience | Training and experience | Training and experience |
| 17 | | | | | | |
| 18 | | | | | | |
| 19 | Personal | | | | | |

| Individual Factors | HERA-JANUS | CREAM | HRMS | SPAR-H | SLIM | HEPI |
|---|---|---|---|---|---|---|
| **20** | | | | Stress and stressors | Workload and stress | Stress |
| **21** | | | | Fitness for duty | Fitness level | |
| **22** | | | | | | |
| **23** | | | | | | |
| **24** | Weather, Environment | Working conditions | | | Environmental factors | Environmental factors |
| **25** | | | | | | |
| **26** | | | | | | |
| **27** | | | | | | Distraction |
| **28** | Team, Communication | Crew collaboration | | | Team dynamics and characteristics | Teamwork |
| **29** | | | | | | |
| **30** | | | | | | |
| **31** | | Number of simultaneous goals | Task complexity, Task organisation | Task complexity | | |
| **32** | | | | | | |
| **33** | | | | | | |
| **34** | | | | | | |
| **35** | | | | | | |
| **36** | | Time of day | | Staffing | | |
| **37** | | | | | | |
| **38** | | | | | | |
| **39** | | | | | | |
| **40** | | | | | | |
| **41** | | | | | | |
| **42** | | | | | | |

# Appendix IV Definitions of proposed R-PSFs

## Personal Factors

*Personal* R-PSFs encompass the factors that characterise and affect the performance of an individual *but are not strongly* related to the precise moment of operation/occurrence. Therefore they are considered as static factors. Personal factors embrace operators' levels of training and experience, physical conditions and other intrinsic characteristics.

### Training-competence

*Training* is defined as the process of bringing a person to an agreed standard of proficiency by practice and instruction in order to conduct a particular job or activity. *Competence* is then defined as the control of the required knowledge, skills and attitudes to perform the specific range of tasks according to prescribed standards (Rail Safety and Standards Board, 2004e).

Well-trained and competent employees are those who are familiar and aware of new technologies, equipment and systems, while they regularly refresh their old skills and knowledge. For instance, train drivers who are adequately trained in order to use the GSM-R communication system.

### Experience

*Experience* is defined as the skills and knowledge that operators acquire from extensive participation and observation that is especially gained over the years of executing their profession. In particular, it is described as the accumulation of information and knowledge gained through training and interactions with the system (Groth, 2009).

An experienced signaller, for instance, is more likely to efficiently resolve the consequences of an unexpected event, such as a sudden signal failure. On the other hand, an inexperienced train driver may apply braking inappropriately when he/she operates on a new type of rolling stock (Rail Safety and Standards Board, 2004b). However it is not only the inexperienced drivers prone to errors. For example, high levels of experience may reduce the awareness of operators, which eventually may result to hazardous events. The level of operators' experience should be indicated every time that experience is identified as contributor R-PSF.

**Familiarity**

*Familiarity* is defined as the reasonable knowledge or acquaintance of operators with respect to their duties or tasks.

Familiarity may overlap with experience; however there is a clear difference between the two terms. For instance, although an operator may be well experienced, he/she may not necessarily be familiar with a particular situation, e.g. operating on a new route. Similar to experience, familiarity can either positively or negatively influence human performance, i.e. completely unfamiliar or too familiar.

In general, familiarity is associated with the frequency of conducting a specific task, operating on a route, applying a procedure or equipment. For example, dedicated to a line drivers or signallers who every day perform their duties on the same route.

**Fit to work**

*Fit to work* is defined as the ability of operators to perform their duties in a safe, secure, productive and effective manner, without facing any health, physical or mental, problems (Gertman et al., 2004). Stress or fatigue related problems are not included in this definition.

Fit to work should account for R-SPFs when people face long-term health problems or use any type of medication (with or without prescription) that may affect their performance.

**Motivation**

*Motivation* is described as the desire, interest or enthusiasm of operators to perform their duties.

Motivated employees are characterised by high willingness and increased effort to complete their tasks and positive attitude and way of thinking concerning their work (Groth, 2009). The opposite, for instance, may occur when operators are paid less in relation to operators from other comparable train operators or infrastructure managers.

**Individual characteristics**

*Individual characteristics* refer to all distinctive characteristics of employees.

The individual characteristics comprise features such as age, gender, physical capabilities, and cultural and behavioural characteristics, e.g. extroverted or introverted.

## Dynamic Personal Factors

*Dynamic personal* R-PSFs embrace the factors that characterise and affect individuals' performance *and are strongly* related to the precise moment of operation/occurrence. These factors include, amongst others, operators' levels of stress, distraction, fatigue and vigilance.

### Distraction-loss of concentration

*Distraction-loss of concentration* refers to any internal or external causes that draw operators' attention from concentrating on conducting their tasks and duties.

Amongst others, causes of distraction can be: other people in/around the cab or the control room, people standing on platforms, use of mobile phones, communication with co-workers for relevant or irrelevant to the operation issues, preoccupation with personal issues, desire to finish duty (Rail Safety and Standards Board, 2004b).

### Expectation

*Expectation* is defined as the strong anticipation that something will happen or be the case. It can be expressed by a well-known pattern of information which leads to a strong belief or mind-set about the appeared information (Rail Safety and Standards Board, 2004b).

Expectation is associated with dedicated to specific routes train drivers who operate on the same line several times per day. Hence, for instance, they tend to anticipate that signals will show certain aspects and change according to the usual pattern. Therefore, they might miss a red signal due to their strong belief that the signal would have changed to green aspect until the time they reach it.

### Perception

*Perception* is defined as the ability of operators to recognise by senses, i.e. sight, hearing, any external objects or phenomena. Perception should be considered as contextual factor when the aim is to see or hear what the information (visual or oral) or instructions (written or oral) say (Rail Safety and Standards Board, 2004b).

Operators perception failures include: misread visual information, e.g. red signal is seen as yellow, mishear oral information and misread written instructions (Rail Safety and Standards Board, 2004b).

**Interpretation**

*Interpretation* is described as the ability of operators to clarify or explain the meaning of information, instructions, etc. Interpretation should account for R-PSF when the aim, opposite to perception, is to understand the meaning of information (visual or oral) or instructions (written or oral) (Rail Safety and Standards Board, 2004b).

Operational failures due to interpretation encompass: misunderstanding of signal aspects, e.g. flashing signals, or misunderstanding of written procedures (Rail Safety and Standards Board, 2004b).

**Stress**

*Stress* is defined as the mental, emotional or physical tension and adverse reaction of operators due to excessive pressure (Rail Safety and Standards Board, 2004e).

Stress can be caused by excessive workload, conducting duties under adverse weather conditions, personal problems, e.g. financial losses, marriage problems, threat of assault, getting involved in accidents or being witness of suicide attempts and finally feeling insecure at the work place, e.g. organisational restructuring or relocation of depot. It should be noted that a certain level of stress is required to achieve good performance during the execution of a task (nominal level) (Gertman et al., 2004). The definition above, however, accounts only for high levels of stress.

**Fatigue**

*Fatigue* characterises extreme tiredness, which results from physical exhaustion due to exertion or prolonged activity.

Fatigue is a very common symptom with operators who perform their duties in shifts (Rail Safety and Standards Board, 2004e). It can be caused, amongst others, by lack of sleep, shift pattern, e.g. long duration of shifts, lack of resting days, inadequate duration of breaks, as well as understaffing and personal life style. Fatigue should not be counted as R-PSF when caused by mental exhaustion, as these cases are described by *Stress*. In addition, Fatigue due to illness should be considered as *Fit to work* R-PSF.

**Vigilance**

*Vigilance* is called the action or state of keenly being alert to possible dangers during the execution of duties.

Vigilance is strongly associated with the feeling of boredom and monotony as well as with under-loaded tasks (Rail Safety and Standards Board, 2004b). Therefore, more prone to low vigilant occurrences may be considered long distance train drivers who run the service with very few stops and meet mostly green signals.

**Situational awareness**

*Situational awareness* is described as the ability of operators to be aware of what is happening in their vicinity (around them) and subsequently what would be likely to happen in future (Rail Safety and Standards Board, 2012a).

Situational awareness involves, for instance, the necessity of signallers to acknowledge the exact location and routes of trains in order to avoid potential conflicts. For train drivers, on the other hand, situational awareness may involve the opening and closing of train doors, acknowledging where passengers are standing (Rail Safety and Standards Board, 2012a).

**Decision making skills**

*Decision making skills* account for operators' capability to make effective and timely decisions among several alternative scenarios and solve problems based on both their knowledge and experience (Rail Safety and Standards Board, 2012a).

A relevant example involves the decision that a train driver should take concerning the adopted driving strategy including appropriate speed and braking when operating under adverse weather conditions (Rail Safety and Standards Board, 2012a). For signallers, this may include the assessment of the impact on the network of the decision to block a line in order to perform the separation of a train (Gibson et al., 2004).

## Task Factors

*Task* R-PSFs simply characterise the features of an executed task (Kyriakidis et al., 2012). This category comprises factors such as monotony, workload, routine, time pressure, task complexity and task instructions.

**Workload**

*Workload* is generally defined as the amount of work to be done by operators in a specified period of time.

Workload could be either physical or mental, depending on the type of task. In addition it can be either low or extensive, while in either case the implications on operators' performance may be substantial, e.g. boredom or increased limits of stress (Rail Safety and Standards Board, 2004e). However, this definition captures only extensive workload, as boredom is covered by *Monotony*.

Extensive workload is likely due to sudden technical failures, e.g. signal or rolling stock, as well as understaffing. For instance, failure of a signal may increase the workload of signallers if it requires manual set up of the signals along the line, while they might also need to reset the routing of trains.

**Monotony**

*Monotony* is described as lack of variety and interest, tedious repetition of a task, which subsequently plausibly results in a general feeling of boredom.

In railway operations, examples of monotony can be expressed by high-speed routes with only few stops, where operators run at constant speed and experience mostly green signals (Rail Safety and Standards Board, 2004b).

**Routine**

*Routine* is the sequence of the usual followed actions or procedures during the execution of a task.

Operational routine is expressed by specific patterns followed throughout the execution of tasks. Therefore, any change of routines may lead to certain safety considerations. An example of an incident caused by such change is derived from personal communication between the author and a train operator. The operator claimed that in 20 years of services, he had only one incident of Signal Passed at Danger (SPAD). The driver said that he was used to prepare a cup of tea at a specific station and immediately after embarking the train, check the dedicated signal, then having a sip of tea and start the service. That specific day, the driver embarked the train without preparing his cup of tea and subsequently he forgot to check the signal before departing; this resulted to a SPAD as the signal was still showing a red aspect.

**Time pressure - Time to respond**

*Time pressure* is defined as the level of operators' mental pressure to complete a task within a specified time period. *Time to respond*, on the other hand, describes the available amount

of time that operators have to respond to and act on an unexpected event or to correct an identified error (Gertman et al., 2004).

Typical examples of performing a task under time pressure are: operators who are asked to perform their duties within a limited amount of time due to an unexpected delay or when they move from the one side of the train (cabin) to the other (rear cabin) to start a journey back.

**Task complexity**

*Task complexity* refers to how difficult a task is to be executed within a given context and environment (Gertman et al., 2004).

Task complexity is associated with complicated station layout, where the track layout requires increased level of attention from train drivers (Rail Safety and Standards Board, 2004b). It could also refer to required physical efforts, such as physical actions that are strenuous to be carried out due to complicated patterns or movements (Gertman et al., 2004).

**Task Instructions**

*Task instructions* refer to directions, orders or recommended rules that are implemented to guide and support operators concerning how they should carry out their tasks or run the services.

In operational terms, instructions should be clear, simple and explicit. For instance, signallers provide explicit instructions via radio communication to train drivers who perform shunting movements.

## Team Factors

*Team* R-PSFs are the factors that characterise and affect the performance of operators as members of a team that works together to achieve a common goal.

**Communication between employees**

*Communication between employees* implies the ability of operators to exchange information effectively via verbal and non-verbal media and stand on their position as necessary (Rail Safety and Standards Board, 2012a).

Effective communication is characterised by quick, clear, explicit, justified and verified exchange of information. An example of ineffective communication may involve a train driver who mishears the number of a signal that a signaller indicated, while neither of the operators confirmed that the correct number was understood. The above definition does not encompass the cases of communication failure due to technical problems.

**Teamwork**

*Teamwork* refers to how operators should work together in a positive, respectful and supportive manner and effectively coordinate their tasks and duties including the division of their responsibilities (Groth, 2009, Rail Safety and Standards Board, 2012a).

Teamwork is expressed by the clear distinction of roles and responsibilities during the execution of a task. The definition also includes the quality of feedback between team members, level of trust, structure and stability of a team, consideration of needs and support of team members and treating each other with respect (Rail Safety and Standards Board, 2012a).

**Team Relations**

*Team relations* refer to the social climate between all team members and how these members interact with each other (Groth, 2009).

The relations between team members are expressed by the levels of harmony and honesty, the level of frictions and how any conflicts are resolved, the level of mutual confidence and trust, and finally the frequency and level of common activities and social events.

**Quality and trust in information**

*Quality and trust in information* refers to the level of trust in persons providing the information and the level of completeness, accuracy and relevance of the shared information (Rail Safety and Standards Board, 2004b, Rail Safety and Standards Board, 2012a).

Examples may involve train drivers who do not trust a particular signaller to share the information or train drivers who convey only part of the relevant information or irrelevant information to their colleagues (Rail Safety and Standards Board, 2004b, Rail Safety and Standards Board, 2012a).

## Organisational Factors

*Organisational* R-PSFs include the factors defined and controlled by an organisation (Groth, 2009). With respect to a single occurrence, they are considered as static; however they tend to change over a longer period of time. Organisational factors embrace the characteristics and attitudes of an organisation as well as certain organisational behaviours that influence the performance of employees. Typical examples are the safety culture and safety management systems. Within these two broad categories, additional organisational features can be identified such as the quality of relevant procedures, leadership, supervision and the level of trust within the organisation. Furthermore, staffing, shift pattern and the provided training are also assigned to organisational factors and are directly associated with the number and type of the assigned tasks, the composition of teams as well as the levels of workload and fatigue. For instance, if the driver is found tired due to long consecutive shifts, then the issue is related to the organisation and the way the shift patterns are set.

**Safety culture (disregard procedures)**

*Safety culture* expresses what emerges as a result of a concerted organisational effort to move all cultural elements towards the goal of safety, including its members, systems, and work activities (Cooper, 2000, Sorensen, 2002). It is the set of values and priorities placed on all aspects of safety by all employees at any level within the organisation (Rail Safety and Standards Board, 2004e).

Indicators of organisations with poor safety culture include, amongst others, repeated cases of employees who disregard rules and procedures and management that favours results over safety, e.g. supports drivers to operate faster so as to increase capacity, risky driving behaviour, use of mobile phone or disobey procedures. On the other hand, organisations with positive safety culture are characterized by consciousness, evaluation and action on safety related matters and is supported by communications established on mutual trust throughout the whole organization (Rail Safety and Standards Board, 2004e).

**Safety Management Systems**

*Safety management systems (SMS)* generally refer to the set of work practices and procedures for monitoring and when necessary improving safety of operations. The EC (2004) define railways SMS as *"the organisation and arrangements established by an infrastructure manager or a railway undertaking to ensure the safe management of its operations"*.

Efficient and positive SMS clearly define the roles and responsibilities within an organisation, set arrangements for safety mechanisms, involve operators and employees of all levels in the process and finally ensure constant safety improvements (Office of Rail Regulation, 2011b).

**Training and training methods**

*Training/training methods* is defined as the continuous process provided by the organisation to its employees in order to assure that they have the appropriate knowledge, skills and attitudes to operate the railway according to the defined standards (Rail Safety and Standards Board, 2004e).

Training methods include lectures, demonstrations and simulation sessions. They are targeting to keep employees familiar to new technologies, refresh their qualifications and old skills as well as their levels of competence. Frequency and quality of training methods, in addition to the level of monitoring of employees who may need additional training are also incorporated within this definition.

**Quality and availability of procedures, standards and regulations**

Q*uality and availability of procedures, standards and regulations* refer to the existence and use of formal operating procedures for the tasks under consideration (Gertman et al., 2004)

High quality procedures are characterised by clarity, simplicity and availability with respect to how they describe the actions that operators are required to perform each of their tasks. For instance, how a train driver should proceed after passing signal at danger or what the appropriate signallers' actions are before, during and after applying an absolute possession on a line. Problems related to procedures, standards and regulations may arise when: (i) there are more than one similar procedure and it may not be clear which procedure should be applied in each of the cases; (ii) procedures, regulations and rules may have recently changed and operators are either unaware of the changes or they have not been adequately trained on the new procedures; (iii) there is lack of procedures; (iv) operators do not follow or obey the procedures (usually refer to new procedures) as they are not explicitly and clearly written.

**Leadership**

*Leadership* is described as the ability of an individual "*to motivate others to forego self-interest in the interest of a collective vision*" (House and Shamir, 1993). In general, it is

shown as the ability of a person to lead, rule, guide and influence the behaviour and actions of a group of people or organisation.

Good and effective leadership in railway organisations guarantees that "s*afety objectives of are established and prioritised, that best practices are applied to meet safety targets, that the effectiveness of the system is constantly checked and monitored, and that corrective and/or proactive countermeasures are taken when necessary"* (European Railway Agency, 2010). For instance, it can be expressed by the ability of the manager to motivate, guide and support operators to act towards achieving the goals of the organisation. Leadership should be considered as contributor R-PSF, when it refers to the higher levels of the hierarchy within the railway organisation.

**Supervision**

*Supervision* is defined as the act of "*managing and monitoring employees, developing them, supporting them and mediating on their behalf*" (Rail Safety and Standards Board, 2004e).

Effective supervision is characterised by supporting operators, helping in resolving conflicts, gathering information about service activities and any external conditions that may affect the safety of operations, checking on the progress of operators and quality of service, evaluating objectively the performance of operators and the effectiveness of the team.

Although supervision may overlap with leadership, for this definition it should be considered as relevant R-PSF, when it refers to the personnel responsible for connecting frontline personnel to the organisation's leadership (Groth, 2009).

**Shift Pattern**

*Shift pattern* defines and includes "*any organisation of working hours that differ from the traditional diurnal work period, while sometimes is synonymous of irregular or odd working hours*" (Costa, 2003).

Operational shift patterns include the: type of shift, e.g. morning or night, duration of shift, duration of breaks, and number of people per shift. Shift patterns affect significantly human performance as it is widely acknowledged that it leads to fatigue phenomena. In addition ineffective shift patterns can be characterized by health and family problems, stress, anxiety, irritability and depression (Rail Safety and Standards Board, 2004e)

**Relations within organisation**

*Relations within organisation* refer to the quality and type of personal relationships between people within the organisation. Relations within organisation indicate the social climate of all members of the organisation and the interactions between them on all different levels of hierarchy.

Relations within organisation may overlap with team relations; therefore this R-PSFs should be considered as contributor R-PSF only when it is identified that relations between employees of different levels in the hierarchy have led to operational safety irregularities.

**Incentives for employees**

*Incentives for employees* are all means that an organisation uses and offers to increase the motivation of its employees to operate and execute their tasks (Rail Safety and Standards Board, 2004b).

It is apparent from its definition that incentives of employees are strongly associated with employees' motivation. Incentives include amongst others: awards, e.g. financial rewards and extra days of holidays, promotions and public acknowledgment. The incentives that an organisation offers should be clearly stated and awarded without bias. For instance, it is equally important for employees to know that they will be rewarded for excellent performance but at the same that employees with a deliberate poor performance will not be acknowledged.

**Communication within organisation-feeling secure**

*Communication within organisation* refers to the quality and clarity of communication between the different levels in the hierarchy of employees, including the quality of feedback and the level of information about any organizational changes (May and Mumby, 2005).

For instance, it includes the clarity of information during a period of re-organisation, while employees may be afraid of losing their positions and the feeling of safety to express a personal opinion related to the way the organisation operates without fear of any consequences.

**Fit to Work Aspect**

*Fit to work aspect* outlines any assessments conducted by organisation to assure that its employees can perform safely a specific job or task. (Canadian Centre for Occupational

Health and Safety, 2006). In addition this definition incorporates organisations policy to encourage employees to not go to work when they feel unwell.

Fit to work aspect is expressed by the adequacy of an organisation to help its employees to: improve their personal fitness; reduce their fatigue and stress levels; overcome any health problems without the fear of any consequences, e.g. losing job.

## System Factors

*System* R-PSFs characterise the features of the railway system and how the interactions between the operators and system characteristics may affect the performance of operators. System factors include the design of railway system, operators' working environment, human machine interface, quality and trust in equipment and the railway communication means.

### System design

*System design* is described as the process to achieve high quality, clarity, practicality, and efficiency architecture, components, modules, interfaces and data in order to satisfy specified predefined requirements for the corresponding system (Churchman, 1971).

System design includes, amongst others, the complexity level of stations or tracks layout and the location of installed signals.

### Human Machine Interface

*Human Machine Interface* covers how information is communicated between operators and equipment (Groth, 2009). It describes the usability, physical access and level of friendliness to users of the machines.

HMI is characterized by the amount and clarity of available information on control panels, the functionality and layout of control panels and displays. It is also expressed by the train on-board equipment and how displays convey the information to drivers, as well as how easy it is for operators to input data and information on the machines when required.

**Working environment**

*Working environment* describes the nature of the physical working conditions such as lighting, noise, temperature (Hollnagel, 1998), as well as the layout and ergonomics of the working area.

Working environment encompasses, for instance, the level of noise, vibrations and temperature within a train cabin, as well as the design of the seats and the total cabin interior. Respectively for control rooms it includes temperature, ventilation and design of work place, e.g. position of displays.

**Trust in equipment**

*Trust in equipment* describes the level of trust that operators show to the output of information from a specific equipment (Rail Safety and Standards Board, 2004b).

The level of trust is affected by the long term, i.e. history of faults, reliability of the specific equipment. For instance, both signallers and train drivers may not trust the aspect of a signal that has very often failed in the past.

**Railway Communication Systems (RCS)**

*Railway communication systems* outline the facilities consisting of the physical plants and equipment for disseminating information, between railway employees and operators.

The railway communication systems include any radio and/or telecommunication means, e.g. GSM-R, used for the communication between operators. The RCS should be taken into account as contributor R-PSFs, when communication between operators fails due to technical failures or bad design, e.g. noise on the communication channels.

## Environmental Factors

*Environmental* factors refer to weather conditions at the time of the execution of a task.

**Weather Conditions**

*Weather conditions* describe the day-to-day meteorological conditions as expressed by the state of the atmosphere in terms of temperature, wind, cloudiness and precipitation.

Adverse weather conditions, for instance, may affect operations negatively, as it can be more difficult for operators to execute their tasks, e.g. operating under heavy train or fog, or braking on very slippery conditions.

**Visibility**

*Visibility* is defined as the distance and clarity that operators can see and identify an object on the line or relevant vicinity in order to perceive its emitted information.

For instance, how far and clear train drivers can see a signal or an obstacle on the track under given weather conditions or due to vegetation (sighting).

# Appendix V The Safety Maturity Model features

Table V-I The Safety Maturity Model questions

| | Questions |
|---|---|
| 1 | Do you publish an annual safety report (if yes, please say how we can download it, or if not downloadable, send us a copy) |
| 2 | How often do you monitor the safety data in your organisation? |
| 3 | How often do you prioritise the importance of safety related incidents? |
| 4 | How do you calibrate the importance of safety related incidents? |
| 5 | How often do budget constraints prevent your organisation from being safe? |
| 6 | Are processes to generate safety goals & policies effective? Are they explicit? |
| 7 | How explicit / prescriptive are safety procedures and plans? |
| 8 | Where do initiatives for improving safety come from? |
| 9 | How often does your organisation compare itself with other industries (e.g. nuclear plants, aviation, oil refineries) and their experience? |
| 10 | Do you keep a register of risks/hazards in order to mitigate them? If so, what category of risks / hazards do you monitor? |

Table V-II Scoring of questions

| | Scoring |
|---|---|
| 1 | Yes=4, No=0 |
| 2 | Yearly=1, Monthly=2, Weekly=3, Daily=3, All=4, Daily & Monthly=3, Quarterly & Yearly=1.5, Monthly & Quarterly=2 |
| 3 | After a serious accident=2, Every year=3, No fixed time=0, After a serious accident & Every year=4 |
| 4 | Actual impact=2, Potential worst-case scenario=3, Significance of trend=1, Actual impact & Potential worst-case scenario & Significance of trend=4, Actual impact & Potential worst-case scenario=3.5, Realistic scenario & Significance of trend=2.5 |
| 5 | All the time=0, Frequently=1, Rarely=2, Very rarely=3, Never=4 |
| 6 | Informal=1, Ineffective=0, Effective= 3, Explicit/detailed=3, Effective & Explicit/detailed=4 |
| 7 | Precise & detailed=3, Allow flexibility=2, Rely on qualified individuals=1, Precise & detailed - Allow flexibility=4, Precise & detailed - Rely on qualified individuals=3 |
| 8 | Safety specialists=4, Frontline operators=2, Frontline managers=3, Senior managers=1 |
| 9 | Never=0, Rarely=1, Annually=4, Always before investment decisions=2 |
| 10 | No register=0, Serious risks only=2, Serious & minor=4 |

Table V-III Safety maturity model features

| | Safety Maturity Model features | | | | |
|---|---|---|---|---|---|
| | **Initial (0-10 score)** | **Compliance – Driven (10-20)** | **Defined Risk (20-25)** | **Managed Risk (26-33)** | **Optimising (34-40)** |
| Safety definition | Safety is defined in terms of technical and procedural solutions and compliance with regulations, but is not seen as a key business risk. | Safety is defined not only in terms of technical and procedural solutions and compliance with regulations but also with safety training objectives. | Safety is seen as a key business risk, covering the safety of customers, staff, contractors and any other person on or adjoining the organisation's property. | The level of safety is measured quantitatively through monitoring precursors and top events and changes in the QRA model. | Management is constantly searching for ways to continue reducing the frequency of accidents. |
| Safety responsibility | The primary responsibility for safety is perceived to lie with the safety department. | Safety responsibility is transferred from the safety department to operations and maintenance. | CEO treats safety management as a critical business issue, with a main board member responsible for safety and all line managers responsible for safety in their functional area. | As in defined risk but line managers are also responsible for identifying and mitigating risks within a defined time period. This is then cascaded down to frontline personnel. | Management targets for year on year reductions in indicators of remaining risk while ensuring that staff responsible for reporting such indicators are not deterred by these targets from reporting all such incidents. |
| Design and equipment | The design of equipment and procedures do not consistently incorporate the principles of defence in depth or the prevention of single point failure. | New equipment and procedures are designed to incorporate the principles of defence in depth and the prevention of single point failures. | A programme is under way to bring all equipment and procedures up to recognised design standards and meet realistic safety targets incorporating the principles of risk minimisation. | The programme is complete, bringing all equipment and procedures up to recognised design standards and meeting realistic safety targets, incorporating the principles of risk minimisation. | Current design standards, procedures and other barriers to risk are constantly reviewed to keep pace with technological change to maintain ALARP. |

| | Safety Maturity Model features | | | | |
|---|---|---|---|---|---|
| | **Initial (0-10 score)** | **Compliance – Driven (10-20)** | **Defined Risk (20-25)** | **Managed Risk (26-33)** | **Optimising (34-40)** |
| Risk management | There is no systematic recognition, quantification or management of risk. | Legacy systems and system-wide risks are not addressed, quantified or managed. | All major safety hazards are identified, registered and the risk evaluated and mitigated. | Risks are aggregated and progress measured using a regularly updated model of quantified probabilistic risk analysis (QRA / PRA) or a similarly quantified system. | Measurement of risk and progress is constantly updated to adopt new methods and changes in best practice. |
| Variability and reliability | There is undocumented variation in operating and maintenance processes. | Operating and maintenance processes are now consistent with little or no variation. | Operating and maintenance processes are consistent with minimal variation. | Potential accident precursor incidents, near misses and any other available indices of the underlying risk level are recorded and monitored, while at the same time making efforts to ensure that all such incidents are recorded. | Progress in the science of human reliability is identified and the latest methods of improving reliability are incorporated to minimise human errors and variability. Automation is used wherever appropriate. |
| Human factors management | Managers perceive that the majority of accidents are solely caused by the unsafe behaviour of front-line personnel. | There is an emphasis on ensuring that qualified personnel are used in all safety critical roles. Personnel responsible for safety lapses are disciplined. | Care is taken to ensure that all job descriptions include safety responsibilities and that personnel are not only qualified but regularly trained to ensure up to date compliance. | Contractors are also required to apply similar safety standards and practices to those of the client, including providing data on safety indicators, incidents, and risks in their area of responsibility. | Appropriate and active communication takes place to ensure that passengers are fully aware of their own responsibilities for safety. |

| | Safety Maturity Model features | | | | |
|---|---|---|---|---|---|
| | **Initial (0-10 score)** | **Compliance – Driven (10-20)** | **Defined Risk (20-25)** | **Managed Risk (26-33)** | **Optimising (34-40)** |
| Attitude to hazards | Many minor hazardous incidents and the occasional accident are seen as unavoidable and as part of the job. | Some minor hazardous incidents and the occasional accident are seen as unavoidable and as part of the job. | Minor hazardous incidents and the occasional accident are seen as unavoidable in the long term, but their probability and frequency is seen as manageable and programmes are in place to lower it. | All existing safety hazards are identified and registered. | The condition and degradation of assets is regularly monitored to ensure that there is no deterioration in risk levels, either of safety or reliability. |
| Role and involvement of staff | Most frontline staff are uninterested in safety. | Organisation is convinced that the involvement of the frontline employee in safety is critical and frontline staff are aware of safety objectives but do not see themselves as responsible for mitigating safety risks outside their own functional responsibility. | Most frontline staff see themselves as responsible for mitigating safety risks within their own function and notifying management of any other hazards and risks whenever they identify any that are not already on the hazard register. | All employees are aware of the impact of their own function on safety and the quantified targets for that function and for the organisation as a whole. | All employees are trained and motivated to look for ways of improving safety and are incentivised to provide ideas to do this |
| Safety performance measurement | Safety performance is measured in terms of major accidents and fatalities. | Safety performance is measured in terms of lagging indicators prescribed by governments or statutory safety bodies. | Safety performance is measured in terms of indicators prescribed by governments or statutory safety bodies and identified by the company to be high priority potential causes or measures of risk of accident. | A full range of indicators comprehensively covers all of the highest priority potential causes or measures of risk of accident. Some at least of these are leading indicators that point to a change in the level of risk. | The range of indicators is regularly reviewed to ensure that new trends are identified relatively quickly and any negative trend is quickly mitigated. |

# Appendix VI Report on near-miss at Hanger Lane junction

## Near-miss at Hanger Lane junction 27 March 2009

### Contents


**Preface** 5
**Key Definitions** 5
**Summary of the Report** 6
- Key facts about the incident 6
- Immediate cause, causal and contributory factors, underlying causes 6
- Severity of consequences 7
- Recommendations 7

**The Incident** 8
- Summary of the incident 8
- The parties involved 8
- Location 8
- External circumstances 8
- Trains 8
- Signalling equipment 10
- Consequences of the incident 19

**The Investigation** 20
- Sources of evidence 20

**Key Information** 21
- Train Operator 21
- Earls Court Control Room 21
- Baker Street Control Room 22
- Methods of communication 22
- Signaller 23
- Rule Book 24
- Previous occurrences of a similar character 26

**Analysis** 27
- Identification of the immediate cause 27
- Identification of causal and contributory factors 27
- Identification of underlying factors 30
- Other factors for consideration 33

## Summary of the Report

### Key facts about the incident

6 At about 17:22 hrs on Friday 27 March 2009, eastbound District Line train 103 passed signal WM1 at danger at low speed. The train stopped with its leading end approximately five metres past the signal, and the train operator (driver of the train) reported the incident to the signaller by telephone. Signal WM1 is located between Ealing Broadway and Ealing Common in west London, and controls the approach to Hanger Lane junction where the Piccadilly and District lines converge.

7 At 17:26 hrs, the signaller gave the train operator authority to proceed across the junction. Shortly afterwards, and before train 103 had started to move, a westbound Piccadilly Line train crossed the junction in front of it creating the potential for a collision. The signaller had initially overlooked the presence of this train, and had then been unable to contact the train operator of train 103 once he became aware of the situation.

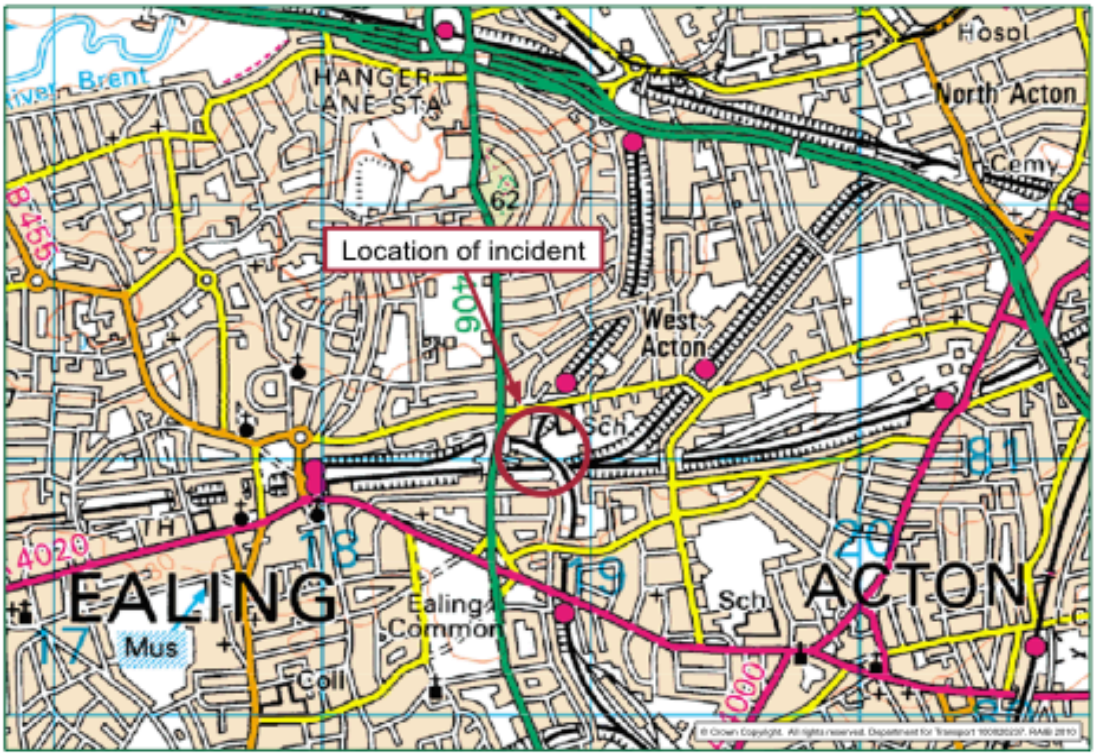



*Figure 1: Extract from Ordnance Survey map showing location of incident*

### Immediate cause, causal and contributory factors, underlying causes

8 The immediate cause of the incident was the signaller giving train 103 the authority to proceed towards Hanger Lane junction before it was safe to do so.

9 Causal factors were:

a. the operator of train 103 passed signal WM1 at danger; and

b. the signaller did not bring all trains to a halt to protect the junction before giving train 103 authority to proceed.

fit to work, par.110

10 The signaller was taking prescribed medication and it is possible that it affected his concentration and judgement. If this caused him to overlook the presence of train 316, the medication was also a causal factor.

11 Contributory factors were:

a. the signaller was under time pressure when dealing with the incident as it affected both the District and Piccadilly Line services and had the potential to cause extensive delays;

safety culture, par. 33

b. the signaller's workload, as he was operating two control desks during the incident; and

par. 103

c. the communication arrangements in place between Earls Court and Baker Street control rooms at the time of the incident, which did not give the signaller's call adequate priority and prolonged the period during which trains 103 and 316 were at risk of collision.

12 Underlying factors were:

system design, par. 116,117

a. The design of points and *track circuits* at Hanger Lane junction;

b. LUL's rules and procedures for incidents involving trains which pass a *semi-automatic signal* at danger, which were fragmented and did not provide guidance on how a signaller was to meet the requirement to ensure that it was safe for a train to proceed; and

c. LUL's arrangements for maintaining the signaller's skills in managing degraded working, which were insufficient.

training methods, supervision, par.81,123

## Severity of consequences

13 Although there were no immediate consequences of the incident, there was the potential for a collision to occur.

## Recommendations

14 Recommendations can be found in paragraph 142. They relate to the following areas:

- guidance provided to assist signallers in dealing with incidents;
- use of simulation techniques during training;
- training of safety critical communication skills;
- issuing of medical advice to managers; and
- preparation of post-incident reports.

# The Incident

## Summary of the incident

15 At approximately 17:22 hrs on 27 March 2009, eastbound District Line train number 103 passed signal WM1 at danger within two minutes of starting its journey from Ealing Broadway to Upminster. The train was travelling at low speed and was stopped by the *train protection system* approximately five metres beyond the signal.

16 The train operator reported the incident by telephone to the signaller controlling the Hanger Lane junction area from a control room at Earls Court. The train operator subsequently received verbal authority to proceed across the junction from the signaller.

17 Before the train operator had moved his train, he observed a westbound Piccadilly Line service (train 316) passing over Hanger Lane junction ahead of him on a conflicting path. Although the trains were separated by a distance of at least 115 metres, there was the potential for a collision to occur as one train was proceeding on clear signals and the other had authority from the signaller to proceed.

## The parties involved

18 LUL is the owner and maintainer of the infrastructure, and the employer of all the staff involved.

19 LUL freely co-operated with the investigation.

## Location

20 Hanger Lane junction is located between Ealing Broadway, North Ealing and Ealing Common stations where the District and Piccadilly lines converge. Having converged, both eastbound lines share the same track (figure 2). The junction is above ground.

21 All *powered points* and signals at Hanger Lane junction are controlled from the Piccadilly Line control room at Earls Court.

## External circumstances

22 The incident occurred during bright sunshine when the sun was low in the western sky.

## Trains

23 District Line train 103 was formed of D78 tube stock introduced from 1980, and Piccadilly Line train 316 was formed of 1973 tube stock introduced from 1975. Neither train was equipped with forward facing CCTV equipment or an on-train data recorder (OTDR).

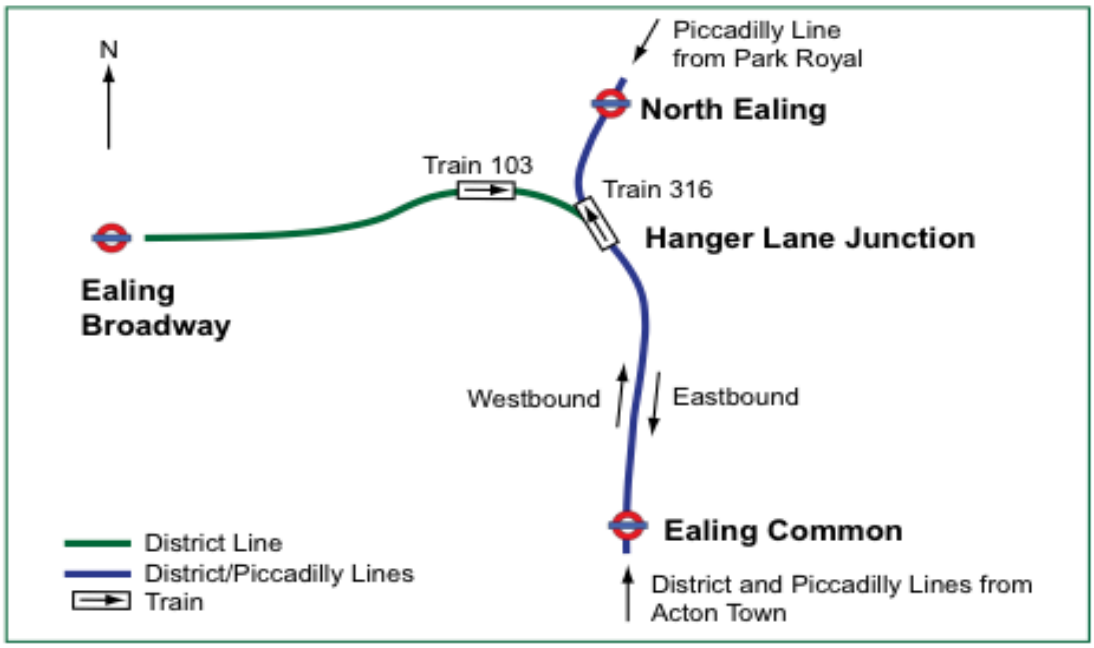



*Figure 2: Simplified map of Hanger Lane junction showing position of trains during incident*

24 Trains operating on the District and Piccadilly lines are fitted with *tripcock* equipment which works in conjunction with the signalling system. A *trainstop* is a device comprising a short arm which is raised when the signal is showing red (figure 3). If a train passes a signal at danger, the trainstop makes contact with the tripcock arm on the train which automatically applies the train's emergency brake. The tripcock equipment can be reset by the train operator following activation after the signaller has given authority to do so. After the tripcock has been reset, a timer limits the train to a maximum speed of 8 mph (13 km/h) for three minutes.

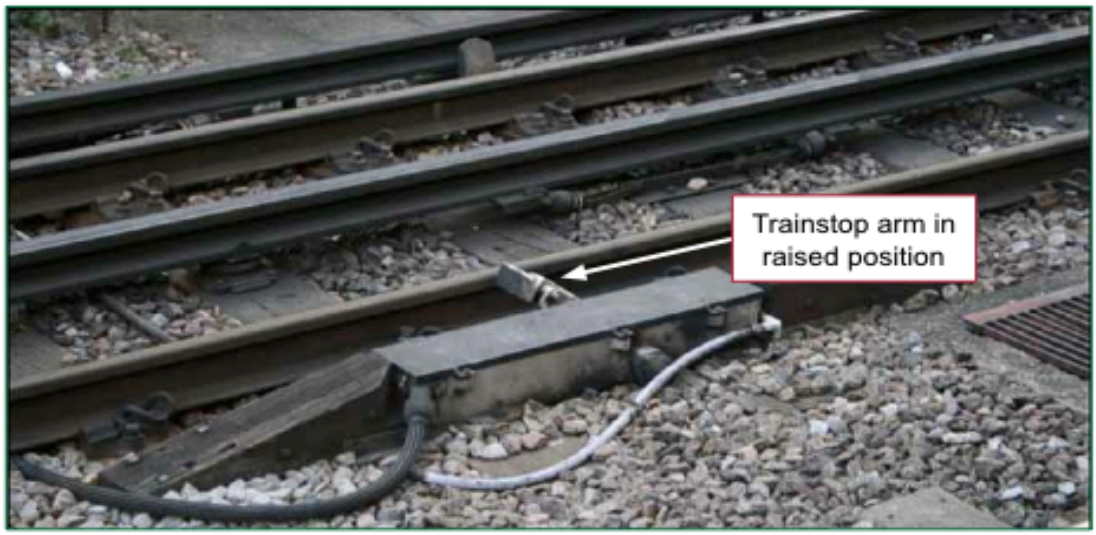



*Figure 3: Trainstop in the raised (stop) position (photograph courtesy of LUL library)*

## Signalling equipment

25 The converging eastbound routes approaching the junction are protected by signals WM1 on the approach from Ealing Broadway, and WM2 at North Ealing station. These signals are referred to as 'semi-automatic' signals, because they can operate in *automatic mode* or can be controlled by the signaller if necessary.

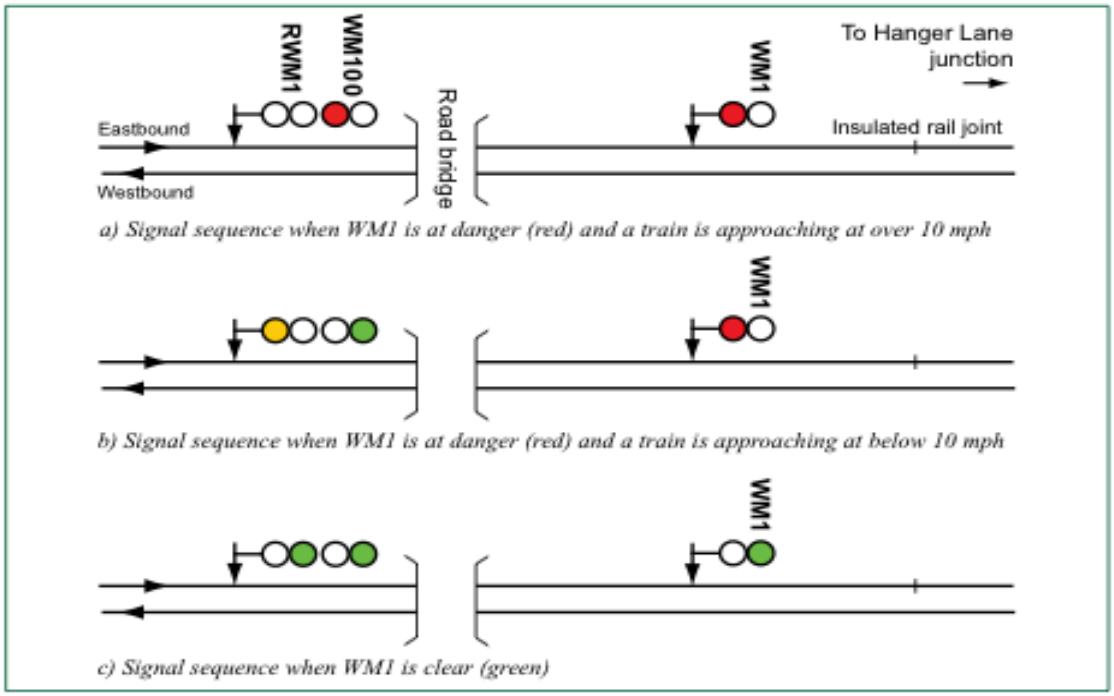



*Figure 4: Signal sequence for an eastbound District Line train approaching Hanger Lane junction*

26 Signal WM1 is preceded by signals WM100 and RWM1. Those two signals are located 74 metres before signal WM1 and mounted together (figures 4 and 5). They provide a train operator with information on the status of signal WM1:

a. Signal WM100, mounted at the top of the post, is provided to reduce the risk of a train passing signal WM1 at danger, and doing so at a speed which would make it unable to stop before reaching the junction. When signal WM1 is at danger, signal WM100 shows a red aspect to approaching trains travelling at more than 10 mph (16 km/h). A timing device, linked to track circuits which can detect a train's position, measures the speed of an approaching train. If signal WM1 remains at danger, signal WM100 will only clear to green once the signalling system has verified that the approaching train's speed is sufficiently low and that the trainstop at signal WM1 is raised.

b. Signal RWM1, mounted below signal WM100, is a *repeater signal* for signal WM1. This signal is designed to show a yellow aspect if signal WM100 is showing green but signal WM1 is at danger, or a green aspect if signal WM1 is green. It shows no aspect if signal WM100 is at danger. Signal RWM1 is provided to assist a train operator if the sighting of signal WM1 becomes obstructed when a train is approaching on the opposite line due to curvature of the track.




*Figure 5: Signal WM100 mounted above signal RWM1. Signal WM1 is beyond the road bridge ahead*

27 The *insulated rail joint* for signal WM1 is located 43.6 metres in advance of (ie beyond) it, as shown in figure 6. This separation is much greater than that found at more modern installations for historical reasons. If a train passed signal WM1 at danger, it would have to travel at least 43.6 metres in order to occupy the next track circuit before the infringement was detected by the signalling system and displayed to the signaller. It could then continue for a further 76 metres before reaching the junction (figure 7). If a train passed signal WM1 at danger but stopped within 43.6 metres, the overrun would not show on the track diagrams in the control room or affect other signals at the junction.

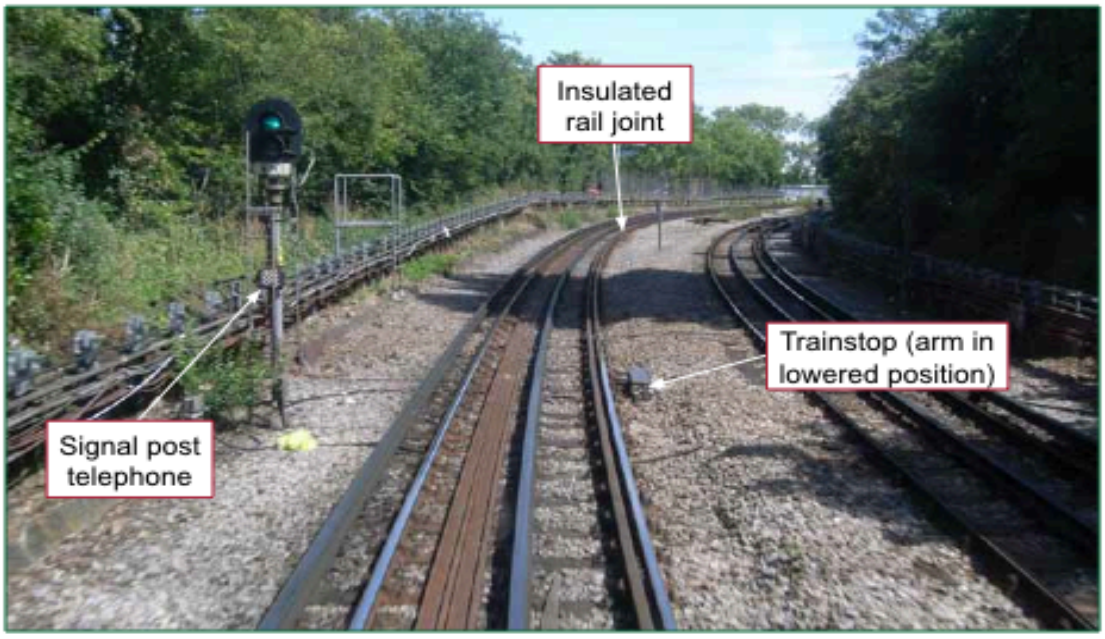



*Figure 6: Signal WM1, with Hanger Lane junction in the distance*

33 The signaller's first control desk duty was to cover the Acton Town desk (controlling the Hanger Lane junction area) from 17:00 hrs. At the same time he also took over the adjacent Northfields desk controlling the Heathrow branch, and worked the two positions simultaneously. This allowed two signallers to take meal breaks. He identified that some trains coming from Heathrow did not accord with train numbers shown on the control room display and instigated an exercise to correct this numbering in liaison with a member of station staff at Acton Town.

34 The District Line train operator, who was to drive train 103, booked on duty at 12:25 hrs at Upminster depot. He had been rostered to work a 'spare' turn to cover any train operator absences. He was subsequently asked to cover duty 542 that involved preparing a set on Upminster depot, prior to working a passenger service from Upminster to Ealing Broadway and return.

35 The train operator left Upminster at 15:32 hrs and had an uneventful journey to Ealing Broadway. After arriving, the train operator took a short break before rejoining his train.

36 At 17:21 hrs, train 103 departed from Ealing Broadway. As the train approached Hanger Lane junction, the train operator observed signal WM100 displaying a red aspect, which indicated that signal WM1 was also at danger. He slowed the train to below 10 mph (16 km/h). This was detected by the signalling system which allowed signal WM100 to clear to a green aspect and signal RWM1 to yellow (figure 4).

*Events during the incident*

37 After observing signal WM100 clear to green, the train operator moved his control lever to apply power and keep the train moving. He then observed that the trainstop at signal WM1 was raised and moved the control lever to the emergency brake position. This action was too late to prevent the train passing signal WM1 at danger. The tripcock tripped and the train came to a stand approximately five metres beyond the signal at 17:23 hrs.

38 The train operator left the driving cab and used a *signal post telephone* (SPT) mounted on the post of signal WM1 to contact the signaller. He reported that he had passed signal WM1 at danger without authority. He did not use the phonetic alphabet, describing the signal as "WM1" rather than "whiskey mike one". The signaller asked the train operator to wait while he contacted the District Line controller, and made arrangements to move the train.

39 At 17:25 hrs, the signaller rang the controller located at Baker Street. The controller was involved with managing a concurrent but unrelated incident at Sloane Square, which required him to deal with a high volume of radio traffic. As a consequence, he was not able to answer the call from the signaller immediately.

40 The controller answered the call after 65 seconds, and the signaller informed him of the incident and that he intended to authorise train 103 to proceed. The controller sought confirmation that the signaller had spoken to the train operator, and agreed that the signaller could authorise the train forwards across the junction. The controller accepted that the points could be treated as spring toggle points and therefore did not need securing.

c. telephone between the signaller and controller: the recipient was able to identify the caller by means of a screen display before answering and was required to give incoming calls from the other control room priority.

## Signaller

*Experience and training*

80 The signaller had over ten years' experience working within the Earls Court control room and also acted as an instructor for new signallers. He was a local union health and safety representative, and as a consequence, typically 25% of his working time was spent on union duties rather than operating a control desk.

81 The signaller sat and passed an annual reassessment in August 2008 as part of LUL's competency development programme. This assessment involved written tests to demonstrate knowledge under the headings 'Communicate using LU protocols' and 'Operate signalling equipment under degraded conditions'. During his time at Earls Court, he had assisted with, but not undertaken a practical safety critical communications course.

*Health, medication and hours worked*

82 The signaller had experienced deterioration in his health towards the end of 2008. His GP prescribed medication to treat his symptoms and advised him not to drive. The signaller was still taking this medication at the time of the incident.

83 Following a short period of absence, the signaller returned to work and requested a referral to LUL's occupational health department. After discussion, the signaller's manager agreed to his request. This was despite written guidance provided by the occupational health department not imposing a restriction on 'signalling duties' for the medication in question despite restricting those 'making decisions which can impact on safety'. This guidance, titled 'Helping managers decide on employee's fitness for work', was issued in January 2005.

84 The signaller attended an appointment with a doctor from the occupational health medical advisory team in January 2009. The doctor reported that while the signaller's condition was improving, he continued to experience symptoms that could affect his concentration. The doctor advised that while the signaller could continue to perform signaller duties at the desk, these were to be limited to five hours per day, and that he could undertake duties that did not require significant levels of concentration for the remainder of his shift. He also recommended that the signaller did not work night shifts pending a further review after five weeks.

## Analysis

### Identification of the immediate cause[4]

98 The immediate cause of the incident was the signaller giving the operator of train 103 authority to proceed before it was safe to do so.

### Identification of causal[5] and contributory[6] factors

Actions of the train operator on train 103

99 When the train operator of train 103 approached signal WM100, it was showing danger and as a consequence, he reduced speed to below 10 mph (16 km/h) which allowed this signal to clear to green. At the same time, signal RWM1, mounted on the same post, showed a yellow aspect indicating that signal WM1 remained at danger.

100 The train operator moved his controller to draw power, but this may have been necessary to keep the train moving as it was running on a curve and climbing a 1:76 uphill gradient. Therefore, this action does not, of itself, indicate that he had misread signal RWM1 or signal WM1.

101 The train operator alleged that he misread signal WM1 due to bright sunlight making it appear to show a green aspect. It is possible for signal WM1 to clear to green after a train has passed signal WM100. The signal head faces due west and the sun would have been low in the sky at this time of day in late March. However, the RAIB considers that it is unlikely that the train operator could have been misled by the aspect shown on signal WM1 for the following reasons:

a. The train operator was able to observe the preceding signals (RWM1 and WM100), exposed to similar lighting conditions, correctly.

b. An inspection by a London Underground supervisor travelling in the cab of a train at the same time of day and in bright sunlight within a week of the incident found that the signal was shaded and that the lit aspect was fully visible.

c. No previous allegations have been made regarding the sighting of this signal in bright sunlight.

d. The signal head is designed to minimise the risk of a false aspect being shown by having no reflective parts.

e. The red aspect was lit. If a train operator observes a signal which appears to be showing a dual aspect (ie both red and green in the case of signal WM1, as it would have been if the sunlight had made the green aspect appear to be lit), he is required by rule book 6, section 3, to treat this as a danger signal and stop his train.

[4] The condition, event or behaviour that directly resulted in the occurrence.

[5] Any condition, event or behaviour that was necessary for the occurrence. Avoiding or eliminating any one of these factors would have prevented it happening.

[6] Any condition, event or behaviour that affected or sustained the occurrence, or exacerbated the outcome. Eliminating one or more of these factors would not have prevented the occurrence but their presence made it more likely, or changed the outcome.

102 It is therefore probable that the train operator either mis-read the signal sequence or was distracted. He observed the raised trainstop, but did not identify that signal WM1 remained at danger until he was too close to it to stop, and he subsequently passed the signal at danger. Not identifying that signal WM1 was at danger was a causal factor in the subsequent near-miss incident.

## Actions of the signaller

### *Accidental disconnection of the signaller's telephone cable*

103 The signaller, after obtaining the controller's permission to move train 103 forwards, then proceeded to instruct the train operator about the movement of train 103 (paragraph 46). It is possible that the signaller intended to hold the operator of train 103 on the telephone and that the train operator either concluded the call prematurely or was cut-off. This conversation was partially recorded in the background of another call, but it is not possible to establish what the signaller instructed the train operator to do. The signaller has alleged that the telephone cable came out of its socket on the control desk, and that the train operator inferred that he had authority to proceed earlier than the signaller had intended as a consequence.

104 The telephone cable is straight and it is possible to pull out the cable. However, this scenario is not supported by the following evidence:

a. the telephone cable is sufficiently long to make it difficult for an operator to pull out the cable while remaining seated at the control desk;

b. no witnesses in the control room observed this event;

c. the signaller did not report the occurrence to the controller, manager or any other member of staff at the time;

d. the signaller, in his subsequent telephone call to the controller, twice confirmed that he had given the operator of train 103 authority to move across the junction (paragraph 50); and

e. the train operator, despite omitting to repeat the message back to the signaller, believed that the call had concluded and that he had been given authority to proceed.

The RAIB does not consider that the alleged disconnection of the telephone cable was a factor in this incident.

### *Signaller's working conditions*

105 The signaller was working under time pressure during the incident as blocking the junction area for an extended period would affect services on the District and Piccadilly lines in both directions. This made the option of authorising train 103 forwards attractive (paragraph 95). The need to resolve the situation quickly in order to minimise disruption was a contributory factor.

106 The signaller's training dictated that he should protect the junction by stopping all train movements before permitting the incident train to move. LUL provide desk-specific training documents detailing the configuration of the signal areas controlled by each desk, but this information did not provide guidance on how to protect specific sections of track.

107 Once the signaller had obtained authority from the controller to move train 103 forwards across the junction rather than setting it back behind the signal, he took action to stop trains 340 and 326 from approaching the junction. Train 316 was closest to the junction, and he had the option of either stopping it at signal WM21 before the junction, or letting it run (ensuring that it had cleared the junction before giving authority for train 103 to proceed). He was aware of this train and observed signal WM22 clearing to green on the signalling diagram, allowing it to leave Ealing Common station, while he was talking to the train operator of train 103 (paragraph 46).

108 The signaller, who has responsibility for all train movements, took no action to prevent train 316 from continuing its journey. Despite there being no destination shown on the display panel, the train's number, prefixed with a 3, indicated that it was a Piccadilly Line service. It would therefore have to cross the junction to reach its next stop at North Ealing station. Had the signaller implemented the correct process, by bringing all trains to a halt to protect the junction before giving train 103 authority to proceed, this incident would have been avoided. This forms part of a signaller's basic training, and not doing so was a causal factor.

109 The signaller appeared to have forgotten about train 316 until his off-duty colleague intervened, 20 seconds after the end of his conversation with the operator of train 103. This may have been due to lack of attention or distraction arising from looking after two desks at once (paragraph 33). The signaller's workload during the incident was a contributory factor.

110 Other possible reasons for this oversight include the signaller's underlying health issues and the prescribed medication he was taking at the time of the incident. In order to determine the likely effects of the signaller's medication on his performance, LUL's occupational health department sought the opinion of a medical specialist with expertise in this area. In parallel, the RAIB commissioned independent advice from an occupational health consultant. Both opinions conclude that the medication being taken by the signaller is 'unlikely' to have caused the overlooking of the position of train 316 when giving train 103 authority to proceed. However, this type of medication can lead to impaired performance[7], with side effects including confusion and impaired concentration, either of which would have impacted on the signaller's performance if they occurred. For this reason, the impact of the medication cannot be ruled out; it is possible that the taking of prescribed medication may have affected the signaller's concentration and judgement. If this did cause him to overlook the presence of train 316, the medication was a causal factor in the incident.

111 The RAIB has analysed the signaller's actual shift pattern against the fatigue and risk index published by the HSE. There is no evidence that the signaller's actual working pattern (paragraph 86) created a risk of fatigue, although this assessment does not evaluate the effects of underlying health issues.

### Signaller's communication arrangements

112 At the time of the incident, signallers were reliant on controllers to contact a train operator via radio if they needed to communicate with them. The only alternative was if a train operator initiated a call to the signaller using a signal post telephone (paragraph 79). These arrangements had been in place since the introduction of train radios in the 1980s, and were originally facilitated by the co-location of signallers and controllers as a signaller could attract a controller's attention immediately in an emergency. The controllers were also able to view the signalling diagram.

113 The removal of District Line controllers to Baker Street took place in February 2008, 15 months before Connect Radio equipment was brought into use in Earls Court control room. During the interim period, if a signaller needed to contact a District Line train operator, he first had to telephone the District Line controller with the attendant risk of delay if the controller did not respond promptly. This arrangement did not contribute to the signaller's decision to authorise train 103 to proceed, but it meant that once he had realised his error, it took almost a minute for the message to reach the train operator via the controller who was distracted by a separate incident on a different part of the network.

114 The communication arrangements in place between Earls Court and Baker Street control rooms at the time of the incident did not give the signaller's call adequate priority. This prolonged the period during which the trains 103 and 316 were at risk of collision and was therefore a contributory factor.

## Identification of underlying factors[8]

### *The arrangements for moving train 103 after the SPAD incident*

115 Hanger Lane junction was equipped with a signalling system which was designed and installed during the 1950s. This included loose points on the eastbound line, used by trains in the trailing direction.

116 The loose points had a direct influence on the way in which the incident was handled, and in particular the options available to the controller and signaller. Their presence meant that the normal choice between working the train in the wrong direction to move it back behind the signal or securing the points ahead of the train did not apply. Although both options required a station supervisor or duty manager to attend site and would have been time-consuming, either would have secured the safety of the movement.

117 The loose points allowed the signaller to move the train forwards with the only protection provided being the manual stopping of all trains approaching the junction.

| |
|---|
| *Rule book 5, section 8: 'Passing a semi-automatic signal at danger (manually driven lines')*<br>• This section states that 'when it is safe to do so, you must give authority to pass the signal at danger' (paragraph 91), but is vague on how this is to be achieved. It does not otherwise deal with trains which have passed a signal at danger without authority. |
| *Rule book 5, section 10: 'Securing and unsecuring points'*<br>• This section refers to spring toggle points and states that they do not need to be secured for movements in the normal direction of travel. However, there is no reference to loose points despite the fact that they are treated in the same way (paragraph 92). |
| *Rule book 7, section 4: 'Signals passed at danger (without authority)'*<br>• This section contains a cross-reference to rule book 5 for instructions for passing an automatic signal at danger. The related cross-reference for semi-automatic signals is omitted.<br>• This section does not detail the actions to be taken by the signaller prior to giving a train operator authority to proceed. A cross-reference to rule book 5, section 8 where this information is given is omitted.<br>• A cross-reference to rule book 5, section 2 'working in the wrong direction', required if an incident train is to be returned to the correct side of a signal is omitted.<br>• A cross-reference to rule book 5, section 10 'Securing and unsecuring points', required if an incident train is close to a junction, is omitted. |

*Table 2: Summary of rule book issues relating to semi-automatic signals passed at danger*

## Training and assessment of signallers

123 Signallers are subject to annual reassessment to ensure that they are capable of performing their duties. This is done using LUL's competence management system, 'Competence Standards & Guidance for service control', issued in April 2008. Section 3.2 of this document addresses the operation of signalling equipment under degraded conditions and describes the method of assessment as involving 'observation and simulation/desk top carried out away from the workplace mainly during Continuous Development Programme'. However, the course attended by the signaller in August 2008 did not involve simulation or desktop exercises for either degraded working or safety critical communications (paragraph 81).

124 A simulator is used for the training of new staff entering Earls Court control room, but this has not been routinely used for developing or enhancing the skills of existing signallers. The lack of such training denies signallers the opportunity to develop their practical skills except through live incidents. In this instance, an error was made which might have been prevented with enhanced training and the lack of such training is therefore considered to be an underlying factor.

# Conclusions

## Immediate cause

130 The immediate cause of the incident was the signaller giving train 103 the authority to proceed towards Hanger Lane junction before it was safe to do so.

## Causal factors

131 Causal factors were:

a. the operator of train 103 passed signal WM1 at danger (paragraph 102); and

b. the signaller did not bring all trains to a halt to protect the junction before giving train 103 authority to proceed (paragraph 108).

132 The signaller was taking prescribed medication and it is possible that it affected his concentration and judgement. If this caused him to overlook the presence of train 316, the medication was also a causal factor (paragraph 110).

## Contributory factors

133 Contributory factors were:

a. the signaller was under time pressure when dealing with the incident as it affected both the District and Piccadilly Line services and had the potential to cause extensive delays (paragraph 105);

b. the signaller's workload, as he was operating two control desks during the incident (paragraph 109); and

c. the communication arrangements in place between Earls Court and Baker Street control rooms at the time of the incident, which did not give the signaller's call adequate priority and prolonged the period during which trains 103 and 316 were at risk of collision (paragraph 114. Refer also to paragraph 141).

## Underlying factors

134 Underlying factors were:

a. The design of points and track circuits at Hanger Lane junction (paragraph 120);

b. LUL's rules and procedures for incidents involving trains which pass a semi-automatic signal at danger, which were fragmented and did not provide guidance on how a signaller was to meet the requirement to ensure that it was safe for a train to proceed (paragraph 122 and table 1, **recommendation 1**); and

c. LUL's arrangements for maintaining the signaller's skills in managing degraded working, which were insufficient (paragraph 124, **recommendation 2**).

## Additional observations[10]

135 LUL's arrangements for developing and maintaining the safety critical communication skills of their staff were insufficient despite action taken in response to an earlier RAIB recommendation (paragraphs 125 and 129, **recommendation 3**).

136 Written guidance on the effects of medication, issued by LUL's occupational health department to assist managers in determining an employee's fitness to work, did not include signalling duties in the list of occupations affected by the medication prescribed to the signaller. This is a clear omission given that staff, whose duties include 'making decisions which can impact on safety', are listed as an affected group (paragraph 83, **recommendation 4**).

137 LUL's medical advisory guidance, in the case of the signaller, did not consider whether it was desirable for a person working restricted hours to be regularly working on rostered rest days (paragraph 85, **recommendation 5**).

138 The existing arrangements for preparing and issuing a signaller's post-incident report did not operate as intended. This was due to the Piccadilly Line Service Manager being unaware of the potential near-miss, as it had not been reported to him, and his confusion over the numbers and directions of the trains involved (paragraph 128 and table 2, **recommendation 6**).

# Appendix VII The Swiss questionnaire (English version)

Zurich, 25/05/2011

Dear Sir / Madam,

My name is Miltos Kyriakidis and I am research student at Lloyds Register Educational Trust - Transport Risk Management Centre of Imperial College London, undertaking a PhD in railway safety.

My research topic is related to Human Reliability Analysis. The aim of my research is to develop a tool, which will be used in order to estimate human error probability of train drivers and signalers, based on a new taxonomy of performance shaping factors (PSFs) focused on railway operations.

To do this, I am currently working with the Forschungsgruppe Organisation – Arbeit Technologie of ETH Zurich and SBB. I am using SBB accident and incident reports to extract the information are needed to develop such taxonomy. To validate the findings of the reports and to finalize the taxonomy I would like to ask you to fill in the following questionnaire.

The questionnaire is consisted of two parts. In the first part you are asked to assess the factors that influence human performance in four different rail scenarios. In the second part, you are asked to indicate the possible interactions amongst the identified PSFs.

Details and instructions are provided in each part of the questionnaire.

Thank you in advance for your time and effort.

Sincerely,

Miltos Kyriakidis

Research Student in Transport Risk Management
Centre for Transport Studies
Department of Civil and Environmental Engineering
Imperial College London
UK
e-mail: m.kyriakidis@imperial.ac.uk

## Performance Shaping Factors Taxonomy (PSFs) Questionnaire

Part 1

In this part you are asked to assess the factors that compromise human performance in four different scenarios. The scenarios have been selected from 85 accident and incident reports based on their frequency and the severity of their consequences. Please assess the factors from the most important to least important one. Indicate as 1 the factor that plays the most important role on operator's performance, while as 18 the least important factor. Finally, in the case that you consider as "important" any other factor(s) which is not included in the list, please add it in the "blank" area.

**Scenario 1:**

Long distance passenger train derails on open line due to excessive speed. **Train driver** did not reduce the speed as the alarm indicated. Competent train driver. No technical failures. Night operation. No adverse weather conditions.

(PSFs evaluation: 1 most important, 18 least important)

| **PSFs** | | **PSFs** | | **PSFs** | |
|---|---|---|---|---|---|
| Communication between employees - information | | Fatigue (sleep lack, shift pattern) | | Quality of procedures | |
| Teamwork - Supervision | | Fit to work - Health | | Stress | |
| Distraction - Concentration | | System design - Human Machine Interface | | Time pressure (time to respond) | |
| Expectation - Routine | | Perception | | Training | |
| Experience - Familiarity | | Safety culture (disregard procedures) | | Weather conditions - Visibility | |
| Situational awareness | | Visibility | | Workload | |
| | | | | | |

**Scenario 2:**

Regional passenger train passes a signal at danger on open line (Hauptsignal überfahrungen). **Train driver** is aware of the signal, does not obey though. Competent driver, familiar with the route. No technical failures. Day operation, good weather conditions.

(PSFs evaluation: 1 most important, 18 least important)

| PSFs | | PSFs | | PSFs | |
|---|---|---|---|---|---|
| Communication between employees - information | | Fatigue (sleep lack, shift pattern) | | Quality of procedures | |
| Teamwork - Supervision | | Fit to work - Health | | Stress | |
| Distraction - Concentration | | System design - Human Machine Interface | | Time pressure (time to respond) | |
| Expectation - Routine | | Perception | | Training | |
| Experience - Familiarity | | Safety culture (disregard procedures) | | Weather conditions - Visibility | |
| Situational awareness | | Visibility | | Workload | |
| | | | | | |

**Scenario 3:**

Freight train passes signal at danger in station due to wrong interpretation (Hauptsignal überfahrungen). **Train driver** is aware of the signal, fails to interpret and keeps the speed constant. Competent driver, familiar with the route. No technical failures. Night operation. No adverse weather conditions.

(PSFs evaluation: 1 most important, 18 least important)

| PSFs | | PSFs | | PSFs | |
|---|---|---|---|---|---|
| Communication between employees - information | | Fatigue (sleep lack, shift pattern) | | Quality of procedures | |
| Teamwork - Supervision | | Fit to work - Health | | Stress | |
| Distraction - Concentration | | System design - Human Machine Interface | | Time pressure (time to respond) | |
| Expectation - Routine | | Perception | | Training | |
| Experience - Familiarity | | Safety culture (disregard procedures) | | Weather conditions - Visibility | |
| Situational awareness | | Visibility | | Workload | |
| | | | | | |

**Scenario 4:**

During a shunting operation in depot a shunting train collides with a commercial train. **Signaller**, responsible for the route, gives authorisation to shunting train to advance onto the track that the commercial train was moving along. At the time of the occurrence signaller was alone in the control room. Procedures available. Competent signaller. No technical failures. Day shift.

(PSFs evaluation: 1 most important, 18 least important)

| **PSFs** | | **PSFs** | | **PSFs** | |
|---|---|---|---|---|---|
| Communication between employees - information | | Fatigue (sleep lack, shift pattern) | | Quality of procedures | |
| Teamwork - Supervision | | Fit to work - Health | | Stress | |
| Distraction - Concentration | | System design - Human Machine Interface | | Time pressure (time to respond) | |
| Expectation - Routine | | Perception | | Training | |
| Experience - Familiarity | | Safety culture (disregard procedures) | | Weather conditions - Visibility | |
| Situational awareness | | Visibility | | Workload | |
| | | | | | |

Part 2

In this section you are asked to indicate the possible interactions amongst the identified factors that influence human performance. The interactions represent the irreversible influence between two factors or how the factors in the left hand column affect the factors in the first row. The reason for irreversible influence lies in the characteristics of the railway environment where one factor may influence another one without any reverse effect. For instance, task complexity can influence driver or signaller's levels of stress, while the opposite influence (impact of stress on task complexity) is simply not logical. To indicate that a factor has an influence on another factor please use the symbol (✓).

| **Most frequent R-PSFs**<br>**Direct influence – Influence of Y upon X**<br>**(Y→X)** | Communication between employees - Information | Teamwork - Supervision | Distraction - concentration | Expectation - Routine | Experience - Familiarity | Situational awareness | Fatigue | Fit to work - Health | System design - HMI | Perception | Safety culture (disregard procedures) | Quality of procedures | Visibility | Stress | Time pressure - Time to respond | Training | Weather conditions | Workload |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Communication between employees - Information | | | | | | | | | | | | | | | | | | |
| Teamwork - Supervision | | | | | | | | | | | | | | | | | | |
| Distraction - concentration | | | | | | | | | | | | | | | | | | |
| Expectation - Routine | | | | | | | | | | | | | | | | | | |
| Experience - Familiarity | | | | | | | | | | | | | | | | | | |
| Situational awareness | | | | | | | | | | | | | | | | | | |
| Fatigue | | | | | | | | | | | | | | | | | | |
| Fit to work - Health | | | | | | | | | | | | | | | | | | |
| System design -Human Machine Interface | | | | | | | | | | | | | | | | | | |
| Perception | | | | | | | | | | | | | | | | | | |
| Safety culture (disregard procedures) | | | | | | | | | | | | | | | | | | |
| Quality of procedures | | | | | | | | | | | | | | | | | | |
| Visibility | | | | | | | | | | | | | | | | | | |
| Stress | | | | | | | | | | | | | | | | | | |
| Time pressure - Time to respond | | | | | | | | | | | | | | | | | | |
| Training | | | | | | | | | | | | | | | | | | |
| Weather conditions | | | | | | | | | | | | | | | | | | |
| Workload | | | | | | | | | | | | | | | | | | |

**Please indicate below your**

Job description:

Number of years in this job:

Previous position:

Number of years in SBB:

**Thank you for your cooperation**

# Appendix VIII Crosstabulations of reference database

Table VIII-1

**R-PSFs (lite) * Type of event Crosstabulation**

| | | | Type of event | | | |
|---|---|---|---|---|---|---|
| | | | Incident | Accident | Serious Accident | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 89 | 115 | 80 | 284 |
| | | Expected Count | 124.4 | 104.7 | 54.9 | 284.0 |
| | | % within R-PSFs (lite) | 31.3% | 40.5% | 28.2% | 100.0% |
| | | Std. Residual | -3.2 | 1.0 | 3.4 | |
| | System design - Visibility - HMI - RCM | Count | 71 | 71 | 45 | 187 |
| | | Expected Count | 81.9 | 69.0 | 36.2 | 187.0 |
| | | % within R-PSFs (lite) | 38.0% | 38.0% | 24.1% | 100.0% |
| | | Std. Residual | -1.2 | .2 | 1.5 | |
| | Fatigue - Shift pattern - Fit to work | Count | 37 | 34 | 38 | 109 |
| | | Expected Count | 47.7 | 40.2 | 21.1 | 109.0 |
| | | % within R-PSFs (lite) | 33.9% | 31.2% | 34.9% | 100.0% |
| | | Std. Residual | -1.6 | -1.0 | 3.7 | |
| | Communication - Teamwork | Count | 103 | 90 | 35 | 228 |
| | | Expected Count | 99.9 | 84.1 | 44.1 | 228.0 |
| | | % within R-PSFs (lite) | 45.2% | 39.5% | 15.4% | 100.0% |
| | | Std. Residual | .3 | .6 | -1.4 | |
| | Quality of procedures | Count | 58 | 43 | 28 | 129 |
| | | Expected Count | 56.5 | 47.6 | 24.9 | 129.0 |
| | | % within R-PSFs (lite) | 45.0% | 33.3% | 21.7% | 100.0% |
| | | Std. Residual | .2 | -.7 | .6 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 117 | 93 | 33 | 243 |
| | | Expected Count | 106.4 | 89.6 | 47.0 | 243.0 |
| | | % within R-PSFs (lite) | 48.1% | 38.3% | 13.6% | 100.0% |
| | | Std. Residual | 1.0 | .4 | -2.0 | |
| | Perception - Interpretation | Count | 48 | 56 | 15 | 119 |
| | | Expected Count | 52.1 | 43.9 | 23.0 | 119.0 |
| | | % within R-PSFs (lite) | 40.3% | 47.1% | 12.6% | 100.0% |
| | | Std. Residual | -.6 | 1.8 | -1.7 | |
| | Training - Experience | Count | 77 | 42 | 13 | 132 |
| | | Expected Count | 57.8 | 48.7 | 25.5 | 132.0 |
| | | % within R-PSFs (lite) | 58.3% | 31.8% | 9.8% | 100.0% |
| | | Std. Residual | 2.5 | -1.0 | -2.5 | |
| | Expectation - Familiarity - Routine | Count | 60 | 28 | 12 | 100 |
| | | Expected Count | 43.8 | 36.9 | 19.3 | 100.0 |
| | | % within R-PSFs (lite) | 60.0% | 28.0% | 12.0% | 100.0% |
| | | Std. Residual | 2.4 | -1.5 | -1.7 | |
| | Quality of information | Count | 27 | 17 | 12 | 56 |
| | | Expected Count | 24.5 | 20.6 | 10.8 | 56.0 |
| | | % within R-PSFs (lite) | 48.2% | 30.4% | 21.4% | 100.0% |
| | | Std. Residual | .5 | -.8 | .4 | |
| | Supervision | Count | 17 | 7 | 7 | 31 |
| | | Expected Count | 13.6 | 11.4 | 6.0 | 31.0 |
| | | % within R-PSFs (lite) | 54.8% | 22.6% | 22.6% | 100.0% |
| | | Std. Residual | .9 | -1.3 | .4 | |
| | Workload - Time pressure - Stress | Count | 30 | 22 | 6 | 58 |
| | | Expected Count | 25.4 | 21.4 | 11.2 | 58.0 |
| | | % within R-PSFs (lite) | 51.7% | 37.9% | 10.3% | 100.0% |
| | | Std. Residual | .9 | .1 | -1.6 | |
| Total | | Count | 734 | 618 | 324 | 1676 |
| | | Expected Count | 734.0 | 618.0 | 324.0 | 1676.0 |
| | | % within R-PSFs (lite) | 43.8% | 36.9% | 19.3% | 100.0% |

Table VIII-2

R-PSFs (lite) * Fatal or Not Crosstabulation

| | | | Fatal or Not | | |
|---|---|---|---|---|---|
| | | | No fatal event | Fatal accident | Total |
| R-PSFs (lite) | Safety culture – SMS | Count | 236 | 48 | 284 |
| | | Expected Count | 247.2 | 36.8 | 284.0 |
| | | % within R-PSFs (lite) | 83.1% | 16.9% | 100.0% |
| | | % of Total | 14.1% | 2.9% | 16.9% |
| | | Std. Residual | -.7 | 1.9 | |
| | System design – Visibility – HMI – RCM | Count | 157 | 30 | 187 |
| | | Expected Count | 162.8 | 24.2 | 187.0 |
| | | % within R-PSFs (lite) | 84.0% | 16.0% | 100.0% |
| | | % of Total | 9.4% | 1.8% | 11.2% |
| | | Std. Residual | -.5 | 1.2 | |
| | Fatigue – Shift pattern – Fit to work | Count | 85 | 24 | 109 |
| | | Expected Count | 94.9 | 14.1 | 109.0 |
| | | % within R-PSFs (lite) | 78.0% | 22.0% | 100.0% |
| | | % of Total | 5.1% | 1.4% | 6.5% |
| | | Std. Residual | -1.0 | 2.6 | |
| | Communication – Teamwork | Count | 204 | 24 | 228 |
| | | Expected Count | 198.5 | 29.5 | 228.0 |
| | | % within R-PSFs (lite) | 89.5% | 10.5% | 100.0% |
| | | % of Total | 12.2% | 1.4% | 13.6% |
| | | Std. Residual | .4 | -1.0 | |
| | Quality of procedures | Count | 110 | 19 | 129 |
| | | Expected Count | 112.3 | 16.7 | 129.0 |
| | | % within R-PSFs (lite) | 85.3% | 14.7% | 100.0% |
| | | % of Total | 6.6% | 1.1% | 7.7% |
| | | Std. Residual | -.2 | .6 | |
| | Distraction – Loss of concentration – Vigilance – Risk awareness | Count | 216 | 27 | 243 |
| | | Expected Count | 211.5 | 31.5 | 243.0 |
| | | % within R-PSFs (lite) | 88.9% | 11.1% | 100.0% |
| | | % of Total | 12.9% | 1.6% | 14.5% |
| | | Std. Residual | .3 | -.8 | |
| | Perception – Interpretation | Count | 112 | 7 | 119 |
| | | Expected Count | 103.6 | 15.4 | 119.0 |
| | | % within R-PSFs (lite) | 94.1% | 5.9% | 100.0% |
| | | % of Total | 6.7% | .4% | 7.1% |
| | | Std. Residual | .8 | -2.1 | |
| | Training – Experience | Count | 115 | 17 | 132 |
| | | Expected Count | 114.9 | 17.1 | 132.0 |
| | | % within R-PSFs (lite) | 87.1% | 12.9% | 100.0% |
| | | % of Total | 6.9% | 1.0% | 7.9% |
| | | Std. Residual | .0 | .0 | |
| | Expectation – Familiarity – Routine | Count | 91 | 9 | 100 |
| | | Expected Count | 87.1 | 12.9 | 100.0 |
| | | % within R-PSFs (lite) | 91.0% | 9.0% | 100.0% |
| | | % of Total | 5.4% | .5% | 6.0% |
| | | Std. Residual | .4 | -1.1 | |
| | Quality of information | Count | 50 | 6 | 56 |
| | | Expected Count | 48.7 | 7.3 | 56.0 |
| | | % within R-PSFs (lite) | 89.3% | 10.7% | 100.0% |
| | | % of Total | 3.0% | .4% | 3.3% |
| | | Std. Residual | .2 | -.5 | |
| | Supervision | Count | 27 | 4 | 31 |
| | | Expected Count | 27.0 | 4.0 | 31.0 |
| | | % within R-PSFs (lite) | 87.1% | 12.9% | 100.0% |
| | | % of Total | 1.6% | .2% | 1.8% |
| | | Std. Residual | .0 | .0 | |
| | Workload – Time pressure – Stress | Count | 56 | 2 | 58 |
| | | Expected Count | 50.5 | 7.5 | 58.0 |
| | | % within R-PSFs (lite) | 96.6% | 3.4% | 100.0% |
| | | % of Total | 3.3% | .1% | 3.5% |
| | | Std. Residual | .8 | -2.0 | |
| Total | | Count | 1459 | 217 | 1676 |
| | | Expected Count | 1459.0 | 217.0 | 1676.0 |
| | | % within R-PSFs (lite) | 87.1% | 12.9% | 100.0% |
| | | % of Total | 87.1% | 12.9% | 100.0% |

Table VIII-3

**R-PSFs (lite) * Years period Crosstabulation**

| | | | Years period | | |
|---|---|---|---|---|---|
| | | | 1997 – 2005 | 2006 – 2011 | Total |
| R-PSFs (lite) | Safety culture – SMS | Count | 73 | 211 | 284 |
| | | Expected Count | 69.0 | 215.0 | 284.0 |
| | | % within R-PSFs (lite) | 25.7% | 74.3% | 100.0% |
| | | Std. Residual | .5 | -.3 | |
| | System design – Visibility – HMI – RCM | Count | 62 | 125 | 187 |
| | | Expected Count | 45.4 | 141.6 | 187.0 |
| | | % within R-PSFs (lite) | 33.2% | 66.8% | 100.0% |
| | | Std. Residual | 2.5 | -1.4 | |
| | Fatigue – Shift pattern – Fit to work | Count | 39 | 70 | 109 |
| | | Expected Count | 26.5 | 82.5 | 109.0 |
| | | % within R-PSFs (lite) | 35.8% | 64.2% | 100.0% |
| | | Std. Residual | 2.4 | -1.4 | |
| | Communication – Teamwork | Count | 41 | 187 | 228 |
| | | Expected Count | 55.4 | 172.6 | 228.0 |
| | | % within R-PSFs (lite) | 18.0% | 82.0% | 100.0% |
| | | Std. Residual | -1.9 | 1.1 | |
| | Quality of procedures | Count | 28 | 101 | 129 |
| | | Expected Count | 31.3 | 97.7 | 129.0 |
| | | % within R-PSFs (lite) | 21.7% | 78.3% | 100.0% |
| | | Std. Residual | -.6 | .3 | |
| | Distraction – Loss of concentration – Vigilance – Risk awareness | Count | 54 | 189 | 243 |
| | | Expected Count | 59.0 | 184.0 | 243.0 |
| | | % within R-PSFs (lite) | 22.2% | 77.8% | 100.0% |
| | | Std. Residual | -.7 | .4 | |
| | Perception – Interpretation | Count | 43 | 76 | 119 |
| | | Expected Count | 28.9 | 90.1 | 119.0 |
| | | % within R-PSFs (lite) | 36.1% | 63.9% | 100.0% |
| | | Std. Residual | 2.6 | -1.5 | |
| | Training – Experience | Count | 15 | 117 | 132 |
| | | Expected Count | 32.1 | 99.9 | 132.0 |
| | | % within R-PSFs (lite) | 11.4% | 88.6% | 100.0% |
| | | Std. Residual | -3.0 | 1.7 | |
| | Expectation – Familiarity – Routine | Count | 23 | 77 | 100 |
| | | Expected Count | 24.3 | 75.7 | 100.0 |
| | | % within R-PSFs (lite) | 23.0% | 77.0% | 100.0% |
| | | Std. Residual | -.3 | .1 | |
| | Quality of information | Count | 11 | 45 | 56 |
| | | Expected Count | 13.6 | 42.4 | 56.0 |
| | | % within R-PSFs (lite) | 19.6% | 80.4% | 100.0% |
| | | Std. Residual | -.7 | .4 | |
| | Supervision | Count | 5 | 26 | 31 |
| | | Expected Count | 7.5 | 23.5 | 31.0 |
| | | % within R-PSFs (lite) | 16.1% | 83.9% | 100.0% |
| | | Std. Residual | -.9 | .5 | |
| | Workload – Time pressure – Stress | Count | 13 | 45 | 58 |
| | | Expected Count | 14.1 | 43.9 | 58.0 |
| | | % within R-PSFs (lite) | 22.4% | 77.6% | 100.0% |
| | | Std. Residual | -.3 | .2 | |
| Total | | Count | 407 | 1269 | 1676 |
| | | Expected Count | 407.0 | 1269.0 | 1676.0 |
| | | % within R-PSFs (lite) | 24.3% | 75.7% | 100.0% |

Table VIII-4

**R-PSFs (lite) * Months period Crosstabulation**

| | | | Months period | | | | |
|---|---|---|---|---|---|---|---|
| | | | Dec - Feb | Mar - May | Jun - Aug | Sep - Nov | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 86 | 53 | 78 | 67 | 284 |
| | | Expected Count | 73.5 | 67.8 | 77.9 | 64.7 | 284.0 |
| | | % within R-PSFs (lite) | 30.3% | 18.7% | 27.5% | 23.6% | 100.0% |
| | | Std. Residual | 1.5 | -1.8 | .0 | .3 | |
| | System design - Visibility - HMI - RCM | Count | 48 | 40 | 43 | 56 | 187 |
| | | Expected Count | 48.4 | 44.6 | 51.3 | 42.6 | 187.0 |
| | | % within R-PSFs (lite) | 25.7% | 21.4% | 23.0% | 29.9% | 100.0% |
| | | Std. Residual | -.1 | -.7 | -1.2 | 2.0 | |
| | Fatigue - Shift pattern - Fit to work | Count | 29 | 25 | 23 | 32 | 109 |
| | | Expected Count | 28.2 | 26.0 | 29.9 | 24.8 | 109.0 |
| | | % within R-PSFs (lite) | 26.6% | 22.9% | 21.1% | 29.4% | 100.0% |
| | | Std. Residual | .1 | -.2 | -1.3 | 1.4 | |
| | Communication - Teamwork | Count | 57 | 66 | 61 | 44 | 228 |
| | | Expected Count | 59.0 | 54.4 | 62.6 | 52.0 | 228.0 |
| | | % within R-PSFs (lite) | 25.0% | 28.9% | 26.8% | 19.3% | 100.0% |
| | | Std. Residual | -.3 | 1.6 | -.2 | -1.1 | |
| | Quality of procedures | Count | 30 | 38 | 30 | 31 | 129 |
| | | Expected Count | 33.4 | 30.8 | 35.4 | 29.4 | 129.0 |
| | | % within R-PSFs (lite) | 23.3% | 29.5% | 23.3% | 24.0% | 100.0% |
| | | Std. Residual | -.6 | 1.3 | -.9 | .3 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 56 | 59 | 80 | 48 | 243 |
| | | Expected Count | 62.9 | 58.0 | 66.7 | 55.4 | 243.0 |
| | | % within R-PSFs (lite) | 23.0% | 24.3% | 32.9% | 19.8% | 100.0% |
| | | Std. Residual | -.9 | .1 | 1.6 | -1.0 | |
| | Perception - Interpretation | Count | 31 | 20 | 40 | 28 | 119 |
| | | Expected Count | 30.8 | 28.4 | 32.7 | 27.1 | 119.0 |
| | | % within R-PSFs (lite) | 26.1% | 16.8% | 33.6% | 23.5% | 100.0% |
| | | Std. Residual | .0 | -1.6 | 1.3 | .2 | |
| | Training - Experience | Count | 41 | 30 | 33 | 28 | 132 |
| | | Expected Count | 34.2 | 31.5 | 36.2 | 30.1 | 132.0 |
| | | % within R-PSFs (lite) | 31.1% | 22.7% | 25.0% | 21.2% | 100.0% |
| | | Std. Residual | 1.2 | -.3 | -.5 | -.4 | |
| | Expectation - Familiarity - Routine | Count | 20 | 27 | 32 | 21 | 100 |
| | | Expected Count | 25.9 | 23.9 | 27.4 | 22.8 | 100.0 |
| | | % within R-PSFs (lite) | 20.0% | 27.0% | 32.0% | 21.0% | 100.0% |
| | | Std. Residual | -1.2 | .6 | .9 | -.4 | |
| | Quality of information | Count | 17 | 17 | 12 | 10 | 56 |
| | | Expected Count | 14.5 | 13.4 | 15.4 | 12.8 | 56.0 |
| | | % within R-PSFs (lite) | 30.4% | 30.4% | 21.4% | 17.9% | 100.0% |
| | | Std. Residual | .7 | 1.0 | -.9 | -.8 | |
| | Supervision | Count | 11 | 6 | 10 | 4 | 31 |
| | | Expected Count | 8.0 | 7.4 | 8.5 | 7.1 | 31.0 |
| | | % within R-PSFs (lite) | 35.5% | 19.4% | 32.3% | 12.9% | 100.0% |
| | | Std. Residual | 1.0 | -.5 | .5 | -1.2 | |
| | Workload - Time pressure - Stress | Count | 8 | 19 | 18 | 13 | 58 |
| | | Expected Count | 15.0 | 13.8 | 15.9 | 13.2 | 58.0 |
| | | % within R-PSFs (lite) | 13.8% | 32.8% | 31.0% | 22.4% | 100.0% |
| | | Std. Residual | -1.8 | 1.4 | .5 | -.1 | |
| Total | | Count | 434 | 400 | 460 | 382 | 1676 |
| | | Expected Count | 434.0 | 400.0 | 460.0 | 382.0 | 1676.0 |
| | | % within R-PSFs (lite) | 25.9% | 23.9% | 27.4% | 22.8% | 100.0% |

Table VIII-5

**R-PSFs (lite) * Time of day (groups) Crosstabulation**

| | | | Time of day (groups) | | | |
|---|---|---|---|---|---|---|
| | | | Day | Night | Dawn - Dusk | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 199 | 73 | 8 | 280 |
| | | Expected Count | 179.9 | 88.3 | 11.8 | 280.0 |
| | | % within R-PSFs (lite) | 71.1% | 26.1% | 2.9% | 100.0% |
| | | Std. Residual | 1.4 | -1.6 | -1.1 | |
| | System design - Visibility - HMI - RCM | Count | 131 | 43 | 7 | 181 |
| | | Expected Count | 116.3 | 57.1 | 7.6 | 181.0 |
| | | % within R-PSFs (lite) | 72.4% | 23.8% | 3.9% | 100.0% |
| | | Std. Residual | 1.4 | -1.9 | -.2 | |
| | Fatigue - Shift pattern - Fit to work | Count | 51 | 53 | 4 | 108 |
| | | Expected Count | 69.4 | 34.1 | 4.6 | 108.0 |
| | | % within R-PSFs (lite) | 47.2% | 49.1% | 3.7% | 100.0% |
| | | Std. Residual | -2.2 | 3.2 | -.3 | |
| | Communication - Teamwork | Count | 146 | 77 | 4 | 227 |
| | | Expected Count | 145.8 | 71.6 | 9.6 | 227.0 |
| | | % within R-PSFs (lite) | 64.3% | 33.9% | 1.8% | 100.0% |
| | | Std. Residual | .0 | .6 | -1.8 | |
| | Quality of procedures | Count | 85 | 34 | 7 | 126 |
| | | Expected Count | 80.9 | 39.7 | 5.3 | 126.0 |
| | | % within R-PSFs (lite) | 67.5% | 27.0% | 5.6% | 100.0% |
| | | Std. Residual | .5 | -.9 | .7 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 146 | 75 | 14 | 235 |
| | | Expected Count | 151.0 | 74.1 | 9.9 | 235.0 |
| | | % within R-PSFs (lite) | 62.1% | 31.9% | 6.0% | 100.0% |
| | | Std. Residual | -.4 | .1 | 1.3 | |
| | Perception - Interpretation | Count | 74 | 30 | 9 | 113 |
| | | Expected Count | 72.6 | 35.6 | 4.8 | 113.0 |
| | | % within R-PSFs (lite) | 65.5% | 26.5% | 8.0% | 100.0% |
| | | Std. Residual | .2 | -.9 | 1.9 | |
| | Training - Experience | Count | 73 | 47 | 8 | 128 |
| | | Expected Count | 82.2 | 40.4 | 5.4 | 128.0 |
| | | % within R-PSFs (lite) | 57.0% | 36.7% | 6.3% | 100.0% |
| | | Std. Residual | -1.0 | 1.0 | 1.1 | |
| | Expectation - Familiarity - Routine | Count | 56 | 37 | 4 | 97 |
| | | Expected Count | 62.3 | 30.6 | 4.1 | 97.0 |
| | | % within R-PSFs (lite) | 57.7% | 38.1% | 4.1% | 100.0% |
| | | Std. Residual | -.8 | 1.2 | .0 | |
| | Quality of information | Count | 39 | 16 | 0 | 55 |
| | | Expected Count | 35.3 | 17.3 | 2.3 | 55.0 |
| | | % within R-PSFs (lite) | 70.9% | 29.1% | .0% | 100.0% |
| | | Std. Residual | .6 | -.3 | -1.5 | |
| | Supervision | Count | 18 | 12 | 1 | 31 |
| | | Expected Count | 19.9 | 9.8 | 1.3 | 31.0 |
| | | % within R-PSFs (lite) | 58.1% | 38.7% | 3.2% | 100.0% |
| | | Std. Residual | -.4 | .7 | -.3 | |
| | Workload - Time pressure - Stress | Count | 33 | 19 | 3 | 55 |
| | | Expected Count | 35.3 | 17.3 | 2.3 | 55.0 |
| | | % within R-PSFs (lite) | 60.0% | 34.5% | 5.5% | 100.0% |
| | | Std. Residual | -.4 | .4 | .4 | |
| Total | | Count | 1051 | 516 | 69 | 1636 |
| | | Expected Count | 1051.0 | 516.0 | 69.0 | 1636.0 |
| | | % within R-PSFs (lite) | 64.2% | 31.5% | 4.2% | 100.0% |

Table VIII-6

R-PSFs (lite) * Responsibility (groups) Crosstabulation

| | | | Responsibility (groups) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Passenger train driver | Freight train driver | Locomotive – Engineering train driver | Metro – Tram driver | Signaller – Controller | Total |
| R-PSFs (lite) | Safety culture – SMS | Count | 87 | 56 | 38 | 13 | 90 | 284 |
| | | Expected Count | 90.5 | 57.4 | 27.3 | 12.2 | 96.6 | 284.0 |
| | | % within R-PSFs (lite) | 30.6% | 19.7% | 13.4% | 4.6% | 31.7% | 100.0% |
| | | Std. Residual | -.4 | -.2 | 2.1 | .2 | -.7 | |
| | System design – Visibility – HMI – RCM | Count | 79 | 55 | 10 | 8 | 35 | 187 |
| | | Expected Count | 59.6 | 37.8 | 18.0 | 8.0 | 63.6 | 187.0 |
| | | % within R-PSFs (lite) | 42.2% | 29.4% | 5.3% | 4.3% | 18.7% | 100.0% |
| | | Std. Residual | 2.5 | 2.8 | -1.9 | .0 | -3.6 | |
| | Fatigue – Shift pattern – Fit to work | Count | 23 | 49 | 8 | 4 | 25 | 109 |
| | | Expected Count | 34.7 | 22.0 | 10.5 | 4.7 | 37.1 | 109.0 |
| | | % within R-PSFs (lite) | 21.1% | 45.0% | 7.3% | 3.7% | 22.9% | 100.0% |
| | | Std. Residual | -2.0 | 5.7 | -.8 | -.3 | -2.0 | |
| | Communication – Teamwork | Count | 36 | 35 | 24 | 7 | 126 | 228 |
| | | Expected Count | 72.6 | 46.1 | 21.9 | 9.8 | 77.5 | 228.0 |
| | | % within R-PSFs (lite) | 15.8% | 15.4% | 10.5% | 3.1% | 55.3% | 100.0% |
| | | Std. Residual | -4.3 | -1.6 | .4 | -.9 | 5.5 | |
| | Quality of procedures | Count | 33 | 20 | 12 | 7 | 57 | 129 |
| | | Expected Count | 41.1 | 26.1 | 12.4 | 5.5 | 43.9 | 129.0 |
| | | % within R-PSFs (lite) | 25.6% | 15.5% | 9.3% | 5.4% | 44.2% | 100.0% |
| | | Std. Residual | -1.3 | -1.2 | -.1 | .6 | 2.0 | |
| | Distraction – Loss of concentration – Vigilance – Risk awareness | Count | 106 | 49 | 18 | 14 | 56 | 243 |
| | | Expected Count | 77.4 | 49.2 | 23.3 | 10.4 | 82.6 | 243.0 |
| | | % within R-PSFs (lite) | 43.6% | 20.2% | 7.4% | 5.8% | 23.0% | 100.0% |
| | | Std. Residual | 3.2 | .0 | -1.1 | 1.1 | -2.9 | |
| | Perception – Interpretation | Count | 51 | 23 | 10 | 3 | 32 | 119 |
| | | Expected Count | 37.9 | 24.1 | 11.4 | 5.1 | 40.5 | 119.0 |
| | | % within R-PSFs (lite) | 42.9% | 19.3% | 8.4% | 2.5% | 26.9% | 100.0% |
| | | Std. Residual | 2.1 | -.2 | -.4 | -.9 | -1.3 | |
| | Training – Experience | Count | 47 | 14 | 12 | 9 | 50 | 132 |
| | | Expected Count | 42.1 | 26.7 | 12.7 | 5.7 | 44.9 | 132.0 |
| | | % within R-PSFs (lite) | 35.6% | 10.6% | 9.1% | 6.8% | 37.9% | 100.0% |
| | | Std. Residual | .8 | -2.5 | -.2 | 1.4 | .8 | |
| | Expectation – Familiarity – Routine | Count | 44 | 19 | 15 | 5 | 17 | 100 |
| | | Expected Count | 31.9 | 20.2 | 9.6 | 4.3 | 34.0 | 100.0 |
| | | % within R-PSFs (lite) | 44.0% | 19.0% | 15.0% | 5.0% | 17.0% | 100.0% |
| | | Std. Residual | 2.2 | -.3 | 1.7 | .3 | -2.9 | |
| | Quality of information | Count | 11 | 11 | 8 | 0 | 26 | 56 |
| | | Expected Count | 17.8 | 11.3 | 5.4 | 2.4 | 19.0 | 56.0 |
| | | % within R-PSFs (lite) | 19.6% | 19.6% | 14.3% | .0% | 46.4% | 100.0% |
| | | Std. Residual | -1.6 | -.1 | 1.1 | -1.6 | 1.6 | |
| | Supervision | Count | 4 | 5 | 3 | 1 | 18 | 31 |
| | | Expected Count | 9.9 | 6.3 | 3.0 | 1.3 | 10.5 | 31.0 |
| | | % within R-PSFs (lite) | 12.9% | 16.1% | 9.7% | 3.2% | 58.1% | 100.0% |
| | | Std. Residual | -1.9 | -.5 | .0 | -.3 | 2.3 | |
| | Workload – Time pressure – Stress | Count | 13 | 3 | 3 | 1 | 38 | 58 |
| | | Expected Count | 18.5 | 11.7 | 5.6 | 2.5 | 19.7 | 58.0 |
| | | % within R-PSFs (lite) | 22.4% | 5.2% | 5.2% | 1.7% | 65.5% | 100.0% |
| | | Std. Residual | -1.3 | -2.5 | -1.1 | -.9 | 4.1 | |
| Total | | Count | 534 | 339 | 161 | 72 | 570 | 1676 |
| | | Expected Count | 534.0 | 339.0 | 161.0 | 72.0 | 570.0 | 1676.0 |
| | | % within R-PSFs (lite) | 31.9% | 20.2% | 9.6% | 4.3% | 34.0% | 100.0% |

Table VIII-7

**R-PSFs (lite) * Location Crosstabulation**

| | | | Location | | | | |
|---|---|---|---|---|---|---|---|
| | | | Open line (incl. curves, gradient, points) | Station (incl. platforms, depots, points) | Yard | Junction | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 108 | 96 | 27 | 36 | 267 |
| | | Expected Count | 108.5 | 89.3 | 23.9 | 45.3 | 267.0 |
| | | % within R-PSFs (lite) | 40.4% | 36.0% | 10.1% | 13.5% | 100.0% |
| | | Std. Residual | .0 | .7 | .6 | -1.4 | |
| | System design - Visibility - HMI - RCM | Count | 89 | 49 | 9 | 30 | 177 |
| | | Expected Count | 71.9 | 59.2 | 15.9 | 30.0 | 177.0 |
| | | % within R-PSFs (lite) | 50.3% | 27.7% | 5.1% | 16.9% | 100.0% |
| | | Std. Residual | 2.0 | -1.3 | -1.7 | .0 | |
| | Fatigue - Shift pattern - Fit to work | Count | 49 | 20 | 7 | 22 | 98 |
| | | Expected Count | 39.8 | 32.8 | 8.8 | 16.6 | 98.0 |
| | | % within R-PSFs (lite) | 50.0% | 20.4% | 7.1% | 22.4% | 100.0% |
| | | Std. Residual | 1.5 | -2.2 | -.6 | 1.3 | |
| | Communication - Teamwork | Count | 80 | 78 | 24 | 36 | 218 |
| | | Expected Count | 88.6 | 72.9 | 19.5 | 37.0 | 218.0 |
| | | % within R-PSFs (lite) | 36.7% | 35.8% | 11.0% | 16.5% | 100.0% |
| | | Std. Residual | -.9 | .6 | 1.0 | -.2 | |
| | Quality of procedures | Count | 56 | 34 | 7 | 26 | 123 |
| | | Expected Count | 50.0 | 41.2 | 11.0 | 20.9 | 123.0 |
| | | % within R-PSFs (lite) | 45.5% | 27.6% | 5.7% | 21.1% | 100.0% |
| | | Std. Residual | .9 | -1.1 | -1.2 | 1.1 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 89 | 79 | 16 | 42 | 226 |
| | | Expected Count | 91.8 | 75.6 | 20.2 | 38.3 | 226.0 |
| | | % within R-PSFs (lite) | 39.4% | 35.0% | 7.1% | 18.6% | 100.0% |
| | | Std. Residual | -.3 | .4 | -.9 | .6 | |
| | Perception - Interpretation | Count | 38 | 47 | 12 | 15 | 112 |
| | | Expected Count | 45.5 | 37.5 | 10.0 | 19.0 | 112.0 |
| | | % within R-PSFs (lite) | 33.9% | 42.0% | 10.7% | 13.4% | 100.0% |
| | | Std. Residual | -1.1 | 1.6 | .6 | -.9 | |
| | Training - Experience | Count | 40 | 54 | 9 | 17 | 120 |
| | | Expected Count | 48.8 | 40.2 | 10.7 | 20.3 | 120.0 |
| | | % within R-PSFs (lite) | 33.3% | 45.0% | 7.5% | 14.2% | 100.0% |
| | | Std. Residual | -1.3 | 2.2 | -.5 | -.7 | |
| | Expectation - Familiarity - Routine | Count | 32 | 25 | 15 | 13 | 85 |
| | | Expected Count | 34.5 | 28.4 | 7.6 | 14.4 | 85.0 |
| | | % within R-PSFs (lite) | 37.6% | 29.4% | 17.6% | 15.3% | 100.0% |
| | | Std. Residual | -.4 | -.6 | 2.7 | -.4 | |
| | Quality of information | Count | 19 | 17 | 7 | 9 | 52 |
| | | Expected Count | 21.1 | 17.4 | 4.7 | 8.8 | 52.0 |
| | | % within R-PSFs (lite) | 36.5% | 32.7% | 13.5% | 17.3% | 100.0% |
| | | Std. Residual | -.5 | -.1 | 1.1 | .1 | |
| | Supervision | Count | 16 | 4 | 2 | 8 | 30 |
| | | Expected Count | 12.2 | 10.0 | 2.7 | 5.1 | 30.0 |
| | | % within R-PSFs (lite) | 53.3% | 13.3% | 6.7% | 26.7% | 100.0% |
| | | Std. Residual | 1.1 | -1.9 | -.4 | 1.3 | |
| | Workload - Time pressure - Stress | Count | 19 | 20 | 5 | 11 | 55 |
| | | Expected Count | 22.3 | 18.4 | 4.9 | 9.3 | 55.0 |
| | | % within R-PSFs (lite) | 34.5% | 36.4% | 9.1% | 20.0% | 100.0% |
| | | Std. Residual | -.7 | .4 | .0 | .5 | |
| Total | | Count | 635 | 523 | 140 | 265 | 1563 |
| | | Expected Count | 635.0 | 523.0 | 140.0 | 265.0 | 1563.0 |
| | | % within R-PSFs (lite) | 40.6% | 33.5% | 9.0% | 17.0% | 100.0% |

Table VIII-8

**R-PSFs (lite) * Location 2 Crosstabulation**

| | | | Location 2 | | | |
|---|---|---|---|---|---|---|
| | | | Open line (incl. curves, gradient) | Points (incl. open lines, stations, depots, junctions, yards) | Station (incl. platforms, depots) | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 106 | 92 | 69 | 267 |
| | | Expected Count | 107.3 | 95.0 | 64.7 | 267.0 |
| | | % within R-PSFs (lite) | 39.7% | 34.5% | 25.8% | 100.0% |
| | | Std. Residual | -.1 | -.3 | .5 | |
| | System design - Visibility - HMI - RCM | Count | 88 | 58 | 31 | 177 |
| | | Expected Count | 71.1 | 63.0 | 42.9 | 177.0 |
| | | % within R-PSFs (lite) | 49.7% | 32.8% | 17.5% | 100.0% |
| | | Std. Residual | 2.0 | -.6 | -1.8 | |
| | Fatigue - Shift pattern - Fit to work | Count | 49 | 32 | 17 | 98 |
| | | Expected Count | 39.4 | 34.9 | 23.8 | 98.0 |
| | | % within R-PSFs (lite) | 50.0% | 32.7% | 17.3% | 100.0% |
| | | Std. Residual | 1.5 | -.5 | -1.4 | |
| | Communication - Teamwork | Count | 80 | 72 | 66 | 218 |
| | | Expected Count | 87.6 | 77.5 | 52.9 | 218.0 |
| | | % within R-PSFs (lite) | 36.7% | 33.0% | 30.3% | 100.0% |
| | | Std. Residual | -.8 | -.6 | 1.8 | |
| | Quality of procedures | Count | 55 | 48 | 20 | 123 |
| | | Expected Count | 49.4 | 43.8 | 29.8 | 123.0 |
| | | % within R-PSFs (lite) | 44.7% | 39.0% | 16.3% | 100.0% |
| | | Std. Residual | .8 | .6 | -1.8 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 88 | 77 | 61 | 226 |
| | | Expected Count | 90.8 | 80.4 | 54.8 | 226.0 |
| | | % within R-PSFs (lite) | 38.9% | 34.1% | 27.0% | 100.0% |
| | | Std. Residual | -.3 | -.4 | .8 | |
| | Perception - Interpretation | Count | 37 | 42 | 33 | 112 |
| | | Expected Count | 45.0 | 39.8 | 27.2 | 112.0 |
| | | % within R-PSFs (lite) | 33.0% | 37.5% | 29.5% | 100.0% |
| | | Std. Residual | -1.2 | .3 | 1.1 | |
| | Training - Experience | Count | 39 | 53 | 28 | 120 |
| | | Expected Count | 48.2 | 42.7 | 29.1 | 120.0 |
| | | % within R-PSFs (lite) | 32.5% | 44.2% | 23.3% | 100.0% |
| | | Std. Residual | -1.3 | 1.6 | -.2 | |
| | Expectation - Familiarity - Routine | Count | 32 | 34 | 19 | 85 |
| | | Expected Count | 34.2 | 30.2 | 20.6 | 85.0 |
| | | % within R-PSFs (lite) | 37.6% | 40.0% | 22.4% | 100.0% |
| | | Std. Residual | -.4 | .7 | -.4 | |
| | Quality of information | Count | 19 | 20 | 13 | 52 |
| | | Expected Count | 20.9 | 18.5 | 12.6 | 52.0 |
| | | % within R-PSFs (lite) | 36.5% | 38.5% | 25.0% | 100.0% |
| | | Std. Residual | -.4 | .3 | .1 | |
| | Supervision | Count | 16 | 10 | 4 | 30 |
| | | Expected Count | 12.1 | 10.7 | 7.3 | 30.0 |
| | | % within R-PSFs (lite) | 53.3% | 33.3% | 13.3% | 100.0% |
| | | Std. Residual | 1.1 | -.2 | -1.2 | |
| | Workload - Time pressure - Stress | Count | 19 | 18 | 18 | 55 |
| | | Expected Count | 22.1 | 19.6 | 13.3 | 55.0 |
| | | % within R-PSFs (lite) | 34.5% | 32.7% | 32.7% | 100.0% |
| | | Std. Residual | -.7 | -.4 | 1.3 | |
| Total | | Count | 628 | 556 | 379 | 1563 |
| | | Expected Count | 628.0 | 556.0 | 379.0 | 1563.0 |
| | | % within R-PSFs (lite) | 40.2% | 35.6% | 24.2% | 100.0% |

Table VIII-9

R-PSFs (lite) * Type of occurrence Crosstabulation

| | | | Type of occurrence | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Trains collision | Train derailment | SPAD | SPAD (ground) | Possession irregularity | Operational irregularity | Train collision with an obstacle | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 119 | 49 | 35 | 8 | 6 | 39 | 19 | 275 |
| | | Expected Count | 92.3 | 39.7 | 62.3 | 13.2 | 9.0 | 37.9 | 20.6 | 275.0 |
| | | % within R-PSFs (lite) | 43.3% | 17.8% | 12.7% | 2.9% | 2.2% | 14.2% | 6.9% | 100.0% |
| | | Std. Residual | 2.8 | 1.5 | -3.5 | -1.4 | -1.0 | .2 | -.4 | |
| | System design - Visibility - HMI - RCM | Count | 81 | 21 | 48 | 9 | 1 | 13 | 10 | 183 |
| | | Expected Count | 61.4 | 26.4 | 41.5 | 8.8 | 6.0 | 25.2 | 13.7 | 183.0 |
| | | % within R-PSFs (lite) | 44.3% | 11.5% | 26.2% | 4.9% | .5% | 7.1% | 5.5% | 100.0% |
| | | Std. Residual | 2.5 | -1.1 | 1.0 | .1 | -2.0 | -2.4 | -1.0 | |
| | Fatigue - Shift pattern - Fit to work | Count | 50 | 16 | 15 | 4 | 5 | 13 | 6 | 109 |
| | | Expected Count | 36.6 | 15.7 | 24.7 | 5.2 | 3.6 | 15.0 | 8.2 | 109.0 |
| | | % within R-PSFs (lite) | 45.9% | 14.7% | 13.8% | 3.7% | 4.6% | 11.9% | 5.5% | 100.0% |
| | | Std. Residual | 2.2 | .1 | -2.0 | -.5 | .7 | -.5 | -.8 | |
| | Communication - Teamwork | Count | 76 | 27 | 41 | 10 | 14 | 37 | 19 | 224 |
| | | Expected Count | 75.2 | 32.3 | 50.7 | 10.8 | 7.4 | 30.8 | 16.8 | 224.0 |
| | | % within R-PSFs (lite) | 33.9% | 12.1% | 18.3% | 4.5% | 6.3% | 16.5% | 8.5% | 100.0% |
| | | Std. Residual | .1 | -.9 | -1.4 | -.2 | 2.4 | 1.1 | .5 | |
| | Quality of procedures | Count | 34 | 21 | 22 | 5 | 5 | 26 | 11 | 124 |
| | | Expected Count | 41.6 | 17.9 | 28.1 | 6.0 | 4.1 | 17.1 | 9.3 | 124.0 |
| | | % within R-PSFs (lite) | 27.4% | 16.9% | 17.7% | 4.0% | 4.0% | 21.0% | 8.9% | 100.0% |
| | | Std. Residual | -1.2 | .7 | -1.1 | -.4 | .5 | 2.2 | .6 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 72 | 35 | 69 | 15 | 6 | 26 | 16 | 239 |
| | | Expected Count | 80.2 | 34.5 | 54.1 | 11.5 | 7.9 | 32.9 | 17.9 | 239.0 |
| | | % within R-PSFs (lite) | 30.1% | 14.6% | 28.9% | 6.3% | 2.5% | 10.9% | 6.7% | 100.0% |
| | | Std. Residual | -.9 | .1 | 2.0 | 1.0 | -.7 | -1.2 | -.4 | |
| | Perception - Interpretation | Count | 36 | 24 | 28 | 6 | 3 | 10 | 10 | 117 |
| | | Expected Count | 39.3 | 16.9 | 26.5 | 5.6 | 3.8 | 16.1 | 8.8 | 117.0 |
| | | % within R-PSFs (lite) | 30.8% | 20.5% | 23.9% | 5.1% | 2.6% | 8.5% | 8.5% | 100.0% |
| | | Std. Residual | -.5 | 1.7 | .3 | .2 | -.4 | -1.5 | .4 | |
| | Training - Experience | Count | 21 | 17 | 43 | 9 | 5 | 23 | 12 | 130 |
| | | Expected Count | 43.6 | 18.8 | 29.5 | 6.3 | 4.3 | 17.9 | 9.7 | 130.0 |
| | | % within R-PSFs (lite) | 16.2% | 13.1% | 33.1% | 6.9% | 3.8% | 17.7% | 9.2% | 100.0% |
| | | Std. Residual | -3.4 | -.4 | 2.5 | 1.1 | .4 | 1.2 | .7 | |
| | Expectation - Familiarity - Routine | Count | 22 | 9 | 44 | 6 | 0 | 8 | 8 | 97 |
| | | Expected Count | 32.5 | 14.0 | 22.0 | 4.7 | 3.2 | 13.4 | 7.3 | 97.0 |
| | | % within R-PSFs (lite) | 22.7% | 9.3% | 45.4% | 6.2% | .0% | 8.2% | 8.2% | 100.0% |
| | | Std. Residual | -1.8 | -1.3 | 4.7 | .6 | -1.8 | -1.5 | .3 | |
| | Quality of information | Count | 19 | 7 | 14 | 4 | 1 | 8 | 3 | 56 |
| | | Expected Count | 18.8 | 8.1 | 12.7 | 2.7 | 1.8 | 7.7 | 4.2 | 56.0 |
| | | % within R-PSFs (lite) | 33.9% | 12.5% | 25.0% | 7.1% | 1.8% | 14.3% | 5.4% | 100.0% |
| | | Std. Residual | .0 | -.4 | .4 | .8 | -.6 | .1 | -.6 | |
| | Supervision | Count | 7 | 5 | 4 | 0 | 5 | 8 | 2 | 31 |
| | | Expected Count | 10.4 | 4.5 | 7.0 | 1.5 | 1.0 | 4.3 | 2.3 | 31.0 |
| | | % within R-PSFs (lite) | 22.6% | 16.1% | 12.9% | .0% | 16.1% | 25.8% | 6.5% | 100.0% |
| | | Std. Residual | -1.1 | .2 | -1.1 | -1.2 | 3.9 | 1.8 | -.2 | |
| | Workload - Time pressure - Stress | Count | 14 | 6 | 9 | 3 | 3 | 15 | 7 | 57 |
| | | Expected Count | 19.1 | 8.2 | 12.9 | 2.7 | 1.9 | 7.8 | 4.3 | 57.0 |
| | | % within R-PSFs (lite) | 24.6% | 10.5% | 15.8% | 5.3% | 5.3% | 26.3% | 12.3% | 100.0% |
| | | Std. Residual | -1.2 | -.8 | -1.1 | .2 | .8 | 2.6 | 1.3 | |
| Total | | Count | 551 | 237 | 372 | 79 | 54 | 226 | 123 | 1642 |
| | | Expected Count | 551.0 | 237.0 | 372.0 | 79.0 | 54.0 | 226.0 | 123.0 | 1642.0 |
| | | % within R-PSFs (lite) | 33.6% | 14.4% | 22.7% | 4.8% | 3.3% | 13.8% | 7.5% | 100.0% |

Table VIII-10

**R-PSFs (lite) * Associated events Crosstabulation**

| | | | Associated events | | | |
|---|---|---|---|---|---|---|
| | | | SPAD | SPAD (ground) | Operational irregularity | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 71 | 6 | 117 | 194 |
| | | Expected Count | 73.5 | 9.6 | 110.9 | 194.0 |
| | | % within R-PSFs (lite) | 36.6% | 3.1% | 60.3% | 100.0% |
| | | Std. Residual | -.3 | -1.2 | .6 | |
| | System design - Visibility - HMI - RCM | Count | 52 | 6 | 57 | 115 |
| | | Expected Count | 43.5 | 5.7 | 65.7 | 115.0 |
| | | % within R-PSFs (lite) | 45.2% | 5.2% | 49.6% | 100.0% |
| | | Std. Residual | 1.3 | .1 | -1.1 | |
| | Fatigue - Shift pattern - Fit to work | Count | 43 | 1 | 28 | 72 |
| | | Expected Count | 27.3 | 3.6 | 41.2 | 72.0 |
| | | % within R-PSFs (lite) | 59.7% | 1.4% | 38.9% | 100.0% |
| | | Std. Residual | 3.0 | -1.4 | -2.1 | |
| | Communication - Teamwork | Count | 33 | 5 | 85 | 123 |
| | | Expected Count | 46.6 | 6.1 | 70.3 | 123.0 |
| | | % within R-PSFs (lite) | 26.8% | 4.1% | 69.1% | 100.0% |
| | | Std. Residual | -2.0 | -.4 | 1.8 | |
| | Quality of procedures | Count | 21 | 3 | 43 | 67 |
| | | Expected Count | 25.4 | 3.3 | 38.3 | 67.0 |
| | | % within R-PSFs (lite) | 31.3% | 4.5% | 64.2% | 100.0% |
| | | Std. Residual | -.9 | -.2 | .8 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 57 | 4 | 63 | 124 |
| | | Expected Count | 47.0 | 6.2 | 70.9 | 124.0 |
| | | % within R-PSFs (lite) | 46.0% | 3.2% | 50.8% | 100.0% |
| | | Std. Residual | 1.5 | -.9 | -.9 | |
| | Perception - Interpretation | Count | 20 | 9 | 41 | 70 |
| | | Expected Count | 26.5 | 3.5 | 40.0 | 70.0 |
| | | % within R-PSFs (lite) | 28.6% | 12.9% | 58.6% | 100.0% |
| | | Std. Residual | -1.3 | 3.0 | .2 | |
| | Training - Experience | Count | 18 | 3 | 31 | 52 |
| | | Expected Count | 19.7 | 2.6 | 29.7 | 52.0 |
| | | % within R-PSFs (lite) | 34.6% | 5.8% | 59.6% | 100.0% |
| | | Std. Residual | -.4 | .3 | .2 | |
| | Expectation - Familiarity - Routine | Count | 20 | 4 | 15 | 39 |
| | | Expected Count | 14.8 | 1.9 | 22.3 | 39.0 |
| | | % within R-PSFs (lite) | 51.3% | 10.3% | 38.5% | 100.0% |
| | | Std. Residual | 1.4 | 1.5 | -1.5 | |
| | Quality of information | Count | 6 | 3 | 20 | 29 |
| | | Expected Count | 11.0 | 1.4 | 16.6 | 29.0 |
| | | % within R-PSFs (lite) | 20.7% | 10.3% | 69.0% | 100.0% |
| | | Std. Residual | -1.5 | 1.3 | .8 | |
| | Supervision | Count | 4 | 0 | 10 | 14 |
| | | Expected Count | 5.3 | .7 | 8.0 | 14.0 |
| | | % within R-PSFs (lite) | 28.6% | .0% | 71.4% | 100.0% |
| | | Std. Residual | -.6 | -.8 | .7 | |
| | Workload - Time pressure - Stress | Count | 6 | 2 | 20 | 28 |
| | | Expected Count | 10.6 | 1.4 | 16.0 | 28.0 |
| | | % within R-PSFs (lite) | 21.4% | 7.1% | 71.4% | 100.0% |
| | | Std. Residual | -1.4 | .5 | 1.0 | |
| Total | | Count | 351 | 46 | 530 | 927 |
| | | Expected Count | 351.0 | 46.0 | 530.0 | 927.0 |
| | | % within R-PSFs (lite) | 37.9% | 5.0% | 57.2% | 100.0% |

Table VIII-11

**R-PSFs (lite) * Immediate cause (groups) Crosstabulation**

| | | | Immediate cause (groups) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Signalling | Excessive speed | Braking | Routing | Points set | Information | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 99 | 30 | 21 | 54 | 14 | 20 | 238 |
| | | Expected Count | 118.2 | 25.2 | 18.3 | 40.7 | 14.0 | 21.7 | 238.0 |
| | | % within R-PSFs (lite) | 41.6% | 12.6% | 8.8% | 22.7% | 5.9% | 8.4% | 100.0% |
| | | Std. Residual | -1.8 | 1.0 | .6 | 2.1 | .0 | -.4 | |
| | System design - Visibility - HMI - RCM | Count | 106 | 18 | 11 | 20 | 8 | 11 | 174 |
| | | Expected Count | 86.4 | 18.4 | 13.4 | 29.7 | 10.2 | 15.9 | 174.0 |
| | | % within R-PSFs (lite) | 60.9% | 10.3% | 6.3% | 11.5% | 4.6% | 6.3% | 100.0% |
| | | Std. Residual | 2.1 | -.1 | -.6 | -1.8 | -.7 | -1.2 | |
| | Fatigue - Shift pattern - Fit to work | Count | 56 | 17 | 9 | 14 | 4 | 3 | 103 |
| | | Expected Count | 51.1 | 10.9 | 7.9 | 17.6 | 6.1 | 9.4 | 103.0 |
| | | % within R-PSFs (lite) | 54.4% | 16.5% | 8.7% | 13.6% | 3.9% | 2.9% | 100.0% |
| | | Std. Residual | .7 | 1.8 | .4 | -.9 | -.8 | -2.1 | |
| | Communication - Teamwork | Count | 74 | 8 | 11 | 37 | 9 | 42 | 181 |
| | | Expected Count | 89.9 | 19.2 | 13.9 | 30.9 | 10.6 | 16.5 | 181.0 |
| | | % within R-PSFs (lite) | 40.9% | 4.4% | 6.1% | 20.4% | 5.0% | 23.2% | 100.0% |
| | | Std. Residual | -1.7 | -2.5 | -.8 | 1.1 | -.5 | 6.3 | |
| | Quality of procedures | Count | 42 | 12 | 6 | 22 | 11 | 12 | 105 |
| | | Expected Count | 52.1 | 11.1 | 8.1 | 17.9 | 6.2 | 9.6 | 105.0 |
| | | % within R-PSFs (lite) | 40.0% | 11.4% | 5.7% | 21.0% | 10.5% | 11.4% | 100.0% |
| | | Std. Residual | -1.4 | .3 | -.7 | 1.0 | 1.9 | .8 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 125 | 23 | 20 | 32 | 12 | 11 | 223 |
| | | Expected Count | 110.7 | 23.6 | 17.1 | 38.1 | 13.1 | 20.4 | 223.0 |
| | | % within R-PSFs (lite) | 56.1% | 10.3% | 9.0% | 14.3% | 5.4% | 4.9% | 100.0% |
| | | Std. Residual | 1.4 | -.1 | .7 | -1.0 | -.3 | -2.1 | |
| | Perception - Interpretation | Count | 54 | 10 | 5 | 16 | 9 | 7 | 101 |
| | | Expected Count | 50.2 | 10.7 | 7.8 | 17.3 | 5.9 | 9.2 | 101.0 |
| | | % within R-PSFs (lite) | 53.5% | 9.9% | 5.0% | 15.8% | 8.9% | 6.9% | 100.0% |
| | | Std. Residual | .5 | -.2 | -1.0 | -.3 | 1.3 | -.7 | |
| | Training - Experience | Count | 55 | 14 | 13 | 16 | 8 | 7 | 113 |
| | | Expected Count | 56.1 | 12.0 | 8.7 | 19.3 | 6.6 | 10.3 | 113.0 |
| | | % within R-PSFs (lite) | 48.7% | 12.4% | 11.5% | 14.2% | 7.1% | 6.2% | 100.0% |
| | | Std. Residual | -.1 | .6 | 1.5 | -.8 | .5 | -1.0 | |
| | Expectation - Familiarity - Routine | Count | 65 | 6 | 6 | 4 | 2 | 1 | 84 |
| | | Expected Count | 41.7 | 8.9 | 6.4 | 14.3 | 4.9 | 7.7 | 84.0 |
| | | % within R-PSFs (lite) | 77.4% | 7.1% | 7.1% | 4.8% | 2.4% | 1.2% | 100.0% |
| | | Std. Residual | 3.6 | -1.0 | -.2 | -2.7 | -1.3 | -2.4 | |
| | Quality of information | Count | 22 | 4 | 2 | 8 | 0 | 10 | 46 |
| | | Expected Count | 22.8 | 4.9 | 3.5 | 7.9 | 2.7 | 4.2 | 46.0 |
| | | % within R-PSFs (lite) | 47.8% | 8.7% | 4.3% | 17.4% | .0% | 21.7% | 100.0% |
| | | Std. Residual | -.2 | -.4 | -.8 | .1 | -1.6 | 2.8 | |
| | Supervision | Count | 6 | 7 | 1 | 8 | 1 | 3 | 26 |
| | | Expected Count | 12.9 | 2.8 | 2.0 | 4.4 | 1.5 | 2.4 | 26.0 |
| | | % within R-PSFs (lite) | 23.1% | 26.9% | 3.8% | 30.8% | 3.8% | 11.5% | 100.0% |
| | | Std. Residual | -1.9 | 2.6 | -.7 | 1.7 | -.4 | .4 | |
| | Workload - Time pressure - Stress | Count | 14 | 4 | 6 | 16 | 7 | 5 | 52 |
| | | Expected Count | 25.8 | 5.5 | 4.0 | 8.9 | 3.1 | 4.7 | 52.0 |
| | | % within R-PSFs (lite) | 26.9% | 7.7% | 11.5% | 30.8% | 13.5% | 9.6% | 100.0% |
| | | Std. Residual | -2.3 | -.6 | 1.0 | 2.4 | 2.3 | .1 | |
| Total | | Count | 718 | 153 | 111 | 247 | 85 | 132 | 1446 |
| | | Expected Count | 718.0 | 153.0 | 111.0 | 247.0 | 85.0 | 132.0 | 1446.0 |
| | | % within R-PSFs (lite) | 49.7% | 10.6% | 7.7% | 17.1% | 5.9% | 9.1% | 100.0% |

Table VIII-12

**R-PSFs (lite) * Degraded Operations 2 Crosstabulation**

| | | | Degraded Operations 2 | | |
|---|---|---|---|---|---|
| | | | No | Yes | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 250 | 34 | 284 |
| | | Expected Count | 240.8 | 43.2 | 284.0 |
| | | % within R-PSFs (lite) | 88.0% | 12.0% | 100.0% |
| | | Std. Residual | .6 | -1.4 | |
| | System design - Visibility - HMI - RCM | Count | 163 | 24 | 187 |
| | | Expected Count | 158.5 | 28.5 | 187.0 |
| | | % within R-PSFs (lite) | 87.2% | 12.8% | 100.0% |
| | | Std. Residual | .4 | -.8 | |
| | Fatigue - Shift pattern - Fit to work | Count | 99 | 10 | 109 |
| | | Expected Count | 92.4 | 16.6 | 109.0 |
| | | % within R-PSFs (lite) | 90.8% | 9.2% | 100.0% |
| | | Std. Residual | .7 | -1.6 | |
| | Communication - Teamwork | Count | 187 | 41 | 228 |
| | | Expected Count | 193.3 | 34.7 | 228.0 |
| | | % within R-PSFs (lite) | 82.0% | 18.0% | 100.0% |
| | | Std. Residual | -.5 | 1.1 | |
| | Quality of procedures | Count | 97 | 32 | 129 |
| | | Expected Count | 109.4 | 19.6 | 129.0 |
| | | % within R-PSFs (lite) | 75.2% | 24.8% | 100.0% |
| | | Std. Residual | -1.2 | 2.8 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 213 | 30 | 243 |
| | | Expected Count | 206.0 | 37.0 | 243.0 |
| | | % within R-PSFs (lite) | 87.7% | 12.3% | 100.0% |
| | | Std. Residual | .5 | -1.1 | |
| | Perception - Interpretation | Count | 104 | 15 | 119 |
| | | Expected Count | 100.9 | 18.1 | 119.0 |
| | | % within R-PSFs (lite) | 87.4% | 12.6% | 100.0% |
| | | Std. Residual | .3 | -.7 | |
| | Training - Experience | Count | 108 | 24 | 132 |
| | | Expected Count | 111.9 | 20.1 | 132.0 |
| | | % within R-PSFs (lite) | 81.8% | 18.2% | 100.0% |
| | | Std. Residual | -.4 | .9 | |
| | Expectation - Familiarity - Routine | Count | 92 | 8 | 100 |
| | | Expected Count | 84.8 | 15.2 | 100.0 |
| | | % within R-PSFs (lite) | 92.0% | 8.0% | 100.0% |
| | | Std. Residual | .8 | -1.8 | |
| | Quality of information | Count | 41 | 15 | 56 |
| | | Expected Count | 47.5 | 8.5 | 56.0 |
| | | % within R-PSFs (lite) | 73.2% | 26.8% | 100.0% |
| | | Std. Residual | -.9 | 2.2 | |
| | Supervision | Count | 26 | 5 | 31 |
| | | Expected Count | 26.3 | 4.7 | 31.0 |
| | | % within R-PSFs (lite) | 83.9% | 16.1% | 100.0% |
| | | Std. Residual | -.1 | .1 | |
| | Workload - Time pressure - Stress | Count | 41 | 17 | 58 |
| | | Expected Count | 49.2 | 8.8 | 58.0 |
| | | % within R-PSFs (lite) | 70.7% | 29.3% | 100.0% |
| | | Std. Residual | -1.2 | 2.8 | |
| Total | | Count | 1421 | 255 | 1676 |
| | | Expected Count | 1421.0 | 255.0 | 1676.0 |
| | | % within R-PSFs (lite) | 84.8% | 15.2% | 100.0% |

Table VIII-13

**R-PSFs (lite) * Shunting operation Crosstabulation**

| | | | Shunting operation | | |
|---|---|---|---|---|---|
| | | | No | Yes | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 233 | 51 | 284 |
| | | Expected Count | 245.9 | 38.1 | 284.0 |
| | | % within R-PSFs (lite) | 82.0% | 18.0% | 100.0% |
| | | Std. Residual | -.8 | 2.1 | |
| | System design - Visibility - HMI - RCM | Count | 172 | 15 | 187 |
| | | Expected Count | 161.9 | 25.1 | 187.0 |
| | | % within R-PSFs (lite) | 92.0% | 8.0% | 100.0% |
| | | Std. Residual | .8 | -2.0 | |
| | Fatigue - Shift pattern - Fit to work | Count | 104 | 5 | 109 |
| | | Expected Count | 94.4 | 14.6 | 109.0 |
| | | % within R-PSFs (lite) | 95.4% | 4.6% | 100.0% |
| | | Std. Residual | 1.0 | -2.5 | |
| | Communication - Teamwork | Count | 179 | 49 | 228 |
| | | Expected Count | 197.4 | 30.6 | 228.0 |
| | | % within R-PSFs (lite) | 78.5% | 21.5% | 100.0% |
| | | Std. Residual | -1.3 | 3.3 | |
| | Quality of procedures | Count | 114 | 15 | 129 |
| | | Expected Count | 111.7 | 17.3 | 129.0 |
| | | % within R-PSFs (lite) | 88.4% | 11.6% | 100.0% |
| | | Std. Residual | .2 | -.6 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 216 | 27 | 243 |
| | | Expected Count | 210.4 | 32.6 | 243.0 |
| | | % within R-PSFs (lite) | 88.9% | 11.1% | 100.0% |
| | | Std. Residual | .4 | -1.0 | |
| | Perception - Interpretation | Count | 107 | 12 | 119 |
| | | Expected Count | 103.0 | 16.0 | 119.0 |
| | | % within R-PSFs (lite) | 89.9% | 10.1% | 100.0% |
| | | Std. Residual | .4 | -1.0 | |
| | Training - Experience | Count | 118 | 14 | 132 |
| | | Expected Count | 114.3 | 17.7 | 132.0 |
| | | % within R-PSFs (lite) | 89.4% | 10.6% | 100.0% |
| | | Std. Residual | .3 | -.9 | |
| | Expectation - Familiarity - Routine | Count | 83 | 17 | 100 |
| | | Expected Count | 86.6 | 13.4 | 100.0 |
| | | % within R-PSFs (lite) | 83.0% | 17.0% | 100.0% |
| | | Std. Residual | -.4 | 1.0 | |
| | Quality of information | Count | 45 | 11 | 56 |
| | | Expected Count | 48.5 | 7.5 | 56.0 |
| | | % within R-PSFs (lite) | 80.4% | 19.6% | 100.0% |
| | | Std. Residual | -.5 | 1.3 | |
| | Supervision | Count | 30 | 1 | 31 |
| | | Expected Count | 26.8 | 4.2 | 31.0 |
| | | % within R-PSFs (lite) | 96.8% | 3.2% | 100.0% |
| | | Std. Residual | .6 | -1.5 | |
| | Workload - Time pressure - Stress | Count | 50 | 8 | 58 |
| | | Expected Count | 50.2 | 7.8 | 58.0 |
| | | % within R-PSFs (lite) | 86.2% | 13.8% | 100.0% |
| | | Std. Residual | .0 | .1 | |
| Total | | Count | 1451 | 225 | 1676 |
| | | Expected Count | 1451.0 | 225.0 | 1676.0 |
| | | % within R-PSFs (lite) | 86.6% | 13.4% | 100.0% |

Table VIII-14

**R-PSFs (lite) * Number of operators Crosstabulation**

| | | | Number of operators | | |
|---|---|---|---|---|---|
| | | | One | More than one | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 238 | 46 | 284 |
| | | Expected Count | 241.0 | 43.0 | 284.0 |
| | | % within R-PSFs (lite) | 83.8% | 16.2% | 100.0% |
| | | Std. Residual | -.2 | .5 | |
| | System design - Visibility - HMI - RCM | Count | 143 | 44 | 187 |
| | | Expected Count | 158.7 | 28.3 | 187.0 |
| | | % within R-PSFs (lite) | 76.5% | 23.5% | 100.0% |
| | | Std. Residual | -1.2 | 2.9 | |
| | Fatigue - Shift pattern - Fit to work | Count | 74 | 35 | 109 |
| | | Expected Count | 92.5 | 16.5 | 109.0 |
| | | % within R-PSFs (lite) | 67.9% | 32.1% | 100.0% |
| | | Std. Residual | -1.9 | 4.5 | |
| | Communication - Teamwork | Count | 197 | 31 | 228 |
| | | Expected Count | 193.4 | 34.6 | 228.0 |
| | | % within R-PSFs (lite) | 86.4% | 13.6% | 100.0% |
| | | Std. Residual | .3 | -.6 | |
| | Quality of procedures | Count | 109 | 20 | 129 |
| | | Expected Count | 109.4 | 19.6 | 129.0 |
| | | % within R-PSFs (lite) | 84.5% | 15.5% | 100.0% |
| | | Std. Residual | .0 | .1 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 219 | 24 | 243 |
| | | Expected Count | 206.2 | 36.8 | 243.0 |
| | | % within R-PSFs (lite) | 90.1% | 9.9% | 100.0% |
| | | Std. Residual | .9 | -2.1 | |
| | Perception - Interpretation | Count | 99 | 20 | 119 |
| | | Expected Count | 101.0 | 18.0 | 119.0 |
| | | % within R-PSFs (lite) | 83.2% | 16.8% | 100.0% |
| | | Std. Residual | -.2 | .5 | |
| | Training - Experience | Count | 123 | 9 | 132 |
| | | Expected Count | 112.0 | 20.0 | 132.0 |
| | | % within R-PSFs (lite) | 93.2% | 6.8% | 100.0% |
| | | Std. Residual | 1.0 | -2.5 | |
| | Expectation - Familiarity - Routine | Count | 91 | 9 | 100 |
| | | Expected Count | 84.8 | 15.2 | 100.0 |
| | | % within R-PSFs (lite) | 91.0% | 9.0% | 100.0% |
| | | Std. Residual | .7 | -1.6 | |
| | Quality of information | Count | 50 | 6 | 56 |
| | | Expected Count | 47.5 | 8.5 | 56.0 |
| | | % within R-PSFs (lite) | 89.3% | 10.7% | 100.0% |
| | | Std. Residual | .4 | -.9 | |
| | Supervision | Count | 26 | 5 | 31 |
| | | Expected Count | 26.3 | 4.7 | 31.0 |
| | | % within R-PSFs (lite) | 83.9% | 16.1% | 100.0% |
| | | Std. Residual | -.1 | .1 | |
| | Workload - Time pressure - Stress | Count | 53 | 5 | 58 |
| | | Expected Count | 49.2 | 8.8 | 58.0 |
| | | % within R-PSFs (lite) | 91.4% | 8.6% | 100.0% |
| | | Std. Residual | .5 | -1.3 | |
| Total | | Count | 1422 | 254 | 1676 |
| | | Expected Count | 1422.0 | 254.0 | 1676.0 |
| | | % within R-PSFs (lite) | 84.8% | 15.2% | 100.0% |

Table VIII-15

**R-PSFs (lite) * Railway protection systems features Crosstabulation**

| | | | Railway protection systems features | | |
|---|---|---|---|---|---|
| | | | AWS | TPWS | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 173 | 106 | 279 |
| | | Expected Count | 170.8 | 108.2 | 279.0 |
| | | % within R-PSFs (lite) | 62.0% | 38.0% | 100.0% |
| | | Std. Residual | .2 | -.2 | |
| | System design - Visibility - HMI - RCM | Count | 110 | 76 | 186 |
| | | Expected Count | 113.9 | 72.1 | 186.0 |
| | | % within R-PSFs (lite) | 59.1% | 40.9% | 100.0% |
| | | Std. Residual | -.4 | .5 | |
| | Fatigue - Shift pattern - Fit to work | Count | 81 | 25 | 106 |
| | | Expected Count | 64.9 | 41.1 | 106.0 |
| | | % within R-PSFs (lite) | 76.4% | 23.6% | 100.0% |
| | | Std. Residual | 2.0 | -2.5 | |
| | Communication - Teamwork | Count | 137 | 85 | 222 |
| | | Expected Count | 135.9 | 86.1 | 222.0 |
| | | % within R-PSFs (lite) | 61.7% | 38.3% | 100.0% |
| | | Std. Residual | .1 | -.1 | |
| | Quality of procedures | Count | 93 | 33 | 126 |
| | | Expected Count | 77.1 | 48.9 | 126.0 |
| | | % within R-PSFs (lite) | 73.8% | 26.2% | 100.0% |
| | | Std. Residual | 1.8 | -2.3 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 131 | 106 | 237 |
| | | Expected Count | 145.1 | 91.9 | 237.0 |
| | | % within R-PSFs (lite) | 55.3% | 44.7% | 100.0% |
| | | Std. Residual | -1.2 | 1.5 | |
| | Perception - Interpretation | Count | 54 | 65 | 119 |
| | | Expected Count | 72.9 | 46.1 | 119.0 |
| | | % within R-PSFs (lite) | 45.4% | 54.6% | 100.0% |
| | | Std. Residual | -2.2 | 2.8 | |
| | Training - Experience | Count | 83 | 46 | 129 |
| | | Expected Count | 79.0 | 50.0 | 129.0 |
| | | % within R-PSFs (lite) | 64.3% | 35.7% | 100.0% |
| | | Std. Residual | .5 | -.6 | |
| | Expectation - Familiarity - Routine | Count | 58 | 40 | 98 |
| | | Expected Count | 60.0 | 38.0 | 98.0 |
| | | % within R-PSFs (lite) | 59.2% | 40.8% | 100.0% |
| | | Std. Residual | -.3 | .3 | |
| | Quality of information | Count | 29 | 24 | 53 |
| | | Expected Count | 32.5 | 20.5 | 53.0 |
| | | % within R-PSFs (lite) | 54.7% | 45.3% | 100.0% |
| | | Std. Residual | -.6 | .8 | |
| | Supervision | Count | 26 | 4 | 30 |
| | | Expected Count | 18.4 | 11.6 | 30.0 |
| | | % within R-PSFs (lite) | 86.7% | 13.3% | 100.0% |
| | | Std. Residual | 1.8 | -2.2 | |
| | Workload - Time pressure - Stress | Count | 31 | 27 | 58 |
| | | Expected Count | 35.5 | 22.5 | 58.0 |
| | | % within R-PSFs (lite) | 53.4% | 46.6% | 100.0% |
| | | Std. Residual | -.8 | 1.0 | |
| Total | | Count | 1006 | 637 | 1643 |
| | | Expected Count | 1006.0 | 637.0 | 1643.0 |
| | | % within R-PSFs (lite) | 61.2% | 38.8% | 100.0% |

Table VIII-16

**R-PSFs (lite) * Railway protection systems features Crosstabulation**

| | | | Railway protection systems features | | |
|---|---|---|---|---|---|
| | | | AWS | TPWS - ATP | Total |
| R-PSFs (lite) | Safety culture - SMS | Count | 173 | 111 | 284 |
| | | Expected Count | 170.5 | 113.5 | 284.0 |
| | | % within R-PSFs (lite) | 60.9% | 39.1% | 100.0% |
| | | Std. Residual | .2 | -.2 | |
| | System design - Visibility - HMI - RCM | Count | 110 | 77 | 187 |
| | | Expected Count | 112.2 | 74.8 | 187.0 |
| | | % within R-PSFs (lite) | 58.8% | 41.2% | 100.0% |
| | | Std. Residual | -.2 | .3 | |
| | Fatigue - Shift pattern - Fit to work | Count | 81 | 28 | 109 |
| | | Expected Count | 65.4 | 43.6 | 109.0 |
| | | % within R-PSFs (lite) | 74.3% | 25.7% | 100.0% |
| | | Std. Residual | 1.9 | -2.4 | |
| | Communication - Teamwork | Count | 137 | 91 | 228 |
| | | Expected Count | 136.9 | 91.1 | 228.0 |
| | | % within R-PSFs (lite) | 60.1% | 39.9% | 100.0% |
| | | Std. Residual | .0 | .0 | |
| | Quality of procedures | Count | 93 | 36 | 129 |
| | | Expected Count | 77.4 | 51.6 | 129.0 |
| | | % within R-PSFs (lite) | 72.1% | 27.9% | 100.0% |
| | | Std. Residual | 1.8 | -2.2 | |
| | Distraction - Loss of concentration - Vigilance - Risk awareness | Count | 131 | 112 | 243 |
| | | Expected Count | 145.9 | 97.1 | 243.0 |
| | | % within R-PSFs (lite) | 53.9% | 46.1% | 100.0% |
| | | Std. Residual | -1.2 | 1.5 | |
| | Perception - Interpretation | Count | 54 | 65 | 119 |
| | | Expected Count | 71.4 | 47.6 | 119.0 |
| | | % within R-PSFs (lite) | 45.4% | 54.6% | 100.0% |
| | | Std. Residual | -2.1 | 2.5 | |
| | Training - Experience | Count | 83 | 49 | 132 |
| | | Expected Count | 79.2 | 52.8 | 132.0 |
| | | % within R-PSFs (lite) | 62.9% | 37.1% | 100.0% |
| | | Std. Residual | .4 | -.5 | |
| | Expectation - Familiarity - Routine | Count | 58 | 42 | 100 |
| | | Expected Count | 60.0 | 40.0 | 100.0 |
| | | % within R-PSFs (lite) | 58.0% | 42.0% | 100.0% |
| | | Std. Residual | -.3 | .3 | |
| | Quality of information | Count | 29 | 27 | 56 |
| | | Expected Count | 33.6 | 22.4 | 56.0 |
| | | % within R-PSFs (lite) | 51.8% | 48.2% | 100.0% |
| | | Std. Residual | -.8 | 1.0 | |
| | Supervision | Count | 26 | 5 | 31 |
| | | Expected Count | 18.6 | 12.4 | 31.0 |
| | | % within R-PSFs (lite) | 83.9% | 16.1% | 100.0% |
| | | Std. Residual | 1.7 | -2.1 | |
| | Workload - Time pressure - Stress | Count | 31 | 27 | 58 |
| | | Expected Count | 34.8 | 23.2 | 58.0 |
| | | % within R-PSFs (lite) | 53.4% | 46.6% | 100.0% |
| | | Std. Residual | -.6 | .8 | |
| Total | | Count | 1006 | 670 | 1676 |
| | | Expected Count | 1006.0 | 670.0 | 1676.0 |
| | | % within R-PSFs (lite) | 60.0% | 40.0% | 100.0% |

# Appendix IX Loglinear analysis tests

Table IX-1

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Responsibility**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 233 | 2817.215 | .000 | 3647.919 | .000 | 0 |
| | 2 | 214 | 599.853 | .000 | 633.412 | .000 | 2 |
| | 3 | 120 | 122.692 | .415 | 106.878 | .799 | 4 |
| K-way Effects[b] | 1 | 19 | 2217.362 | .000 | 3014.507 | .000 | 0 |
| | 2 | 94 | 477.161 | .000 | 526.534 | .000 | 0 |
| | 3 | 120 | 122.692 | .415 | 106.878 | .799 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-2

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Location**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 155 | 1659.762 | .000 | 1900.639 | .000 | 0 |
| | 2 | 138 | 322.441 | .000 | 317.821 | .000 | 2 |
| | 3 | 72 | 111.980 | .002 | 106.838 | .005 | 4 |
| K-way Effects[b] | 1 | 17 | 1337.321 | .000 | 1582.818 | .000 | 0 |
| | 2 | 66 | 210.461 | .000 | 210.983 | .000 | 0 |
| | 3 | 72 | 111.980 | .002 | 106.838 | .005 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-3

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Immediate causes**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 311 | 3151.870 | .000 | 5078.813 | .000 | 0 |
| | 2 | 290 | 563.878 | .000 | 548.214 | .000 | 2 |
| | 3 | 168 | 150.908 | .824 | 128.823 | .989 | 4 |
| K-way Effects[b] | 1 | 21 | 2587.992 | .000 | 4530.599 | .000 | 0 |
| | 2 | 122 | 412.970 | .000 | 419.391 | .000 | 0 |
| | 3 | 168 | 150.908 | .824 | 128.823 | .989 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-4

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Special Operational Conditions**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 77 | 2068.992 | .000 | 2367.527 | .000 | 0 |
| | 2 | 62 | 221.295 | .000 | 233.139 | .000 | 2 |
| | 3 | 24 | 23.660 | .481 | 20.872 | .646 | 3 |
| K-way Effects[b] | 1 | 15 | 1847.697 | .000 | 2134.389 | .000 | 0 |
| | 2 | 38 | 197.635 | .000 | 212.267 | .000 | 0 |
| | 3 | 24 | 23.660 | .481 | 20.872 | .646 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-5

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Shunting Operations**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 77 | 2186.910 | .000 | 2433.148 | .000 | 0 |
| | 2 | 62 | 231.836 | .000 | 240.008 | .000 | 2 |
| | 3 | 24 | 55.277 | .000 | 57.150 | .000 | 3 |
| K-way Effects[b] | 1 | 15 | 1955.074 | .000 | 2193.140 | .000 | 0 |
| | 2 | 38 | 176.559 | .000 | 182.858 | .000 | 0 |
| | 3 | 24 | 55.277 | .000 | 57.150 | .000 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-6

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Number of Operators**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 77 | 2242.080 | .000 | 2622.098 | .000 | 0 |
| | 2 | 62 | 390.943 | .000 | 430.507 | .000 | 2 |
| | 3 | 24 | 30.498 | .169 | 28.266 | .249 | 4 |
| K-way Effects[b] | 1 | 15 | 1851.138 | .000 | 2191.591 | .000 | 0 |
| | 2 | 38 | 360.445 | .000 | 402.241 | .000 | 0 |
| | 3 | 24 | 30.498 | .169 | 28.266 | .249 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-7

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Railway Protection Systems**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 77 | 1237.366 | .000 | 1179.191 | .000 | 0 |
| | 2 | 62 | 215.930 | .000 | 208.896 | .000 | 2 |
| | 3 | 24 | 57.821 | .000 | 55.975 | .000 | 3 |
| K-way Effects[b] | 1 | 15 | 1021.436 | .000 | 970.295 | .000 | 0 |
| | 2 | 38 | 158.109 | .000 | 152.920 | .000 | 0 |
| | 3 | 24 | 57.821 | .000 | 55.975 | .000 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-8

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Location - Shunting Operations**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 311 | 855.430 | .000 | 2050.520 | .000 | 0 |
| | 2 | 293 | 154.775 | 1.000 | 184.401 | 1.000 | 2 |
| | 3 | 210 | 48.115 | 1.000 | 40.678 | 1.000 | 4 |
| | 4 | 72 | .000 | 1.000 | .000 | 1.000 | 2 |
| K-way Effects[b] | 1 | 18 | 700.655 | .000 | 1866.119 | .000 | 0 |
| | 2 | 83 | 106.660 | .041 | 143.723 | .000 | 0 |
| | 3 | 138 | 48.115 | 1.000 | 40.678 | 1.000 | 0 |
| | 4 | 72 | .000 | 1.000 | .000 | 1.000 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-9

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Location - Railway Protection Systems**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 311 | 2087.613 | .000 | 2443.219 | .000 | 0 |
| | 2 | 293 | 718.807 | .000 | 737.172 | .000 | 2 |
| | 3 | 210 | 307.546 | .000 | 297.805 | .000 | 6 |
| | 4 | 72 | 82.306 | .191 | 75.130 | .377 | 5 |
| K-way Effects[b] | 1 | 18 | 1368.806 | .000 | 1706.047 | .000 | 0 |
| | 2 | 83 | 411.261 | .000 | 439.367 | .000 | 0 |
| | 3 | 138 | 225.240 | .000 | 222.675 | .000 | 0 |
| | 4 | 72 | 82.306 | .191 | 75.130 | .377 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

Table IX-10

**K-Way and Higher-Order Effects for Type of Events - R-PSFs lite - Shunting Operations - Railway Protection Systems**

| | | | Likelihood Ratio | | Pearson | | |
|---|---|---|---|---|---|---|---|
| | K | df | Chi-Square | Sig. | Chi-Square | Sig. | Number of Iterations |
| K-way and Higher Order Effects[a] | 1 | 155 | 2470.277 | .000 | 2891.254 | .000 | 0 |
| | 2 | 139 | 454.672 | .000 | 479.468 | .000 | 2 |
| | 3 | 86 | 184.385 | .000 | 190.459 | .000 | 4 |
| | 4 | 24 | 17.792 | .813 | 13.788 | .951 | 4 |
| K-way Effects[b] | 1 | 16 | 2015.605 | .000 | 2411.786 | .000 | 0 |
| | 2 | 53 | 270.287 | .000 | 289.009 | .000 | 0 |
| | 3 | 62 | 166.593 | .000 | 176.672 | .000 | 0 |
| | 4 | 24 | 17.792 | .813 | 13.788 | .951 | 0 |

a. Tests that k-way and higher order effects are zero.
b. Tests that k-way effects are zero.

# Appendix X Material to support SME's common understanding of R-PSFs

## The R-PSFs categories

### Dynamic Personal factors

Factors that characterise and affect individual's performance and are strongly related to the precise moment of the operation.

Distraction - Loss of concentration, Expectation, Stress, Perception, Interpretation, Stress, Fatigue, Vigilance, Situational awareness, Fit to work (health), Decision making.

### Personal factors

Factors that characterise and affect individual's performance **but are not** strongly related to the precise moment of the operation.

Training - competence, Familiarity, Experience, Motivation, Individual characteristics (e.g. age)

### Team factors

Factors that characterise and affect individual's performance as a member of a team which works together to achieve a common goal.

Communication, Teamwork, Team relations, Quality and Trust in Information.

### System factors

Factors that characterise the system and affect individual's performance through the interaction between the system and the operator.

System design (e.g. signal location), Human Machine Interface, Working Environment, Trust in equipment, Railway communication means.

### Organisational factors

Factors that are defined by or are under the control of the organisation.

Safety culture, Safety Management Systems, Quality of procedures, Leadership, Supervision, Shift pattern, Relations within organisation, Incentives for employees, Communication within organisation – feeling secure, Fit to work aspect

### Task factors

Factors that characterise the features of an executed task.

Workload, Monotony, Routine, Time pressure-time to respond, Task complexity, Task instructions.

### Environmental factors

Factors primarily related to weather conditions over the execution of a task.

Weather conditions, Visibility.

## The R-PSFs elements

### Training - Experience

Sufficient knowledge of the specific tasks to be undertaken and the risks which the work will entail. Knowledge and skills especially that gained in a particular profession.

Operator (involved in a task) trained, qualified and competent so that operates its business safely and effectively.

### Familiarity - Expectation - Routine

A good knowledge of something. A strong belief that something will happen or be the case.

The level of familiarity with a task, route, scenario, procedure or equipment. A known pattern of information, leading to a strong expectation about what information appeared. For example, dedicated train drivers as they execute the same route for several times per day.

### Distraction - Loss of concentration- Situational awareness - Vigilance

Anything that prevents operator from concentrating on what he/she is doing

Other people in/around the cab. Driver may be distracted by people/events or preoccupied with an issue of concern. Communication with other co-workers. Use of mobile phone. Desire to finish duty.

### Fatigue - Shift pattern - Fit to work

Extreme tiredness resulting from mental or physical exertion or illness or prolonged activity.

Level of fatigue due to factors such as lack of sleep, too much work, shift pattern, duration of breaks, staffing levels, personal life **but not** related to training or experience.

### Perception - Interpretation

The ability to see, hear, or become aware of something through the senses or the way in which something is regarded, understood, or interpreted.

Misread visual information or mishear auditory information. Misinterpret auditory (verbal and non-verbal) or visual information.

### Workload - Stress - Time pressure

The amount of work to be done by someone in a period of time. The state of mental or emotional strain resulting from adverse or demanding circumstances or a great worry caused by a difficult situation.

The levels of workload due to multi-tasking requirements, disruption/ faults, complex route, unexpected events etc. The available time to respond to an unexpected event or to correct an identified error.

### Supervision

Supervise and inspect work, maintain discipline in order to preserve the company's objectives.

Managing, developing, supporting staff, and mediating on their behalf. Gathering information about work activities and external conditions affecting the work; checking on the progress and quality of the work; evaluating the performance of individuals and the effectiveness of the organizational unit.

### System design - HMI

The process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements.

Whether the system was designed and constructed correctly or there are inconsistencies, e.g. signals installation, station design, rolling stock equipment, displays, usability, design and quality of equipment.

### Communication – Teamwork

The quality of communication between employees when no technical failure is identified and how people work together and coordinate to achieve safe performance.

Refers to the level of quick, explicit, clear communication, oral or hand signaling and to clear distinction of roles-responsibilities, collaboration and quality of feedback between crew members, level of trust and general social climate.

### Information quality & trust

Trust in person providing information and/or quality & availability of information given in procedure via verbal communication (or other method).

Refer to accurate, up to date, available, relevant and explicit information and/or the level of trust for the person that provides the information.

### Safety culture – SMS

Safety culture characterises the set of attitude, values and priorities placed on all aspects of safety by the organisation.

A safety management system defines roles and responsibilities, sets arrangements for safety mechanisms, involves workers in the process and ensures continuous improvement.

### Procedures quality & availability

The existence and use of formal operating procedures for the tasks under consideration. The official rules or directives made and maintained by the corresponding authorities.

Refers to the existence, clarity, availability, accuracy, and how easy the procedures are to be followed.

# Appendix XI The R-PSFs "weighting" and "ranking" questionnaire

**Human Performance Railway Operational Index Questionnaire**

Dear Madam / Sir,

This questionnaire is created for the purpose of obtaining information on the factors that influence the performance of train drivers during railway operations, such as the level of training, the given procedures or the level of fatigue.

I would, therefore, like to ask you to fill in the following questionnaire. It will only take few minutes of your time to fill it in, which will contribute to our effort to assess in a more comprehensive way the human factors aspect into railway safety. The information you provide will be used as a part of a PhD research project is currently being undertaken at Imperial College London.

The questionnaire comprises five parts. In the first part you are kindly asked to provide some generic personal information in a strictly confidential manner. In the second and third parts you are asked to assess both the categories and the factors that most significantly influence human performance for a railway operational scenario related to SPADs. In the fourth part, you are asked to rate how good or bad each of the factors is with respect to the ability of an operator to conduct a certain action. Finally, your feedback about the questionnaire is asked in the fifth part.

Details and instructions are provided in each part of the questionnaire.

Upon completion please enclose the questionnaire in the attached envelope. Therefore, anonymity for all responders and confidentiality of data are assured.

Thank you in advance for your time and effort.

Sincerely,
Miltos Kyriakidis

Research Student
LRET Transport Risk
Management Center
Centre for Transport Studies
Imperial College London
SW2 2AZ
e-mail:
m.kyriakidis@imperial.ac.uk
pwp: www.miltoskyriakidis.com

## SECTION 1. Introductory details

| | | | | | | |
|---|---|---|---|---|---|---|
| **1.1** | **Date** | | **1.2** | **What is your grade?** | _____ | |
| **1.3** | **Which is your job title** | | | | | |
| | | Train Driver | | Safety analyst | | |
| | | Signaller | | Human Factors expert | | |
| | | Controller | | Other (please indicate) | | |
| **1.4** | **Where is your main place of work? (dedicated to one line?)** | | | _____ | | |
| **1.5** | **What is your gender?** | | | | | |
| | | Male | | Female | | |
| **1.6** | **Please indicate your age group** | | | | | |
| | | Under 25 | | 25–34 | | 35–44 |
| | | 45–56 | | Over 56 | | |
| **1.7** | **How long have you been working in the railway industry?** | | | | | |
| | | Less than 1 year | | 1 to 5 years | | 6 to 10 years |
| | | 11 to 19 years | | 20 years or more | | |
| **1.8** | **How long have you been in your current posting?** | | | | | |
| | | Less than 1 year | | 1 to 5 years | | 6 to 10 years |
| | | 11 to 19 years | | 20 years or more | | |
| **1.9** | **Have you ever been involved to a SPAD before? (if "No", ignore question 1.10)** | | | | | |
| | Yes | | No | | | |
| **1.10** | **If you have been involved to a SPAD, can you recall what was the case** | | | | | |
| | | Fail to detect signal | | Wrong signal interpretation | | |
| | | Wrong signal observation | | Misread previous signal | | |
| | | Late signal observation | | Braking | | |
| | | Anticipation of signal clearance | | Other (please indicate) _____ | | |
| **1.11** | **Based on your experience a SPAD is more likely to happen** | | | | | |
| | | Inside a tunnel | | On open line | | |

**SECTION 2. R-PSFs Categories "weighting" assessment**

In this part you are asked to assess the categories of R-PSFs that affect Human Performance Integrity for the generic scenario described below.

**Scenario - Description for Train Drivers**

**Train driver fails to stop the train before passes signal at danger on open line (not tunnel), although no technical failures are reported. Good weather conditions. Good visibility. Day operation.**

To assess the influence of each R-PSF category, please keep in mind the following question:

"*Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how much more important is it with respect to the influence on the reference factor category?*"

For the assessment, you have to use the fundamental scale given below:

| The fundamental scale | | |
|---|---|---|
| **Intensity of importance** | **Definition** | **Explanation** |
| **1** | **Equal importance** | Two factors contribute equally to the objective |
| 2 | Weak | Compromise between values 1 & 3 |
| **3** | **Moderate importance** | Experience and judgment slightly favour one factor over another |
| 4 | Moderate plus | Compromise between values 3 & 5 |
| **5** | **Strong importance** | Experience and judgment strongly favour one factor over another |
| 6 | Strong plus | Compromise between values 5 & 7 |
| **7** | **Very strong demonstrated importance** | A factor is favoured very strongly over another; its dominance demonstrated in practice |
| 8 | Very very strong | Compromise between values 7 & 9 |
| **9** | **Extreme importance** | The evidence favouring one factor over another is of highest possible order of affirmation |

Table 2-1: *Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the* ***personal*** *factors?*

| Personal | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Personal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Task | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

Table 2-2: *Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the **dynamic personal** factors?*

| Dynamic Personal | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Dynamic Personal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Task | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

Table 2-3: *Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the **task** factors?*

| Task | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Task |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

Table 2-4: *Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the **team** factors?*

| Team | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Team |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| System | 9 | 8 | 7 | 6 | 5 | 4 | **3** | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Task | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

Table 2-5: *Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the* ***organisational*** *factors?*

| Organisational | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Organisational |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | **9** | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | **5** | 6 | 7 | 8 | 9 | System |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | **6** | 7 | 8 | 9 | Task |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | **8** | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | **4** | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | System |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | **4** | 5 | 6 | 7 | 8 | 9 | Team |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| System | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Task | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

Table 2-6: *Considering the influence of each factor category on the performance of the operator, of the two factor categories in each row which one is considered more important and how more important is it with respect to the influence on the* ***system*** *factors?*

| System | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | System |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Organisational |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Dynamic Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Personal |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Organisational | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Task |
| Personal | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |
| Task | 9 | 8 | 7 | 6 | 5 | 4 | 3 | **2** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Team |

**SECTION 3. R-PSFs Elements "weighting" assessment**

In this part you are asked to assess the R-PSFs elements that affect Human Performance Integrity for the generic scenario described below.

**Scenario - Description for Train Drivers**

**Train driver fails to stop the train before passes signal at danger on open line (not tunnel), although no technical failures are reported. Good weather conditions. Good visibility. Day operation.**

To assess the influence of each R-PSF, please keep in mind the following question:

"*Considering the influence of each factor on the performance of the operator, of the two factor elements in each row which one is considered more important and how much more important is it with respect to the influence on the reference factor?*"

For the assessment, you have to use the fundamental scale given below:

| The fundamental scale | | |
|---|---|---|
| **Intensity of importance** | **Definition** | **Explanation** |
| **1** | **Equal importance** | Two factors contribute equally to the objective |
| 2 | Weak | Compromise between values 1 & 3 |
| **3** | **Moderate importance** | Experience and judgment slightly favour one factor over another |
| 4 | Moderate plus | Compromise between values 3 & 5 |
| **5** | **Strong importance** | Experience and judgment strongly favour one factor over another |
| 6 | Strong plus | Compromise between values 5 & 7 |
| **7** | **Very strong demonstrated importance** | A factor is favoured very strongly over another; its dominance demonstrated in practice |
| 8 | Very very strong | Compromise between values 7 & 9 |
| **9** | **Extreme importance** | The evidence favouring one factor over another is of highest possible order of affirmation |

Table 3-1: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it, with respect to the influence on* ***training****?*

| Training | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Training |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Procedures | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | (2) | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | (8) | 9 | Supervision |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-2: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***familiarity****?*

| Familiarity | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Familiarity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | (4) | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | (8) | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | (5) | 6 | 7 | 8 | 9 | Information |

Table 3-3: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***workload****?*

| Workload | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Workload |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | (7) | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | (3) | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | (3) | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-4: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***distraction****?*

| Distractrion | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Distractrion |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Fatigue | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | (7) | 8 | 9 | Training |
| Communication | 9 | 8 | 7 | (6) | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-5: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***fatigue****?*

| Fatigue | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Fatigue |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-6: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***perception****?*

| Perception | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Perception |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | (6) | 7 | 8 | 9 | Training |
| Communication | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-7: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***communication****?*

| Communication | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Communication |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | (7) | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | (6) | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |

Table 3-8: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on the* ***quality and trust in information****?*

| Information | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Information |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |

Table 3-9: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***safety culture****?*

| Safety culture | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Safety culture |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-10: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on the* ***quality of procedures****?*

| Procedures | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Procedures |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Safety culture | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-11: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***supervision****?*

| Supervision | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | Supervision |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | 4 | 3 | (2) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

Table 3-12: *Considering the influence of each factor on the performance of the operator, of the two factors in each row which one is considered more important and how more important is it with respect to the influence on* ***system design****?*

| System design | Extreme | | Very strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very strong | | Extreme | System design |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Fatigue |
| Distraction | 9 | 8 | 7 | 6 | 5 | 4 | (3) | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Fatigue | 9 | 8 | 7 | 6 | 5 | 4 | (3) | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Perception |
| Procedures | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Safety culture |
| Procedures | 9 | 8 | 7 | 6 | 5 | (4) | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Safety culture | 9 | 8 | 7 | 6 | (5) | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Supervision |
| Familiarity | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Training |
| Communication | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | (1) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Information |

## SECTION 4. R-PSFs Elements "rating"

In this part, you are asked to rate the condition (optimal) of each R-PSF in the specific task context under consideration.

To rate the influence of each R-PSF on each action, please keep in mind the following question:

"*Considering that each R-PSF affects the ability of the operator to conduct a certain action, please rate how good or bad do you think that each R-PSF is with respect to the action required*".

Please use a 0 to 100 scale (increment of 10) to rate each R-PSF, where 0 shows the worst and 100 the best case. For example, "*how would you rate "training" with respect to not signal detection*"?

| | **Rating of Railway Performance Shaping Factors** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Scenario - Errors** | **R-PSFs** | | | | | | | | | | | |
| **SPAD**<br>Train driver, open line (not tunnel), day operation, good weather conditions, good visibility | **Training** (incl. experience) | **Familiarity** (incl. expectation and routine) | **Distraction** (incl. concentration, vigilance, situational | **Fatigue** (incl. shift pattern and fitness to work) | **Perception** (incl. interpretation) | **Workload** (incl. stress and time pressure) | **Communication** (incl. teamwork) | **Quality and Trust in Information** | **Safety culture** (incl. SMS) | **Quality of Procedures** | **Supervision** | **System design** (incl. HMI) |
| **Job at time of SPAD**<br>Train driver fails to stop the train before passes signal at danger | 0 = no training,<br>50= some training,<br>100= very good training | 0 = no familiarity<br>50 = some level of familiarity<br>100 = high levels of familiarity | 0 = very distracted<br>50 = some distraction<br>100 = no distraction | 0 = too fatigued<br>50 = some level of fatigue<br>100 = no fatigued | 0 = no perception<br>50 = some perception<br>100 = high perception levels | 0 = high workload<br>50 = some level of workload<br>100 = no workload | 0 = poor quality<br>50 = average quality<br>100 = very good quality | 0 = poor quality<br>50 = average quality<br>100 = high level quality | 0= poor safety culture<br>50 = some safety culture<br>100= very good safety culture | 0 = poor quality<br>50 = average quality<br>100 = very good quality | 0 = no supervision distracted<br>50 = some supervision<br>100 = very good supervision | 0 = poor system design<br>50 = average quality<br>100 = very good system design |
| 1. **Signal detection** | 80 | 80 | 10 | 30 | 20 | 60 | 70 | 70 | 70 | 80 | 80 | 50 |
| 2. **Signal interpretation** | 80 | 60 | 20 | 40 | 10 | 60 | 70 | 70 | 70 | 80 | 80 | 50 |
| 3. **Action executed** | 70 | 40 | 10 | 30 | 20 | 50 | 60 | 40 | 70 | 50 | 50 | 50 |

Finally, please indicate in a 0 (min) to 100 (max) scale how much your performance would change under the following conditions with respect to the described scenario? Please choose only one of the three options.

| Conditions | Performance | | |
|---|---|---|---|
| | Improved | Reduced | No influence |
| **Tunnel operation (compare to over ground)** | | | 0 |
| **Intermediate weather conditions (e.g. light rain, fog)** | | | 0 |
| **Adverse weather conditions** | | | 0 |
| **Degraded operation** | | | 0 |
| **Night operation** | | | 0 |

**SECTION 5. Some feedback about the Questionnaire**

We would be grateful for your feedback.

| | | |
|---|---|---|
| **4.1** | About how long has it taken you to complete this questionnaire?_____minutes | |
| **4.2** | Are there any important issues relating to your work or work environment that we have not asked or discussed about? | |
| | Yes | No |
| | If Yes, please state what these issues are:<br>_____<br>_____ | |
| **4.3** | If you have any further comments (both positive and negative) about the questionnaire would be greatly appreciated | |
| | _____<br>_____ | |

# Appendix XII The ratings of the R-PSFs per human error and classification group

**Human error "signal not detected"**

Table XII-1 R-PSFs mean ratings per TOC

| | TOCs | | |
|---|---|---|---|
| | **TOC1** | **TOC2** | **TOC3** |
| Distraction | 37 | 25 | 31 |
| Fatigue | 36 | 30 | 35 |
| Perception | 41 | 28 | 33 |
| Procedures | 51 | 44 | 45 |
| Safety Culture | 64 | 53 | 49 |
| Supervision | 46 | 35 | 33 |
| Familiarity | 32 | 65 | 42 |
| Training | 65 | 57 | 65 |
| System Design | 47 | 36 | 53 |
| Workload | 46 | 49 | 27 |
| Communication | 51 | 45 | 45 |
| Information | 49 | 46 | 53 |

Table XII-2 R-PSFs mean ratings per operational experience (train drivers only)

| | Operational experience | | | | |
|---|---|---|---|---|---|
| | < 1 year | 1 to 4 years | 5 to 10 years | 11 to 19 years | 20 years or more |
| Distraction | 30 | . | 37 | 19 | 38 |
| Fatigue | 36 | . | 30 | 24 | 43 |
| Perception | 27 | . | 39 | 36 | 36 |
| Procedures | 46 | . | 63 | 37 | 47 |
| Safety Culture | 70 | . | 64 | 47 | 51 |
| Supervision | 43 | . | 47 | 32 | 37 |
| Familiarity | 17 | . | 33 | 53 | 63 |
| Training | 66 | . | 61 | 70 | 56 |
| System Design | 47 | . | 47 | 43 | 45 |
| Workload | 44 | . | 36 | 38 | 47 |
| Communication | 33 | . | 66 | 43 | 49 |
| Information | 31 | . | 70 | 53 | 45 |

Table XII-3 R-PSFs mean ratings per age group (train drivers only)

| | Age group | | | |
|---|---|---|---|---|
| | 24 to 34 | 35 to 44 | 45 to 56 | Over 56 |
| Distraction | 26 | 22 | 39 | 33 |
| Fatigue | 30 | 19 | 46 | 30 |
| Perception | 24 | 22 | 46 | 38 |
| Procedures | 40 | 54 | 49 | 35 |
| Safety Culture | 63 | 58 | 58 | 35 |
| Supervision | 29 | 46 | 43 | 25 |
| Familiarity | 16 | 36 | 64 | 48 |
| Training | 60 | 71 | 61 | 53 |
| System Design | 37 | 56 | 47 | 33 |
| Workload | 36 | 31 | 51 | 43 |
| Communication | 27 | 57 | 56 | 33 |
| Information | 33 | 67 | 53 | 25 |

Table XII-4 R-PSFs mean ratings per gender and SPAD involvement (train drivers only)

| | Gender | | SPAD involvement | |
|---|---|---|---|---|
| | Male | Female | No | Yes |
| Distraction | 31 | 40 | 34 | 25 |
| Fatigue | 33 | 40 | 36 | 30 |
| Perception | 34 | 40 | 36 | 32 |
| Procedures | 45 | 70 | 48 | 46 |
| Safety Culture | 55 | 67 | 56 | 57 |
| Supervision | 38 | 47 | 44 | 27 |
| Familiarity | 45 | 43 | 40 | 59 |
| Training | 63 | 57 | 67 | 52 |
| System Design | 47 | 33 | 48 | 38 |
| Workload | 42 | 37 | 38 | 51 |
| Communication | 48 | 43 | 48 | 48 |
| Information | 49 | 57 | 51 | 46 |

**Human error "signal wrongly interpreted"**

Table XII-5 R-PSFs mean ratings per TOC

| | TOCs | | |
|---|---|---|---|
| | TOC1 | TOC2 | TOC3 |
| Distraction | 39 | 33 | 36 |
| Fatigue | 48 | 33 | 44 |
| Perception | 54 | 35 | 46 |
| Procedures | 48 | 47 | 38 |
| Safety Culture | 59 | 45 | 44 |
| Supervision | 49 | 37 | 28 |
| Familiarity | 59 | 43 | 49 |
| Training | 69 | 49 | 55 |
| System Design | 52 | 36 | 39 |
| Workload | 41 | 49 | 40 |
| Communication | 51 | 46 | 44 |
| Information | 46 | 50 | 48 |

Table XII-6 R-PSFs mean ratings per operational experience (train drivers only)

| | Operational experience | | | | |
|---|---|---|---|---|---|
| | **< 1 year** | **1 to 4 years** | **5 to 10 years** | **11 to 19 years** | **20 years or more** |
| Distraction | 31 | . | 37 | 32 | 41 |
| Fatigue | 56 | . | 31 | 33 | 47 |
| Perception | 51 | . | 51 | 44 | 39 |
| Procedures | 39 | . | 60 | 38 | 45 |
| Safety Culture | 61 | . | 59 | 40 | 47 |
| Supervision | 44 | . | 50 | 32 | 36 |
| Familiarity | 54 | . | 44 | 57 | 49 |
| Training | 77 | . | 63 | 53 | 50 |
| System Design | 49 | . | 44 | 39 | 43 |
| Workload | 37 | . | 39 | 47 | 48 |
| Communication | 34 | . | 66 | 42 | 48 |
| Information | 37 | . | 51 | 56 | 47 |

Table XII-7 R-PSFs mean ratings per age group (train drivers only)

| | Age group | | | |
|---|---|---|---|---|
| | **24 to 34** | **35 to 44** | **45 to 56** | **Over 56** |
| Distraction | 21 | 29 | 45 | 43 |
| Fatigue | 47 | 27 | 48 | 45 |
| Perception | 50 | 46 | 47 | 33 |
| Procedures | 27 | 59 | 48 | 33 |
| Safety Culture | 51 | 53 | 53 | 30 |
| Supervision | 24 | 51 | 43 | 25 |
| Familiarity | 47 | 53 | 53 | 48 |
| Training | 76 | 66 | 49 | 50 |
| System Design | 39 | 42 | 47 | 40 |
| Workload | 36 | 31 | 53 | 50 |
| Communication | 34 | 53 | 55 | 30 |
| Information | 39 | 53 | 54 | 30 |

Table XII-8 R-PSFs mean ratings per gender and SPAD involvement (train drivers only)

| | Gender | | SPAD involvement | |
|---|---|---|---|---|
| | Male | Female | No | Yes |
| Distraction | 37 | 27 | 38 | 32 |
| Fatigue | 42 | 40 | 45 | 34 |
| Perception | 46 | 37 | 48 | 39 |
| Procedures | 44 | 50 | 44 | 46 |
| Safety Culture | 51 | 43 | 52 | 47 |
| Supervision | 40 | 37 | 44 | 29 |
| Familiarity | 52 | 40 | 51 | 51 |
| Training | 59 | 57 | 62 | 51 |
| System Design | 46 | 13 | 44 | 41 |
| Workload | 44 | 37 | 40 | 51 |
| Communication | 48 | 40 | 48 | 46 |
| Information | 49 | 40 | 47 | 51 |

**Human error "correct action not taken"**

Table XII-9 R-PSFs mean ratings per TOC

| | TOCs | | |
|---|---|---|---|
| | TOC1 | TOC2 | TOC3 |
| Distraction | 32 | 22 | 27 |
| Fatigue | 32 | 30 | 32 |
| Perception | 47 | 34 | 38 |
| Procedures | 38 | 37 | 27 |
| Safety Culture | 59 | 48 | 38 |
| Supervision | 38 | 40 | 22 |
| Familiarity | 58 | 37 | 50 |
| Training | 50 | 44 | 38 |
| System Design | 51 | 37 | 45 |
| Workload | 42 | 40 | 26 |
| Communication | 45 | 40 | 39 |
| Information | 42 | 47 | 45 |

Table XII-10 R-PSFs mean ratings per operational experience (train drivers only)

| | Operational experience | | | | |
|---|---|---|---|---|---|
| | < 1 year | 1 to 4 years | 5 to 10 years | 11 to 19 years | 20 years or more |
| Distraction | 30 | . | 24 | 21 | 33 |
| Fatigue | 36 | . | 19 | 29 | 38 |
| Perception | 41 | . | 34 | 43 | 41 |
| Procedures | 21 | . | 51 | 30 | 36 |
| Safety Culture | 54 | . | 57 | 40 | 50 |
| Supervision | 24 | . | 47 | 30 | 35 |
| Familiarity | 61 | . | 36 | 58 | 43 |
| Training | 53 | . | 46 | 38 | 44 |
| System Design | 43 | . | 49 | 47 | 43 |
| Workload | 43 | . | 29 | 41 | 35 |
| Communication | 27 | . | 60 | 40 | 41 |
| Information | 30 | . | 47 | 56 | 43 |

Table XII-11 R-PSFs ratings per age group (train drivers only)

| | Age group | | | |
|---|---|---|---|---|
| | 24 to 34 | 35 to 44 | 45 to 56 | Over 56 |
| Distraction | 19 | 11 | 39 | 35 |
| Fatigue | 24 | 17 | 41 | 43 |
| Perception | 31 | 33 | 47 | 48 |
| Procedures | 19 | 43 | 39 | 28 |
| Safety Culture | 51 | 46 | 55 | 35 |
| Supervision | 19 | 39 | 41 | 23 |
| Familiarity | 51 | 53 | 47 | 45 |
| Training | 51 | 44 | 43 | 38 |
| System Design | 34 | 50 | 49 | 35 |
| Workload | 33 | 29 | 43 | 38 |
| Communication | 24 | 50 | 50 | 23 |
| Information | 33 | 51 | 51 | 28 |

Table XII-12 R-PSFs mean ratings per gender and SPAD involvement

| | Gender | | SPAD involvement | |
|---|---|---|---|---|
| | Male | Female | No | Yes |
| Distraction | 28 | 27 | 31 | 19 |
| Fatigue | 32 | 23 | 34 | 25 |
| Perception | 42 | 23 | 42 | 37 |
| Procedures | 34 | 40 | 32 | 42 |
| Safety Culture | 51 | 40 | 48 | 54 |
| Supervision | 35 | 23 | 34 | 35 |
| Familiarity | 50 | 40 | 52 | 41 |
| Training | 46 | 30 | 45 | 43 |
| System Design | 48 | 10 | 46 | 42 |
| Workload | 39 | 17 | 34 | 43 |
| Communication | 43 | 30 | 42 | 42 |
| Information | 46 | 33 | 42 | 51 |