

1     **Application of Human Factors Analysis and Classification System (HFACS) to**  
2                     **UK rail safety of the line incidents**

3                     Ruth Madigan <sup>a</sup>, David Golightly <sup>b</sup>, & Richard Madders <sup>c</sup>

4  
5     <sup>a</sup> Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, United Kingdom

6     <sup>b</sup> Human Factors Research Group, Innovative Technology Research Centre, Department of  
7     Mechanical, Materials and Manufacturing Engineering, University of Nottingham, University Park,  
8     Nottingham, NG7 2RD.

9     <sup>c</sup> Arcadia Alive Ltd., 8 The Quadrant, 99 Parkway Avenue, Sheffield, S9 4WG

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21

## 22 **Abstract**

23 Minor safety incidents on the railways cause disruption, and may be indicators of more serious safety  
24 risks. The following paper aimed to gain an understanding of the relationship between active and  
25 latent factors, and particular causal paths for these types of incidents by using the Human Factors  
26 Analysis and Classification System (HFACS) to examine rail industry incident reports investigating such  
27 events. 78 reports across 5 types of incident were reviewed by two authors and cross-referenced for  
28 interrater reliability using the index of concordance. The results indicate that the reports were strongly  
29 focused on active failures, particularly those associated with work-related distraction and  
30 environmental factors. Few latent factors were presented in the reports. Different causal pathways  
31 emerged for memory failures for events such a failure to call at stations, and attentional failures which  
32 were more often associated with signals passed at danger. The study highlights a need for the rail  
33 industry to look more closely at latent factors at the supervisory and organisational levels when  
34 investigating minor safety of the line incidents. The results also strongly suggest the importance of a  
35 new factor – operational environment – that captures unexpected and non-routine operating  
36 conditions which have a risk of distracting the driver. Finally, the study is further demonstration of the  
37 utility of HFACS to the rail industry, and of the usefulness of the index of concordance measure of  
38 interrater reliability.

39 **Keywords:** HFACS, System Analysis, Rail, Accident Investigation,

## 40 **1. Introduction**

41 In the period from 2001 to 2014 there were 803 fatalities (excluding suicides) and 5794 major  
42 injuries on the UK rail network (Department for Transport, 2014). Although, the rail industry has an  
43 excellent safety record in comparison to other forms of transport (Department for Transport, 2014),  
44 the Office of Road & Rail has put forward a safety vision for zero workforce and industry-caused  
45 passenger fatalities, and an ever-decreasing overall safety risk (ORR, 2014). If we are to move towards

46 a realisation of this vision, it is important to gain a detailed understanding of all of the factors which  
47 contribute to accidents and incidents in order to put appropriate controls in place.

48         Recent analyses have argued that human error was a causal factor in the occurrence of many  
49 serious and fatal rail accidents, both in the UK (French & Cope, 2012) and across Europe (Kyriakidis,  
50 Pak, & Majumdar, 2015). On top of these more serious incidents, there are many hundreds of minor  
51 incidents within the UK rail industry, many of which are also attributed to driver error. These include  
52 speed exceedances and signals passed at danger (SPADs) that did not lead to any accidents, along with  
53 trains that stop short or overshoot their platform, or fail to call altogether. These types of incident are  
54 extremely costly for organisations due to fines and infrastructure costs, along with disruption leading  
55 to negative public opinion. The most recent National Rail Passenger Survey showed that  
56 punctuality/reliability was the factor with the biggest impact on overall customer satisfaction, and  
57 how a train company dealt with delays had the biggest impact on overall dissatisfaction (NRPS, 2016).  
58 Additional costs arise as these incidents often require a driver to be removed from duty for an  
59 investigation and possibly retraining. Furthermore, the concern is that a minor event is an indicator of  
60 the risk of a more serious incident in the future (Reason, 1997; Hollnagel, 2014).

61         The opportunity for minor safety of the line events to occur is huge. For example, the number  
62 of approaches to red signals annually in the UK may be in the region of 7.5m (Gibson, Mills, Basacik,  
63 & Harrison, 2015). Few of these result in actual SPADs, and error probability for SPADs or events such  
64 as wrong side door openings (Basacik and Gibson, 2015) suggests error rates may be approaching the  
65 limits of performance. Therefore, careful analysis of events is required if new levels of safety are to be  
66 achieved, and there is a need for rail companies to understand what causes these events, so that  
67 potential courses of remedial action can be identified including training, technical or procedural  
68 change.

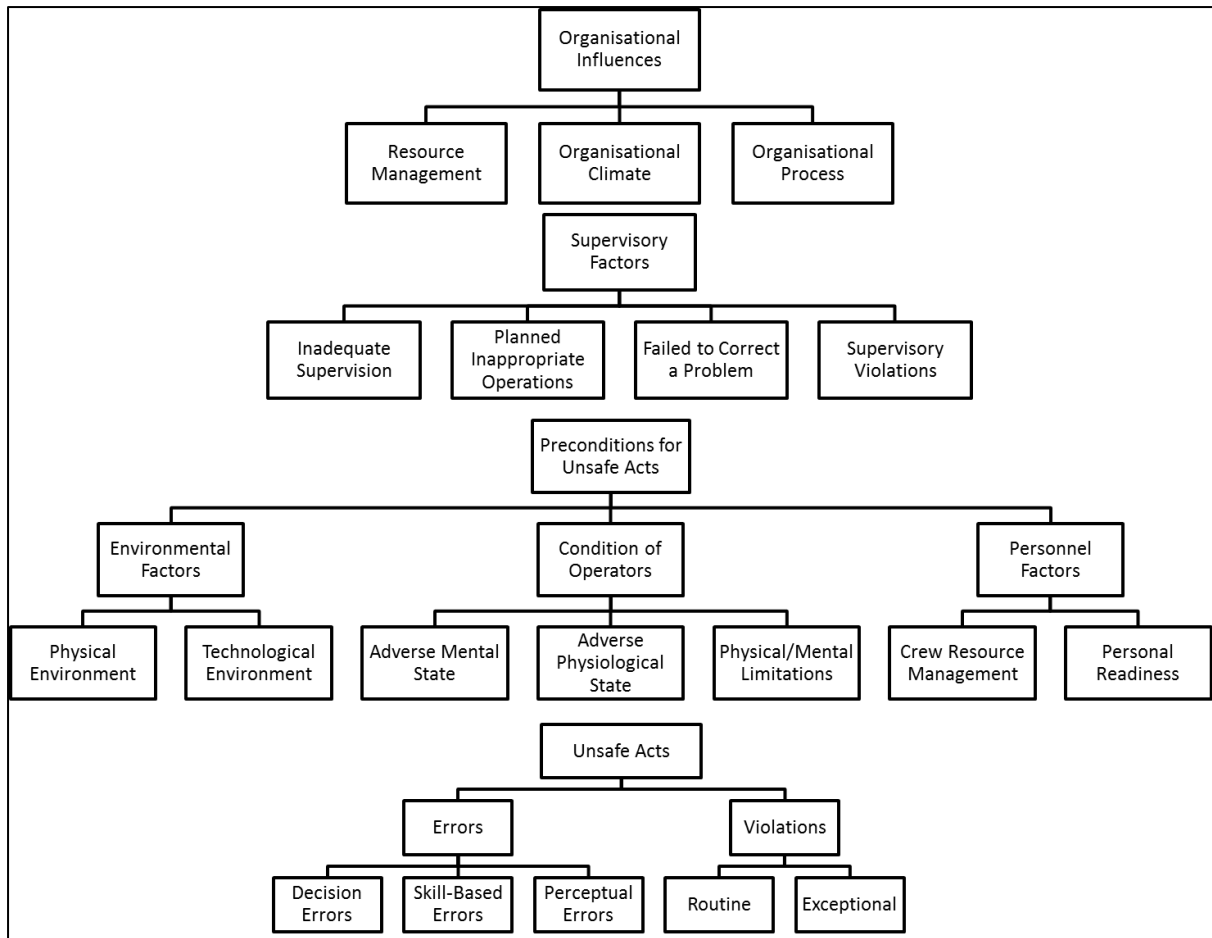
69         Contemporary human factors approaches to system safety have been used to provide greater  
70 insights into the causes of accidents in many safety-critical domains (Lenné, Salmon, Liu, & Trotter,

71 2012). Much of this work has been based on Reason's (1990) Generic Error Modelling System (GEMS),  
72 which defines two broad categories of error: active and latent failures. Active errors are associated  
73 with the front-line operators of a system, and their effects usually become evident almost  
74 immediately. Latent (or hidden) errors refer to the errors of designers or managers, and their adverse  
75 consequences may lie dormant within the system for a long time, only becoming evident when they  
76 combine with other factors to breach the system's defences. Reason (1990) noted that latent errors  
77 may pose the greatest risk to system safety because unless they are identified they remain in a system  
78 despite attempts to resolve an issue through rectifying the immediate performance issue (e.g. through  
79 non-systemic equipment fixes or training). Thus, one of the most important aspects of Reason's model  
80 is the argument that human error is a consequence, not a cause, of latent failures; and that "it is only  
81 by understanding the context that provoked an error can we hope to limit its reoccurrence" (Reason,  
82 1997, p.126). As a result of the issues outlined above, there is currently a strong emphasis on tackling  
83 human factors within the rail industry (e.g. Atkins, 2003; FRA, 2007; Lawton & Ward, 2005; RSSB,  
84 2009), and as part of this process it is vital that both the active and latent failures which contribute to  
85 railway incidents are understood.

## 86 **1.1 Human Factors Analysis & Classification System**

87 A number of studies have used different frameworks to look at the factors contributing to  
88 specific types of railway incident i.e. SPADs (e.g. Edkins & Pollock, 1997; Lawton & Ward, 2005; Rjabovs  
89 and Palacin, 2016), and specific types of error e.g. communication errors (Murphy, 2001). Read, Lenné,  
90 and Moss (2012) used the Contributing Factors Framework to investigate the associations between  
91 factors involved in Australian rail accidents and found that task demand factors (e.g. high workload,  
92 distraction) were significantly associated with skill-based errors; knowledge and training deficiencies  
93 significantly associated with mistakes; and violations significantly linked to social environmental  
94 factors. Currently, the UK rail sector is working towards a database of trends and themes in human  
95 performance and incident underlying causes for a sample of high risk Great British (GB) rail incidents.

96 This database uses the Incident Factor Classification System (IFCS) of 10 factors that may shape human  
97 performance in rail incidents (Gibson et al., 2015). However, one of the most common frameworks for  
98 analysis, based on Reason's (1990) model, is the Human Factors Analysis and Classification System  
99 (HFACS; Wiegmann & Shappell, 2003). HFACS describes four levels of failure based on Reason's Swiss  
100 Cheese Model (Reason, Hollnagel, & Paries, 2006): unsafe acts, preconditions for unsafe acts, unsafe  
101 supervision, and organisational influences (see Figure 1). Critically, this model specifies that in order  
102 for an incident to occur, failures in defences at all levels of the system must line up, thus highlighting  
103 the importance of identifying the factors which contribute at each level. The unsafe acts level focuses  
104 on identifying any errors or violations made by front line workers that led to an accident or incident  
105 occurring. Within the error category there are three subcategories of skill-based error, decision error,  
106 and perceptual error. Decision errors can be further broken down into rule-based and choice-based  
107 decisions, and skill-based errors can be broken down to attentional and memory failures. Within the  
108 violations category there are two subcategories of routine and exceptional violations.



109

110 **Figure 1: The HFACS framework (Wiegmann & Shappell, 2003)**

111 The second level of the HFACS framework is “preconditions for unsafe acts”. These refer to  
 112 the immediate underlying conditions that contribute to the occurrence of unsafe acts. This level  
 113 comprises three categories: condition of operators, environmental factors, and personnel factors.  
 114 Each of these categories has a number of subcategories as shown in Figure 1. The third level within  
 115 HFACS is “unsafe supervision”. This considers the situations where supervision was either lacking or  
 116 unsuitable and has four categories of inadequate supervision, planned inappropriate operations,  
 117 failure to correct a problem, and supervisory violations. The fourth and final level within many  
 118 applications of HFACS models is organisational influences. This level looks at the failures occurring at  
 119 the higher managerial levels of the organisation which contributed to an accident, focusing on the  
 120 subcategories of resource management, organisational climate and organisational process.

121 Typically, HFACS is used as a retrospective tool for analysing accident and incident reports,  
122 and the different failures which contributed to an accident at all four levels are identified. Although  
123 originally designed to classify aviation accidents (Wiegmann & Shappell, 2001; 2003), HFACS has now  
124 been applied successfully in numerous safety critical industries including maritime (Celik & Cebi,  
125 2009), mining (Lenné et al., 2012; Patterson & Shappell, 2010), medicine (ElBardissi, Wiegmann,  
126 Dearani, Daly, & Sundt, 2007) and rail (Baysari, McIntosh, & Wilson, 2008; Reinach & Viale, 2006), with  
127 researchers making various adaptations to the model to make it more suitable in different contexts.  
128 One criticism of HFACS has been its failure to consider contributory factors outside of the organisation  
129 involved, such as government policy, or local authority oversights (Salmon, Cornelissen, & Trotter,  
130 2012). For that reason, some versions have gone beyond the organisational level to include 'external  
131 influences' which take account of issues such as legislation gaps, administration oversights, and design  
132 flaws (e.g. Chen, Wall, Davies, Yang, Wang, & Chou, 2013; Reinach & Viale, 2006).

133 Overall, the results of previous studies provide strong support for the use of HFACS as a tool  
134 for understanding incidents in the rail industry. However, only two published studies have applied  
135 HFACS in this context. Reinach and Viale (2006) used an adapted version called HFACS-RR to examine  
136 six railyard switching incidents in the US and identified 36 probable contributing factors for these  
137 incidents. Baysari et al. (2008) investigated 40 publicly available railway incident and accident reports  
138 in Australia and identified 330 contributing factors. More than half of the incidents identified resulted  
139 from an equipment failure. In the remaining cases, skill-based errors (HFACS Level 1), adverse mental  
140 state (Level 2), and equipment/facility resources (Level 4) emerged as the most common contributory  
141 factors.

142 Both Baysari et al. (2008) and Read et al.'s (2012) studies focus on external inquiries into major  
143 accidents by relevant transport bodies (e.g. Australian Safety Transport Bureau), while the Reinach  
144 and Viale (2006) study focuses solely on switching yard incidents. However, to date, no published  
145 study has focused on the hundreds of minor incidents linked to train drivers every year, such as signals

146 passed at danger or failure to call at stations. As previously noted, these incidents can have extremely  
147 damaging consequences in terms of both infrastructure costs and negative public opinion. In addition,  
148 the causal pattern of these incidents is often similar to that of more serious incidents (Wright & Van  
149 der Schaaf, 2004). Although human error is often identified as a causal factor within these incidents,  
150 there has been little effort to gain a systematic understanding of the latent factors which contribute,  
151 and whether or not these differ depending on the type of incident which occurs. Studies across other  
152 industries e.g. outdoor activity incidents, have shown the potential to identify multiple contributory  
153 factors, both active and latent, from similar minor events, thus emphasizing the potential explanatory  
154 power of these incidents (e.g. Salmon, Goode, Lenné, Finch & Cassell, 2014; Salmon, Goode, Taylor,  
155 Lenné, Dallat, & Finch, in press). Therefore, gaining an understanding of minor safety-of-the-line  
156 incidents is important to provide rail companies with the tools to prevent similar and more serious  
157 incidents occurring in the future.

158 HFACS was chosen as the tool for the purposes of this study into the analysis of safety of the  
159 line incidents. This was due to the number of studies generally that have used HFACS, its wide  
160 availability and research base that makes its application clear and results transferrable, and its prior  
161 use within the rail sector.

## 162 **1.2 Reliability and Report Quality**

163 Although a number of strengths of the HFACS model have been identified, including its  
164 detailed classification of the organisational context (Baysari, Caponecchia, McIntosh, & Wilson, 2009),  
165 and its ability to provide safety professionals with a theoretically based tool for accident investigations  
166 (Wiegmann & Shappell, 2001); a number of papers have identified some concerns with the reliability  
167 of the model. Beaubien and Baker (2002) and Olsen (2011) criticized the validation evidence  
168 supporting the usefulness of the HFACS system, as it was all collected and analysed by the developers  
169 of the framework. However, other authors have now successfully used and proven the system in a  
170 variety of industries (Baysari et al., 2008; Lenné et al., 2012, Li & Harris, 2006; Reinach & Viale, 2006).



171 Another concern raised by Olsen (2011) is the use of incorrect statistics for the reporting of HFACS  
172 reliability levels. It is argued that Cohen's Kappa is an inadequate measure of reliability, as it is based  
173 on the argument that coders who are coding randomly will agree by chance a certain percentage of  
174 the time, and that this should be deducted from the agreement that is not achieved by chance.  
175 However in incident classification systems, coders are not randomly assigning codes but are actually  
176 trying to identify the same causal factors, and therefore agreements are not chance events (Olsen,  
177 2011). For this reason, Olsen argues that the correct method for calculating inter-coder consensus is  
178 to calculate the index of concordance which takes into account both the total number of agreements  
179 and the total number of disagreements of raters' codes. An additional issue is that a number of authors  
180 have highlighted difficulties with the clarity of error codes within HFACS, particularly in derivatives of  
181 HFACS such as HFACS-ADF (Olsen & Shorrock, 2010) and HFACS-DoD (O'Connor & Walker, 2011).  
182 Baysari et al. (2008) reported a large difference in the number of errors identified by the three raters  
183 in their study, with percentage agreement ranging from 40-75%, and as a result they only reported  
184 the ratings of the first author in their paper. Thus, in this paper the index of concordance is used to  
185 evaluate the reliability of HFACS as a tool for the categorisation of UK rail incident reports by two  
186 Human Factors experts.

187 As outlined in Section 1.1, one of the main benefits of HFACS is in identifying latent factors  
188 that can contribute to accident causation. However, this is dependent on the quality of investigation  
189 and subsequent reporting of accidents. While significant rail accidents are subject to extensive  
190 reporting, it was unclear whether it would be possible to identify latent features of accidents, at both  
191 organisational levels and beyond, in the type of reports generated for minor safety of the line  
192 incidents, or whether these investigations focus more on surface-level features relating to unsafe acts  
193 and their preconditions. Rjabovs and Palacin (2015) found that there was a tendency not to attribute  
194 systemic, physical or design factors to the causation of SPADs in a metro environment, and it is likely  
195 that a similar issue might arise when looking at other types of rail transport. Therefore, this paper also

196 aimed to measure the quality and depth of the information contained in minor incident investigation  
197 reports.

### 198 **1.3 Purpose of current study**

199 This paper presents an application of HFACS as an analysis tool to aid with the understanding of  
200 the factors that contribute to minor operational incidents in the UK rail. It aims to investigate the  
201 breakdown of causal factors for these incidents, and in doing so evaluate whether the patterns found  
202 in Baysari et al. (2008) are replicated in the UK rail industry. The study focuses on incidents which have  
203 previously been defined as being caused by Human Error and addresses five key questions:

- 204 1. Can HFACS help us to identify the precursors of minor operational incidents?
- 205 2. Are there any differences in the causation paths of different types of incident e.g. SPAD vs  
206 station overrun?
- 207 3. What is the breakdown of active and latent factors that contribute to this type of incident and  
208 does this vary across incident types?
- 209 4. What is the quality of reporting of minor incidents in the rail industry? Is report content  
210 sufficient to support the identification of latent factors of incident causation, including  
211 organisational and regulatory?
- 212 5. How reliably can two independent Human Factors experts' code investigation reports using  
213 HFACS?

## 214 **2. Method**

### 215 **2.1 Data Sources**

216 Incident investigation reports were collected from seven of the UK's Train Operating  
217 Companies (TOCs). These incidents had all been previously classified by the TOCs as involving some  
218 form of human error. A total of 74 investigation reports were included, all relating to minor safety-of-  
219 the-line incidents occurring between January 2012 and May 2014. None of the incidents included in

220 this study had been investigated by the Rail Accident Investigation Branch (RAIB), who investigate any  
221 accidents causing death, serious injuries, or extensive damage, or incidents which had the potential  
222 to lead to these serious effects. 5 main types of incident were included:

- 223 • Signals passed at danger (SPADs, N=21)
- 224 • Fail to call incidents, where a train failed to stop at a booked station (N=15)
- 225 • Station Overruns, where a train overran the booked platform at a station (N=19)
- 226 • Stop Short incidents, where a train came to a stop at a station before all carriages were at the  
227 platform (N=10)
- 228 • TPWS Activations, where, for example, a train driver failed to acknowledge a speed restriction  
229 warning (N=9)

## 230 **2.2 Data Coding & Analysis**

231 Investigation reports were independently coded by two Human Factors researchers. Prior to  
232 commencing the HFACS coding, information about each incident was extracted, including a  
233 description of the incident type, the location, and date. Each coder also rated the quality of the  
234 investigation report as low, medium, or high depending on the amount of information included in the  
235 report and the evidence provided for any conclusions drawn. Each report was then read in its entirety  
236 and each contributing/safety factor identified in the incident narrative was mapped to a unique HFACS  
237 category following the procedure identified by Baysari et al. (2008) of using the definitions and tables  
238 provided in Wiegmann and Shappell (2003) and the flow-charts included in Viale and Reinach (2006).  
239 For example, in one report the investigator described a sign that was obscured by undergrowth. This  
240 was extracted as a contributory factor and coded under the Physical Environment HFACS code. The  
241 presence or the absence of each HFACS category was assessed in each accident report narrative. More  
242 than one category or sub-category could be identified at each level. However, to avoid over-  
243 representation from any single accident, each HFACS sub-category was counted a maximum of only  
244 once per accident (Li & Harris, 2006). To begin the analysis process, each analyst first independently

245 coded 10 incidents. This coding was then discussed in detail to ensure a joint understanding prior to  
246 independently analysing the rest of the papers. Where disagreements in the final codes arose, these  
247 were discussed until a consensus was reached.

248           Once the initial analysis had begun, it became apparent that a total of 18 of the contributory  
249 factors identified as belonging in the Environmental Factors category did not fit into either the physical  
250 or technological environment, but rather could be described as arising from the operational  
251 environment. These factors related to unscheduled operational occurrences that were a departure  
252 from the operational norm, and examples included situations where there was a highly unusual  
253 signalling pattern, or a train was re-routed. Therefore, an additional subcategory of Operational  
254 Environment was included for this analysis (see Table 1 for examples).

255

256 **Table 1: Examples of report elements that were included in the Operational Environment category**

- 
1. A signalling fault led to modified working on the train route, requiring the use of hand signals to communicate with the signaller.
  2. A possession on a line led to the driver being directed onto a route that they were not familiar with.
  3. An unusual signalling sequence led to a driver being directed to a different platform than usual.
- 

257           Initial analysis of the incident characteristics and HFACS data were performed using frequency  
258 counts. Further analysis to evaluate the associations between HFACS levels and incident types were  
259 conducted using Chi Square analysis and adjusted standardized residuals (ASR). The ASR provides a  
260 measure of the strength of the difference between observed and expected values in situations when  
261 a cross-tabulation result is associated with more than one degree of freedom i.e. larger than a 2X2  
262 contingency table. An ASR with a value of 2 or greater indicates a lack of fit of the null hypothesis in a  
263 given cell (Sharpe, 2015).

264           In order to evaluate interrater reliability the index of concordance was used to provide a  
265 percentage agreement, following the procedure set out in Olsen and Shorrocks (2010). The proportion

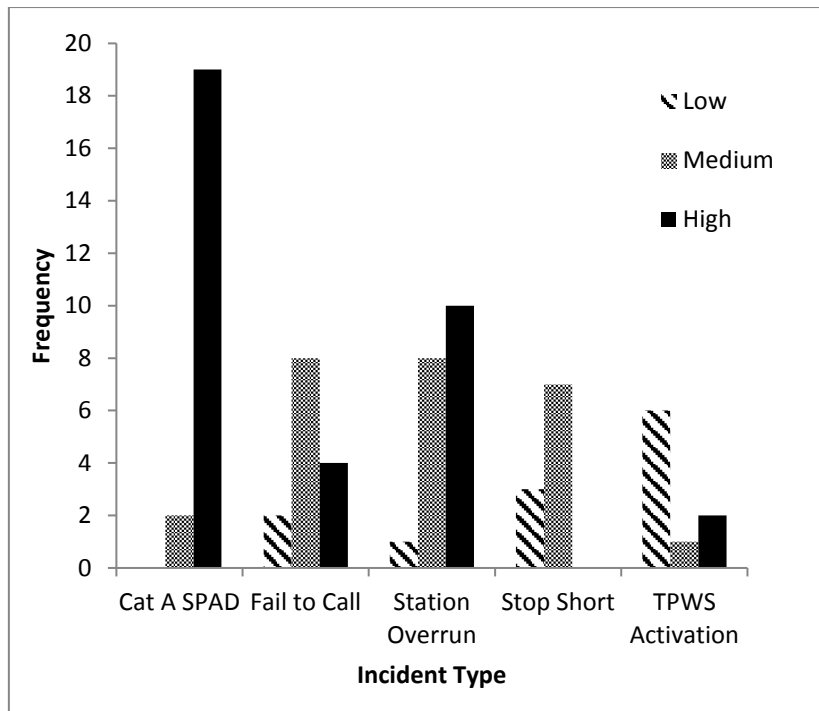
266 of agreeing pairs of codes out of all the possible pairs of codes is calculated as follows: (agreements)  
267 / (agreements + disagreements). Interrater consensus can then be reported as a figure between 0 and  
268 1 or as a percentage. This method takes into account the cases where coders disagreed, along with  
269 providing a method for including situations where there was a difference in the number of codes  
270 assigned between coders. A criterion of 70% agreement between coders was adopted as a reasonable  
271 minimum, in accordance with Wallace and Ross (2006) and Olsen and Shorrock (2010).

## 272 **3. Results**

### 273 **3.1 Inter-Rater Reliability & Quality of Reports**

274 Prior to resolution of any discrepancies in coding between the two raters, the Index of  
275 Concordance was used to evaluate inter-rater reliability (Table 2). The results show that inter-coder  
276 consistency was well above the 70% threshold at both the descriptor and category levels for all  
277 variables other than Adverse Mental state where the consistency was 68.92%. This discrepancy will  
278 be discussed further in Section 3.2.

279 It should be noted that the quality of the incident reports for each of these incident types  
280 varied quite substantially across incident types (see Figure 2), leading to the identification of fewer  
281 contributory factors where the quality was low. Reports categorised as being of low quality generally  
282 contained only tick box information with no supporting data, medium quality reports contained a good  
283 description of the incident with support data and information, but generally did not have a systematic  
284 approach to evaluating human factors. High quality reports contained a good level of support data  
285 and an attempt to systematically evaluate contributory human factors. In general Category A SPADs,  
286 Station Overrun and Fail to Call reports tended to be of a high or medium quality, whereas TPWS  
287 Activation and Stop Short reports tended to have less detail.



288

289

Figure 2: Quality of investigation reports across incident types

290 **Table 2: Inter-rater reliability (prior to resolution) and Frequency counts (post-resolution for each HFACS**  
 291 **category<sup>a</sup>**

Error Categories	Error Subcategories	% Agreement	Frequency <sup>a</sup>	% Reports
<b>Operator Acts</b>				
Skill Based	Attention	77.03	42	56.76
	Memory	81.08	31	41.89
Decision Error	Poor Choice	86.49	9	12.16
Perceptual Error		98.65	1	1.35
Violation	Routine Violation	98.65	2	2.70
	Exceptional Violation	98.65	1	1.35
	Acts of Sabotage	100	0	0
<b>Preconditions to Unsafe Acts</b>				
Environmental Factor	Physical Environment	97.30	6	8.11
	Technological Environment	83.79	13	17.57
	Operational Environment	72.97	18	24.32
Personnel Factor	Crew Resource Management	97.30	6	8.11
	Personal Readiness	91.89	7	9.46
Condition of Operator	Adverse Mental State	68.92	63	85.14
	Adverse Physiological State	90.54	12	16.22
	Physical/Mental Limitations	90.54	10	13.51
<b>Supervisory Factors</b>				
Inadequate Supervision		97.30	2	2.70
Planned Inappropriate Operations		100.00	0	0
Failure to Correct Known Problem		91.89	6	8.11
Supervisory Violations		100.00	2	0
<b>Organisational Factors</b>				
Resource Management		94.59	2	2.70
Organisational Climate		100.00	2	2.70
Organisational Process		93.24	4	5.41

Organisational Violations	100.00	0	0
---------------------------	--------	---	---

292 <sup>a</sup> More than one category could be identified at each of the HFACS levels

293 **3.2 Can HFACS help us to identify the precursors of minor operational incidents which**  
 294 **have the potential to lead to more serious events?**

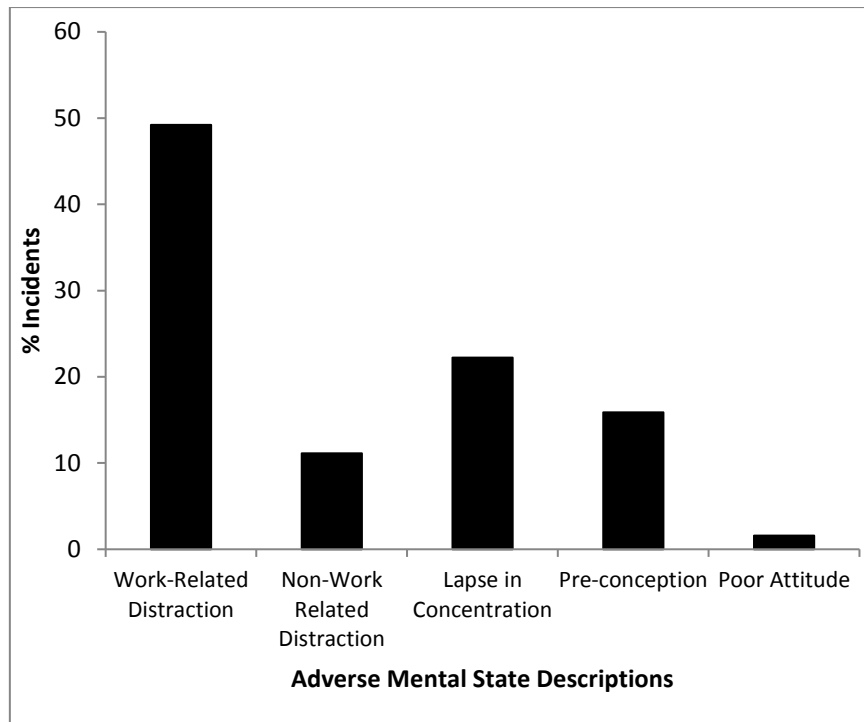
295 It was possible to code all of the contributing factors using our edited version of HFACS  
 296 including Operational Environment. The presence of HFACS codes in the 74 incidents is presented in  
 297 Table 2. A total of 228 contributory factors were identified, with an average of 4.05 factors (SD=1.07)  
 298 per incident.

299 Unsafe acts were identified in all of the reports investigated. The most frequent Level 1 unsafe  
 300 acts were skill-based errors (87.84%). Of these skill-based errors, the majority involved some type of  
 301 attentional failure (56.76% incidents) such as failing to notice the status of a signal or getting  
 302 distracted. 41.89% of the skill based errors involved an issue with memory e.g. forgetting a station  
 303 stop. A decision error was identified in 12.16% of reports, all of which involved a poor choice e.g. not  
 304 making any attempt to stop at a station because of weather conditions. Finally, only 4.05% of unsafe  
 305 acts involved a violation, 2 of which were routine violations e.g. drivers always stopping at a certain  
 306 incorrect part of a platform to avoid passengers having to walk out in the rain, and one of which was  
 307 an exceptional violation involving a failure to clarify instructions.

308 One or more of the Level 2 preconditions for unsafe acts were evident in almost all incidents  
 309 investigated, with one exception (a TPWS Activation). Adverse mental state was identified as a  
 310 precondition in 85.16% incidents. Operational environment (24.32%), technological environment  
 311 (17.57%), adverse physiological state (16.22%), and physical/mental limitations (13.51%) were all also  
 312 identified as Level 2 contributory factors in 10 or more incidents. Unlike the pattern for other  
 313 industries, crew resource management was not a pre-dominant causal factor, and only emerged in  
 314 8.11% incidents.



315 As adverse mental state was deemed to be quite a broad category, and was also the category  
316 where the inter-rater reliability was lowest, it was decided to explore the themes which emerged  
317 within this category further (see Figure 3). Five main themes emerged. The most commonly identified  
318 adverse mental state was work-related distraction, which occurred when drivers claimed to have been  
319 distracted by thinking about something which had occurred during work hours - including problems in  
320 the environment, time pressures, or previous driving patterns. Non-work related distraction occurred  
321 when the driver was distracted by thinking about non-work issues e.g. relationship problems. Lapses  
322 in concentration occurred when the driver claimed to have stopped concentrating on the task for no  
323 particular reason. A preconception refers to situations in which the driver had made an incorrect  
324 assumption about what would happen next. Finally, poor attitude – not following procedures correctly  
325 to avoid having a fault on record - was identified as contributory factor in one incident. As Figure 3  
326 shows, drivers were considerably more likely to be distracted by work-related issues than non-work  
327 related ones. Of the 31 cases in which work-related distraction was identified, environmental issues  
328 were also identified in 18 of these reports (58.06%), suggesting a strong link between any unexpected  
329 changes to the driving environment and the propensity for the driver to lose focus. The weaker inter-  
330 rater reliability of adverse mental state can be accounted for by the fact that one rater was more  
331 inclined to only identify the environmental code in these cases, where the other rater selected both  
332 categories.



333

334

**Figure 3: Breakdown of themes emerging in Adverse Mental State category**

335

Finally, Level 3 supervisory factors and Level 4 organisational factors were both only identified in 10.81% investigations. Failure to correct a problem (8.11%) was the most common supervisory factor, usually resulting from a failure to implement development changes identified in previous incidents.

337

338

The most common Organisational Factor was organisational process (5.41%), usually arising from poor practice and procedures.

339

340 **3.3 Are there any differences in the causation paths of different types of incident?**

341 **Table 3: Frequency counts across Incident Types for each HFACS category**

<b>Error Categories</b>	<b>Error Subcategories</b>	<b>% Cat A SPAD</b>	<b>% Fail to Call</b>	<b>% Station Overrun</b>	<b>% Stop Short</b>	<b>% TPWS Activation</b>
<b>Operator Acts</b>						
Skill Based	Attention	95.24	26.67	42.11	20.00	88.89
	Memory	9.52	86.67	63.16	30.00	0
Decision Error	Poor Choice	4.76	13.33	5.26	40.00	11.11
Perceptual Error		4.76	0	0	0	0
Violation	Routine Violation	4.76	0	0	10.00	0
	Exceptional Violation	4.76	0	0	0	0
<b>Preconditions to Unsafe Acts</b>						
Environmental Factor	Physical Environment	4.76	13.33	5.26	0	22.22
	Technological Environment	23.98	13.33	15.79	30.00	0
	Operational Environment	47.61	33.33	5.26	20.00	0
Personnel Factor	Crew Resource Management	14.29	13.33	5.26	0	0
	Personal Readiness	14.29	0	15.79	0	11.11
Condition of Operator	Adverse Mental State	85.71	93.33	94.74	70.00	66.67
	Adverse Physiological State	9.52	6.67	26.31	10.00	33.33
	Physical/Mental Limitations	19.05	6.67	5.26	30.00	11.11
<b>Supervisory Factors</b>						
Inadequate Supervision		4.76	0	0	0	11.11
Planned Inappropriate Operations		0	0	0	0	0
Failure to Correct Known Problem		19.05	0	0	20.00	0
Supervisory Violations		0	0	0	0	0
<b>Organisational Factors</b>						
Resource Management		9.52	0	0	0	0
Organisational Climate		0	0	10.52	0	0
Organisational Process		4.76	0	5.26	10.00	11.11
Organisational Violations		0	0	0	0	0

342

343 Table 3 shows that there was a difference in the pattern of contributory factors for each of  
 344 the five incident types. In order to determine where significant differences between the groups  
 345 emerged a series of chi-square analyses were conducted. Three of these relationships reached  
 346 significance and these are explored further in Table 4 and Figure 4.

347 **Table 4: Significant associations between HFACS categories and incident type**

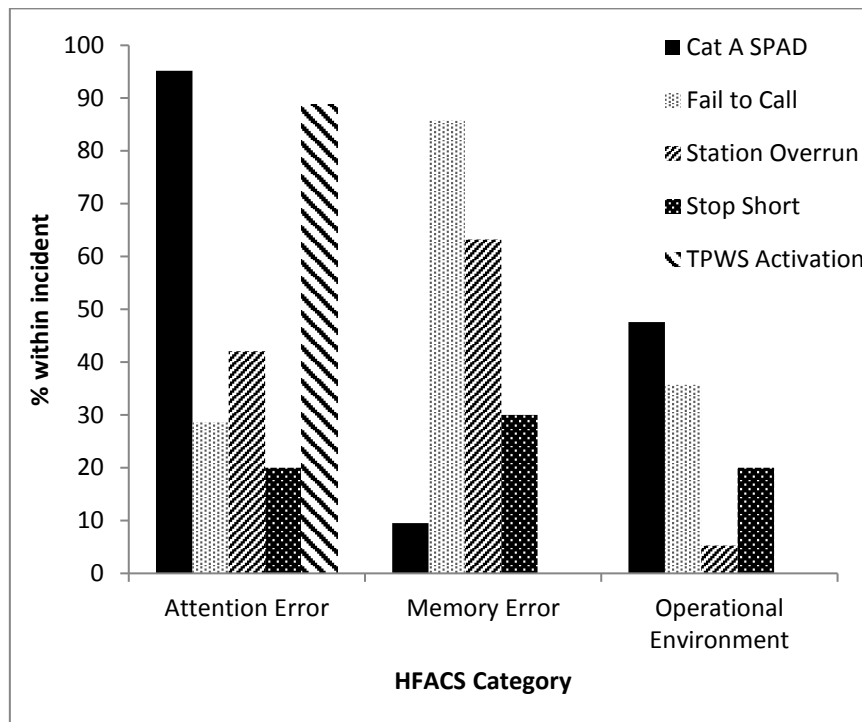
Incident Type		Attention Error	Memory Error	Operational Environment
		$\chi^2=28.26$ (df=4) p<0.001	$\chi^2=31.05$ , df=4, p<0.001	$\chi^2=13.79$ (df=4), p<0.01
Category A SPAD	Observed	20	2	10
	Expected	12.1	8.3	5.2
	<b>ASR<sup>a</sup></b>	<b>4.1</b>	<b>-3.4</b>	<b>2.9</b>
Fail to Call	Observed	4	12	5
	Expected	8.1	5.6	3.5
	<b>ASR<sup>a</sup></b>	<b>-2.4</b>	<b>3.9</b>	<b>1.1</b>
Station Overrun	Observed	8	12	1
	Expected	10.9	7.5	4.7
	<b>ASR<sup>a</sup></b>	<b>-1.6</b>	<b>2.4</b>	<b>-2.3</b>
Stop Short	Observed	2	3	2
	Expected	5.8	4.0	2.5
	<b>ASR<sup>a</sup></b>	<b>-2.6</b>	<b>-0.7</b>	<b>-0.4</b>
TPWS Activation	Observed	8	0	0
	Expected	5.2	3.6	2.2
	<b>ASR<sup>a</sup></b>	<b>2.0</b>	<b>-2.6</b>	<b>-1.8</b>

348 <sup>a</sup>ASR = adjusted standardized residual

349 At level 1 of the HFACS framework, attention and memory errors were both significantly  
 350 associated with incident type. For Category A SPADs (ASR=4.1) and TPWS Activations (ASR=2.0),  
 351 attentional errors were over-represented. However for Fail to Call (ASR=3.9) and Station Overrun  
 352 incidents (ASR=2.4), memory errors were over-represented. This suggests that attention and memory  
 353 errors lead to different outcomes, and thus different initiatives will have to be taken to address each  
 354 incident type.

355 At level 2 of the HFACS framework, operational environment was the only variable to be  
 356 significantly associated with incident type. This category was significantly over-represented in  
 357 Category A SPADs (ASR=2.9). This suggests that Category A SPADs are more likely to occur after some  
 358 change in the operational environment.

There were no significant associations between Level 3 and Level 4 factors and incident type.



360

361 **Figure 4. Percentages related to significant associations between HFACS categories and incident type**

362 **4. Discussion**

363 The aim of the study was to examine the active and latent causal factors of minor safety of  
 364 the line incidents, using the HFACS methodology, and one purpose of the research was to understand  
 365 the utility of HFACS for the task at hand. A number of specific research questions were outlined, which  
 366 are addressed below.

367 **4.1 Can HFACS help us to identify the precursors of minor operational incidents?**

368 74 minor incident investigations were analysed using HFACS to identify the factors which  
 369 contribute to the occurrence of these types of events. In total, 228 contributory factors were identified  
 370 and classified from the reports. The findings provide some initial evidence that the pattern of  
 371 contributory factors for minor incidents is similar to that identified in more serious incidents (e.g.  
 372 Baysari et al., 2008, Read et al., 2012), at least in terms of the Level 1 and Level 2 contributory factors.

373 Consistent with previous research in both rail (Baysari et al., 2008, Reinach & Viale, 2006) and  
374 other sectors (e.g. ElBardissi et al., 2007; Lenné et al., 2012; Li & Harris), skill-based errors emerged as  
375 the most common contributory factor at Level 1, with more attentional than memory errors arising.  
376 However, unlike previous studies, very few violations occurred, with only 2 routine violations and 1  
377 exceptional violation identified. This suggests that minor incidents are more likely to be caused by an  
378 error or mistake than by a deliberate breach of rules. Adverse mental state was the most common  
379 Level 2 category, followed by operational environment, technological environment, and adverse  
380 physiological state. Baysari et al. (2008) also identified adverse mental state as the most common  
381 precondition and, indeed, adverse mental state and environmental factors consistently emerge as  
382 strong contributory factors across a range of sectors, although the order of importance may vary (e.g.  
383 Li & Harris, 2006; Shappell et al., 2007, Lenné et al., 2012). However, in both aviation and medicine,  
384 Crew Resource Management (CRM) also emerges as a common contributory factor (e.g. ElBardissi et  
385 al., 2007; Li et al., 2008; Shappell et al., 2007), which was not identified in this study. This is most likely  
386 a result of the more solitary nature of the train driver role compared to that of an airline pilot or  
387 medical surgeon.

388 Adverse mental state was the most commonly identified category across all of the incidents  
389 investigated. As it is quite a broad category, a deeper analysis was deemed necessary and it was,  
390 therefore, further broken down into 5 main themes. This analysis showed that distraction due to work-  
391 related issues was the single biggest contributory factor. Some caution should be taken in interpreting  
392 this result, as this finding arises from self-report aspects of the report and it is possible that drivers  
393 were unable to accurately remember, or chose to misrepresent what they had been thinking about  
394 prior to an incident. However, the fact that environmental factors, in particular operational  
395 environment, were also identified as a causal factor in over half of the reports suggests that work-  
396 related distraction is a real issue in incident causation.

397 Linked to this, one of the key findings of this study was the importance of the operational  
398 environment. The items in this category were environmental factors that were not overtly physical  
399 (e.g. weather) or technical (e.g. faulty equipment), but altered driving conditions based on operational  
400 circumstances - such as other late running trains in the area causing the incident-involved train to run  
401 on cautionary signals, or a temporary change to the station calling pattern. While these situations are  
402 well within the driver's required competency, they were a deviation from planned or routine action.  
403 Cognitively, changes to the operational environment create a situation where the driver moves from  
404 a skill-based, proactive feedforward mode of control (Rasmussen, 1983; Hollnagel and Woods, 2005),  
405 to a more rule-based, and cognitively effortful (and error prone), reactive mode of control. To amplify  
406 the risk, this change of mode takes place just at that point where the driver is likely to be late or trying  
407 to preserve tight performance allowances in the timetable. Thus, they have the paradox of needing to  
408 work faster at a time when the environment demands, cognitively, that they take longer. Baysari et  
409 al.'s (2008) analysis of Australian railway incidents identified a similar issue, and they advocated the  
410 creation of an extra category of Task Factors at the preconditions for unsafe acts level – many of the  
411 factors they identified could also be considered as part of the Operational Environment.

412 The problems identified in these analyses are not unique to the rail industry, and indeed similar  
413 incidents can easily be identified in other industries. For example, in aviation a flight path may have  
414 to be changed at short notice, or in medicine a routine operation may become more complex due to  
415 unforeseeable complications. Thus, the addition of the category of Operational Environment to HFACS  
416 would provide an additional opportunity to understand the impact of alterations to planned routine  
417 on the propensity for incidents to occur.

418 On the whole, these results highlight the potential power of minor incidents to provide valuable  
419 insights into common causal factors, at least at the unsafe acts and preconditions levels, and to  
420 reinforce some of the similarities (importance of skill-based error, and adverse mental state) and  
421 differences (few violations, increased emphasis on context including operational environment,

422 reduced emphasis on CRM) between train driving and other domains. This highlights that a simple  
423 transfer of initiatives, such as training programmes, from other domains (e.g. aviation) into train  
424 driving is not always appropriate, and indicates where adaption (e.g. an emphasis on attentional over  
425 CRM-type support) is required.

426 **4.2 Are there any differences in the causation paths of different types of incident e.g.**  
427 **SPAD vs station overrun.**

428 Although Li et al. (2013) had compared contributory factors across aircraft type, pilot rank, and  
429 flight phase; this is the first study to investigate the causes of specific incident types within a single  
430 study. Our results indicate that different types of railway incidents appear to have different causal  
431 pathways, at least in terms of the factors immediately preceding the incident. Of particular interest is  
432 the fact that any change in the Operational Environment e.g. a change in diagrammed stops, an  
433 unusual sequence of restrictive aspects; was found to be significantly linked to the occurrence of a  
434 SPAD. Although the SPAD investigations included in this study were relatively minor events with no  
435 major repercussions, similar circumstances have been identified in more serious incidents. As far back  
436 as 1997, a study of over 100 Australian railway incidents identified that the expectation of a green  
437 signal was one of the most common skill-based errors contributing to drivers passing a red signal  
438 (Edkins & Pollock, 1997), and recent major incident investigations have re-iterated this finding (e.g.  
439 RAIB, 2014). Similarly, Rjabovs and Palacin (2016) found that unfamiliar tasks and locations may play  
440 a role in safety of the line incidents in a metro environment. In our paper it may not be that the  
441 location was unfamiliar as such, but that the conditions in which the location was experienced may be  
442 unfamiliar or, at least, a divergence from the norm. This highlights the importance of providing  
443 additional support to drivers in situations which are more cognitively effortful, suggesting that  
444 interventions which specifically address the methods of communicating and alerting drivers to areas  
445 of importance during changes to the operational environment could be successful in reducing the  
446 occurrence of SPADs.



447 In addition, it appears that Category A SPADs and TPWS activations (which have the capacity  
448 to escalate to become a SPAD) were both more likely to be caused by an attentional failure, while Fail  
449 to Call and Station Overrun incidents were more likely to be caused by a memory failure. The fact that  
450 different causal paths are emerging suggests that companies need to take different approaches to  
451 how they address these incidents and, in some cases, technical solutions will be required, similar to  
452 the ones reported by Basacik and Gibson (2015) for wrong side door openings. Read, Lenné, and Moss  
453 (2012) found that task demand factors (e.g. high workload, distraction) were significantly associated  
454 with skill-based errors in Australian rail accidents. We have further broken this down to show how the  
455 impacts of different types of skill-based error (i.e. memory versus attention) can vary, suggesting that  
456 safety interventions need to be carefully targeted to maximise their benefits. For example, technical  
457 systems to more clearly alert drivers of diagrammed station stops may be beneficial in preventing Fail  
458 to Call and Station Overrun incidents, whereas improving communication of the likely risk areas during  
459 non-routine running may reduce the risk of a SPAD.

460 **4.3 What is the breakdown of active and latent factors that contribute to this type of**  
461 **incident, and does this vary across incident types?**

462 Active factors dominated the causes identified from the incident analysis. Due to the small  
463 number of organisational and supervisory factors identified, it was impossible to identify any causal  
464 paths originating at these levels. In addition, some of the reports around TPWS activations, Stop Short,  
465 and Fail to Call incidents were of a low quality containing minimal information, which was usually  
466 related solely to driver error – no Supervisory or Organisational Factors were identified in any of the  
467 Fail to Call reports. In these reports it was often quite difficult to build a picture of the events which  
468 led up to the incident. Although, these incidents are often seen as quite minor, and companies have  
469 to make trade-offs in terms of the costs associated with detailed investigations; being able to address  
470 the causes of these minor incidents and eliminate them is likely to significantly reduce the risk of a  
471 more serious incident occurring (Wright & van der Schaaf, 2004), and result in greater savings in the  
472 long run. The fact that it was possible to identify differences in causal pathways from even basic quality

473 minor investigations provides evidence of the importance of using minor events and near misses to  
474 further our understanding of how safety systems can be improved.

475 It is important to note, however, that even in reports with extensive data e.g. for SPADs or Station  
476 Overruns, there were still few references to organisational and supervisory issues, and many that were  
477 identified were cases where a driver had not yet completed relevant training after a prior incident  
478 (classified as 'Failure to correct known problem'). This indicates an issue with the focus of reporting,  
479 discussed next. Certainly, the perception of driver error as captured in the reports is that the issues lie  
480 in active factors, and this reinforced by train operating companies' interest in Non-Technical Skills  
481 programmes.

#### 482 **4.4 What is the quality of reporting of minor incidents in the rail industry?**

483 Building on the point above, one of the questions entering into this study was whether reports  
484 contained enough detail to identify issues arising at the supervisory, organizational and regulatory  
485 levels. In practice the number of examples of this kind of factors in the data were few and far between.  
486 This is one of the major drawbacks of using HFACS as a tool to investigate more minor accidents, as  
487 several studies have found that systems approaches are hugely dependent on the quality of the data  
488 provided (e.g. Lenné et al., 2012). The majority of the investigations reported in this study were carried  
489 out by front-line supervisors rather than dedicated accident investigators, and thus it is perhaps  
490 unsurprising that these supervisors might be reluctant to find fault with themselves and, in many  
491 cases, their employers. Research shows that latent errors pose the greatest risk to system safety  
492 (Reason, 1990; 1997), and it is a key characteristic that these latent errors are the pre-conditions that  
493 enable active errors to occur. It is therefore important that organisations are able to identify these  
494 latent errors to mitigate against potentially serious accidents occurring in the future.

495 However, it is important not to appear too critical of reporting. Of all 74 reports identified by train  
496 operators as being related to human error, all did cover human error and presented issues that fitted  
497 naturally within HFACS. None presented information that suggested a significant misclassification of

498 the report (e.g. that it was primarily a technical fault). This suggests a good level of understanding of  
499 basic human factors within the industry, and further work could help to refine or expand that  
500 understanding to seek out more latent factors. Further work to develop investigation and reporting  
501 around supervisory, organisation and external factors should not just look to support accident analysis  
502 using HFACS. This level of reporting would also help assist in the identification of causes of accident  
503 using systems-orientated approaches such as STAMP (Leveson, 2004) and Accimap (Rasmussen,  
504 1997).

#### 505 **4.5 How reliably can two independent Human Factors experts code investigation reports** 506 **using HFACS**

507 On the whole, the research team found HFACS to be a straightforward tool to use, although it was  
508 not without its flaws. Previous research had identified problems with inter-rater reliability, and  
509 difficulties in identifying the level at which factors should be categorised (Olsen, 2011; Olsen &  
510 Shorrock, 2010; Baysari et al., 2008). Olsen (2011) investigated the success of air traffic controllers and  
511 human factors specialists in applying HFACS consistently and found that neither group achieved  
512 acceptable agreement levels between raters. However, this was not a problem in the current study,  
513 with inter-rater reliability reaching an acceptable level in all categories other than Adverse Mental  
514 State, where it was just below the 70% agreement level advocated. Prior to beginning the coding  
515 process, both raters had spent some time agreeing on their interpretation of each of the categories  
516 and this may have aided the coding process. Also, all incidents had already been classified by the train  
517 operating companies as relating to human error, which again may have reduced some of the scope  
518 for variance.

#### 519 **4.6 Limitations**

520 A limitation of this study, particularly for TPWS activation and Stop Short events was the lack  
521 of data in the reports, and, as noted above, all of the reports lacked information on supervisory and  
522 organisational factors. This, coupled with a modest sample size of 74 investigation reports, limits the

523 depth of conclusions that can be drawn from the reports regarding causal factors. As noted under data  
524 quality, a second factor is the potential bias in the reports through the reliance on the skills of the line  
525 managers and supervisors as investigators. These investigators could not be assumed to have  
526 extensive training or knowledge of Human Factors, and may have a personal relationship with the  
527 driver they were interviewing. Thirdly, putting aside the role of the investigator, the drivers were asked  
528 to recall their thoughts and mental states at the time of the incident. This is also likely to be biased,  
529 and caution must be taken when interpreting any self-report data. A final limitation is that HFACS was  
530 the only interpretation tool used in the study. While the aims of the study were practical, rather than  
531 a study of methodology, it might be useful to compare different tool outputs e.g. Accimap (Rasmussen,  
532 1997), STAMP (Leveson, 2004), along with the Incident Factor Classification Study which is being  
533 adopted in the UK rail section (Gibson et al., 2015).

## 534 **5. Conclusions**

535 The current study successfully applies HFACS to provide a retrospective analysis of minor  
536 incident investigations in the rail industry. Such examination of minor incidents provides a much wider  
537 scope for us to interpret accident causal pathways, as these incidents occur much more frequently  
538 than more serious incidents. By highlighting the differences in the causes of different incident types,  
539 a greater level of understanding of the mechanisms required to prevent future incidents is achieved.

540 Active failures, specifically those related to attention and adverse mental state, dominate the  
541 results, suggesting that measures to reduce safety of the line incidents should be targeted at these  
542 areas. However, it is important to stress that training approaches should not be the only solution, and  
543 more systemic solutions are also required. Currently, supervisory and organisational issues are under-  
544 represented in the reports, and therefore more efforts should be made to identify latent factors that  
545 might be setting up the preconditions for active failures. Uncovering these latent errors may need rail  
546 companies to refine the current approach to minor incident investigation, in order to ensure that all  
547 factors can be identified, not only those relating to the competency or attitude of the driver.

548 Finally, this study has also identified the importance of the operational environment in  
549 shaping risk. Gibson et al (2015) put the case that as an aggregate, performance may be approaching  
550 a ceiling, and that further investigation is required to target specific locations or circumstances that  
551 might lead to error. From this analysis, we argue that operational environment may be one of those  
552 factors. To test this, one could compare the risk of SPAD for signals approached at red when  
553 operational conditions were out of the norm, from those approached in normal circumstances. If  
554 operational environment is a factor, then SPAD risk will be found to be higher. Also, it would also be  
555 interesting to investigate whether similar differences emerge in the causal factors of incidents on  
556 different types of routes (e.g. high-speed trains versus metro-links).

## 557 **Acknowledgements**

558 This project was co-funded by Innovate UK, ESRC and EPSRC as a Knowledge Transfer Partnership.  
559 We are also grateful to the UK Train Operating Companies who provided access to and information  
560 about their incident investigation processes, without whom this research would not have been  
561 possible.

562

## 563 **References**

564 Atkins. (2003). Research programme management rail-specific HRA technique for driving tasks user  
565 manual. *Rail Safety Standards Board Research Catalogue*.

566 Basacik, D. & Gibson, H. (2015) Where is the platform? Wrong side door release at stations. In:  
567 Sharples, S., Shorrock, S. and Waterson, P. (Eds) *Contemporary Ergonomics and Human Factors*  
568 *2015, Proceedings of the International Conference on Ergonomics & Human Factors 2015, Daventry,*  
569 *Northamptonshire, UK, 13-16 April 2015. London: Taylor and Francis.*

570 Baysari, M.T., Caponecchia, C., McIntosh, A.S., & Wilson, J.R. (2009). Classification of errors  
571 contributing to rail incidents and accidents: A comparison of two human error identification  
572 techniques. *Safety Science*, 47(7), 948-957.

573 Baysari, M.T., McIntosh, A.S., & Wilson, J.R. (2008). Understanding the human factors contribution to  
574 railway accidents and incidents in Australia. *Accident Analysis & Prevention*, 40(5), 1750-1757.

575 Beaubien, J.M., & Baker, D.P. (2002). A review of selected aviation human factors taxonomies,  
576 accident/incident reporting systems and data collection tools. *International Journal of Applied*  
577 *Aviation Studies*, 2(2), 11-36.

578 Celik, M., & Cebi, S. (2009). Analytical HFACS for investigating human errors in shipping accidents.  
579 *Accident Analysis & Prevention*, 41(1), 66-75.

580 Chen, S-T., Wall, A., Davies, P., Young, Z., Wang, J., & Chou, Y-H. (2013). A human and organisational  
581 factors (HOFs) analysis method for marine casualties using HFACS-Maritime Accidents (HFACS-MA).  
582 *Safety Science*, 60, 105-114.

583 Department for Transport (2014). *Transport Statistics Great Britain 2014*. Retrieved from  
584 <https://www.gov.uk/government/statistics/transport-statistics-great-britain-2014>.

585 Edkins, G.D. & Pollock, C.M. (1997). The influence of sustained attention on railway accidents.  
586 *Accident Analysis & Prevention*, 29(4), 533-539.

587 ElBardissi, A.W., Wiegmann, D.A., Dearani, J.A., Daly, R.C., & Sundt, T.M. (2007). Application of the  
588 human factors analysis and classification system methodology to the cardiovascular surgery  
589 operating room. *The Annals of Thoracic Surgery*, 83(4), 1412-1419.

590 FRA. (2007). *Role of Human Factors in Rail Accidents*. Washington DC: U.S. Government Printing  
591 Office.

592 French, S., & Cope, J. (2012). A review of human factors identified in investigations by the Rail  
593 Accident Investigation Branch (RAIB). *Paper Presented at the International Railway Safety*  
594 *Conference*, London, UK.

595 Gibson, W. H., Mills, A., Basacik, D., & Harrison, C. (2015). The Incident Factor Classification System  
596 and Signals Passed at Danger. *Paper Presented at the 5th Conference of Rail Human Factors*, London,  
597 UK.

598 Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management*. Surrey, UK:  
599 Ashgate.

600 Hollnagel, E. & Woods, D.D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems*  
601 *Engineering*. Boca Raton, FL: Taylor & Francis.

602 Kyriakidis, M., Pak, K.T., & Majumdar, A., (2015). Railway accidents due to Human Error: A historic  
603 analysis of the UK Railways (1945-2012). *Transportation Research Record: Journal of the*  
604 *Transportation Research Board*, forthcoming.

605 Lawton, R., & Ward, N.J. (2005). A systems analysis of the Ladbroke Grove rail crash. *Accident*  
606 *Analysis & Prevention*, 37(2), 235-244.

607 Lenné, M.G., Salmon, P.M., Liu, C.C., & Trotter, M. (2012). A systems approach to accident causation  
608 in mining: An application of the HFACS method. *Accident Analysis & Prevention*, 48, 111-117.

609 Leveson, N.G. (2004). A new accident model for engineering safer systems. *Safety Science*, 42, 237-  
610 270.

611 Li, W.C., & Harris, D. (2006). Pilot error and its relationship with higher organizational levels: HFACS  
612 analysis of 523 accidents. *Aviation, Space, and Environmental Medicine*, 77(10), 1056-1061.

613 Li, W.C., & Harris, D. (2013). Identifying training deficiencies in military pilots by applying the human  
614 factors analysis and classification system. *International Journal of Occupational Safety and*  
615 *Ergonomics*, 19(1), 3-18.

616 Murphy, P. (2010). The role of communications in accidents and incidents during rail possessions.  
617 In: *Engineering Psychology and Cognitive Ergonomics: Aerospace and Transportation Systems*. Paper  
618 presented at the Third International Conference on Engineering Psychology and Cognitive  
619 Ergonomics: Aerospace and Transportation Systems. Edinburgh, Scotland.

620 NRPS (2016). Rail Passenger Satisfaction at a glance: Great Britain – Autumn 2015. Accessed at  
621 [http://www.transportfocus.org.uk/research/publications/national-rail-passenger-survey-nrps-at-a-](http://www.transportfocus.org.uk/research/publications/national-rail-passenger-survey-nrps-at-a-glance-great-britain-wide-autumn-2015)  
622 [glance-great-britain-wide-autumn-2015](http://www.transportfocus.org.uk/research/publications/national-rail-passenger-survey-nrps-at-a-glance-great-britain-wide-autumn-2015) on 27th January 2016

623 O'Connor, P., & Walker, P. (2011). Evaluation of a human factors analysis and classification system as  
624 used by simulated mishap boards. *Aviation, Space, and Environmental Medicine*, 82(1), 44-48.

625 Olsen, N.S. (2011). Coding ATC incident data using HFACS: Inter-coder consensus. *Safety Science*,  
626 49(10), 1365-1370.

627 Olsen, N.S., & Shorrock, S.T. (2010). Evaluation of the HFACS-ADF safety classification system: inter-  
628 coder consensus and intra-coder consistency. *Accident Analysis & Prevention*, 42(2), 437-444.

629 ORR (2014). 2013-14 Annual Statistical Release: Safety Key Statistics. Retrieved 19<sup>th</sup> August, 2015  
630 from [http://orr.gov.uk/ data/assets/pdf file/0006/14784/key-safety-statistics-release-2013-14.pdf](http://orr.gov.uk/data/assets/pdf_file/0006/14784/key-safety-statistics-release-2013-14.pdf)

631 Patterson, J.M. & Shappell, S.A. (2010). Operator error and system deficiencies: analysis of 508  
632 mining incidents and accidents from Queensland, Australia using HFACS. *Accident Analysis &*  
633 *Prevention*, 42(4), 1379-1385.

634 RAIB (2014). Unauthorised entry of a train onto a single line at Greenford, 20 March 2014. Accessed  
635 at



636 [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/408651/141222\\_R](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/408651/141222_R)  
637 [292014\\_Greenford.pdf](#) on 25th July 2016.

638 Rasmussen, J. (1983). Skills, rules, and knowledge: signals, signs and symbols, and other distinctions  
639 in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(3), 257-266.

640 Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science*,  
641 27, 182-213.

642 Read, G.J., Lenné, M.G., & Moss, S.A. (2012). Associations between task, training and social  
643 environmental factors and error types involved in rail incidents and accidents. *Accident Analysis &*  
644 *Prevention*, 48, 416-422.

645 Reason, J.T. (1990). *Human Error*. Cambridge: Cambridge University Press.

646 Reason, J.T. (1997). *Managing the risks of organizational accidents*. Ashgate: Aldershot, UK.

647 Reason, J.T., Hollnagel, E., & Paries, J. (2006). Revisiting the Swiss cheese model of accidents. *Journal*  
648 *of Clinical Engineering*, 27, 110-115.

649 Reinach, S., & Viale, A. (2006). Application of a human error framework to conduct train  
650 accident/incident investigations. *Accident Analysis & Prevention*, 38(2), 396-406.

651 Rjabovs, A., & Palacin, R. (2015). Attitudes of Metro Drivers Towards Design of Immediate Physical  
652 Environment and System Layout. *Urban Rail Transit*,1(2), 104-111.

653 Rjabovs, A., & Palacin, R. (2016). The influence of system design-related factors on the safety  
654 performance of metro drivers. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal*  
655 *of Rail and Rapid Transit*, 0954409716630007.

656 RSSB (2009). An analysis of formal inquiries and investigations to identify human factors issues:  
657 Human factors review of railway incidents. *Rail Safety and Standards Board Research Catalogue*.

658 Salmon, P.M., Cornelissen, M., & Trotter, M.J. (2012). Systems-based accident analysis methods: A  
659 comparison of Accimap, HFACS and STAMP. *Safety Science*, 50, 1158-1170.

660 Salmon, P., Goode, N., Lenné, M.G., Finch, C.F., & Cassell, E. (2014). Injury causation in the great  
661 outdoors: A systems analysis of led outdoor activities. *Accident Analysis and Prevention*, 63, 111-120.

662 Salmon, P.M., Goode, N., Taylor, N., Lenné, M.G., Dallat, C.E., & Finch, C.F. (in press). Rasmussen's  
663 legacy in the great outdoors: A new incident reporting and learning system for led outdoor activities.  
664 *Applied Ergonomics*. Advance online publication. doi:  
665 <http://dx.doi.org/10.1016/j.apergo.2015.07.017>

666 Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D.A. (2007). Human  
667 error and commercial aviation accidents: An analysis using the Human Factors Analysis and  
668 Classification System. *Human Factors*, 49, 227-242.

669 Sharpe, D. (2015). Your chi-square test is statistically significant: Now what? *Practical Assessment,*  
670 *Research & Evaluation*, 20, 1-10.

671 Viale, A., & Reinach, S. (2006). A pilot examination of a joint railroad management-labour approach  
672 to root cause analysis of accidents, incidents, and close calls in a diesel and car repair shop  
673 environment. Federal Railroad Administration.

674 Wallace, B. & Ross, A. (2006). *Beyond Human Error: Taxonomies and Safety Science*. CRC Press, Boca  
675 Raton, FL.

676 Wiegmann, D.A., & Shappell, S.A. (2001). Human error analysis of commercial aviation accidents:  
677 Application of the Human Factors Analysis and Classification System (HFACS). *Aviation, Space, and*  
678 *Environmental Medicine*, 72(100), 1006-1016.

679 Wiegmann, D.A. & Shappell, S.A. (2003). *A Human Error Approach to Aviation Accident Analysis: The*  
680 *Human Factors Analysis and Classification System*. Ashgate: Aldershot.

681 Wright, L., & Van der Schaaf, T. (2004). Accident versus near miss causation: A critical review of the  
682 literature, and empirical test in the UK railway domain, and their implications for other sectors.  
683 *Journal of Hazardous Materials*, 111(1), 105-110.

684

685

686

687

688