A framework for human factors analysis of railway on-train data

Nora Balfe
Centre for Innovative Human Systems, Trinity College Dublin
Ireland

Railway operations are increasingly captured using digital technologies, such as on-train-data-recorders (OTDR) which record all control inputs by the train driver along with other metrics including speed, distance travelled, and GPS coordinates. Already widely used for fleet management and fault finding, the data may also have a potential human factors application in analysing and improving railway operations, for example by providing leading indicators of train driver performance or highlighting infrastructure sections correlated with poor driving performance across all drivers. This research explores the possible use of such data, and the barriers to be overcome in its application, including the size of the data sets, the reliability of the data and the identification of useful features or metrics within the data. A framework of a typical train journey is presented, breaking the journey into segments within which the OTDR data can be analysed. The key metrics available from the OTDR that may be applicable to each journey segment are discussed, along with the potential benefits of utilising the data in this context and a roadmap for future research in the area.

Introduction

Human factors research on the train driving task dates back as least as far as Branton, who in 1979 published a paper discussing the nature of train driving and the need for drivers to anticipate future actions, develop internal representations of the railway (called route knowledge), and test these representations against reality. Today, rail human factors is a growing and vibrant discipline, which examines the ways in which human performance in railway operations can be supported (Wilson & Norris, 2005). Changes in technology generate parallel changes in the role and demands of the human operators, many of which are positive but some of which have unexpected consequences and the human factors discipline is constantly developing methods and approaches to study, analyse and improve the system. This paper presents and discusses the possible contribution of on train data recorders (OTDR) in this context.

Digital train event recorders are increasingly placed on board rail vehicles, capturing information on train movements and component status, but currently this detailed information is not widely used outside of incident investigations and maintenance management (de Fabris et al., 2008; Mosimann & Rios, 2014). Walker and Strathie

In D. de Waard, A. Toffetti, R. Wiczorek, A. Sonderegger, S. Röttger, P. Bouchner, T. Franke, S. Fairclough, M. Noordzij, and K. Brookhuis (Eds.) (2017). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2016 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

(2014) suggest that train recorder data is an underused but potentially important data source for understanding human performance and detecting risks in advance of accidents. This is particularly important in the context of the current safety performance of the rail industry, characterised by very few major incidents and relatively stable safety performance indicators. There are no major, obvious changes to be made to improve safety and new ways of looking at and using data are needed to give further insight into issues and potential improvements, as in the Safety II approach (Hollnagel, 2014), which proposes that human performance variability is a key part of system resilience and by studying how this variability contributes to good system performance we can better manage safety.

As in other safety critical industries, human performance is a major contributor to safe railway operations, with an analysis by Evans (2011) finding that 69% of fatal railway accidents across seven European countries between 1980 and 2009 were primarily related to human performance. The train-driving task is primarily visual-spatial, involving constant perception and processing of information (Gillis, 2007) and the majority of train driver physical actions are driven by information received (e.g. moving the traction handle in response to a change in the speedometer), placing a strong requirement for visual attention on the train driver. Key tasks involve processing information collected from both inside and outside the cab and applying route knowledge to correctly control the speed and braking of the train (Doncaster, 2012; Hamilton & Clarke, 2005; Buksh et al., 2013). However, despite the apparent simplicity of the task, Naweed (2014) describes the train-driving task as complex, dynamic, and opaque. Although the basic tasks may be described reasonably simply, the actual practice involves changing conditions, event densities, and performance pressures that drive adjustments in motor skills and problem solving strategies. The complexity is driven by sometimes conflicting goals of time-accuracy, comfort, and speed regulation and the trade-offs required to optimise the overall journey. The dynamism comes from the constant need to regulate speed and finally, the opacity is due to the gaps in information when working with lineside signalling; drivers must use their route knowledge to infer future requirements. Thus, train driver performance is not simply a matter of perceiving and responding to stimuli as suggested by the use of simple information processing models, but is driven by continuous, proactive predication and planning (Elliott et al., 2007).

A key question in monitoring and managing driver performance is, what are the attributes of good train driving? There is very little information in the literature directly addressing this question, although some papers list the attributes of good drivers (e.g. strong mental models, ability to anticipate, good concentration; Russell & Long, 2005). Some experimental studies have used measures such as use of braking power, adherence to speed limits, variability of speed, and reaction to warnings as dependent variables linked to driver performance (Dorrian et al., 2005, 2006; Dunn & Williamson, 2012; Robinson et al., 2015) but there has not been specific validation of such measures as general indicators of driver performance, or any work to integrate different measures into a unified model. Driver performance measures are starting to be widely used in the road transport domain, where they can be collected by in-vehicle data recorders (IVDR) These devices can identify undesirable driving behaviours, such as speed, hard braking, accelerating, sharp turning, and swift lane changes (Albert et al., 2011) and, in specifically equipped vehicles, even driver gaze and distance to lateral road marking (Pérez et al, 2010). In current implementations, this information is typically communicated to driver live while they are driving or as a performance summary at the end of the trip. Research has suggested that such feedback is linked to improved performance (e.g. Musicant et al., 2010; Donmez et al., 2008) and similar benefits may be attainable from the use of OTDR to provide feedback on train driver performance. Professional road drivers tend also to be monitored on energy efficiency, which may also be possible in the rail environment using OTDRs.

Research in OTDR use in railway performance monitoring is more limited than IVDR, but a small number of studies have been published. Two studies (Walker & Strathie, 2014; Rashidy et al., 2016) analyse OTDR data to explore train driver interaction with warning systems. Interesting findings include a high false alarm rate (Walker & Strathie, 2014) and a consistently speedy response to alarms from some drivers (Rashidy et al., 2016). Such research provides highly useful insights into the effectiveness of real-world warning devices and highlights areas for possible improvements, but only utilises a small portion of the available data. Strathie and Walker (2015) have also applied link analysis and associated graph theory to the analysis of on train data recorders. The analysis linked each control action by a driver to their next control action in a diagrammatic form, and facilitated the further analysis of driver styles and differences. The results found that there was a difference in the number of links between elements of the control interface (i.e. some drivers moved between more pairs of controls than other drivers). The number of links was found to be fairly consistent within some drivers, i.e. they reflected a stable driving style of that particular driver, but variable across others. However, the data cannot be used to determine whether this reflects an inherently unstable driving style or external factors motivating different behaviours on each journey analysed. Other analyses, such as the sociometric status of a node, the mean number of throttle moves per journey, the network diameter (or number of links in a chain of movements), all also showed promise for differentiating between driver styles. These indicators have not yet been linked to ‘good’ driver performance, but many ultimately provide a way to differentiate between good and poor driving behaviours.

Finally, Green et al. (2011) proposed the use of OTDR data in assessing driver performance. They note that such data is used in a qualitative fashion in current competence management, but that there is very little attempt to classify drivers according to their driving ability or risk. They identified five routine events for an initial implementation of driver performance monitoring, but did not provide any justification or evidence supporting these as valid metrics able to distinguish between different levels of driver performance:

1. Speed at which power notch 4 is selected when accelerating (passenger comfort)
2. Percentage of time in a braking sequence that the driver selects brake step 3 (passenger comfort)
3. Speed over TPWS grids approaching a Permanent Speed Restriction (PSR) (train speed)

4. Speed through a PSR as a percentage of the maximum speed (train speed)
5. Mean speed when an AWS horn is received (train speed)

These metrics can all be automatically calculated from the OTDR downloads, and displayed on a dashboard showing each drivers' performance. In addition, software could also detect error events, such as wrong-side door releases, stopping at the wrong stop mark, and speeding. The same statistics could also be analysed by route, to highlight areas of concern on the infrastructure.

To date, the limited research into the use of OTDR for performance monitoring has been piecemeal and/or unvalidated. A more coherent, scientific approach is necessary to guide future research. The aim of this paper is to present a framework within which to analyse and research OTDR application in human performance monitoring in a more structured fashion than has thus far been achieved. The paper will also discuss the potential use of the signals available in monitoring and improving human performance on the rail network.

OTDR Data

The OTDR record a wide range of signals; only those potentially relevant to driver performance will be discussed in this paper. The data are logged each time one of the monitored signals changes status, i.e. if any signal changes, all signal statuses are logged. The majority of the signals are digital, recording in bitcode format. The exceptions to this are the timestamp on each recording, the distance travelled in kilometres, the latitude and longitude at each recording, the system speed in km/h, and the brake pressure in bar. The distance travelled is an additive signal over the life of the unit (or until reset), so to calculate the distance travelled in terms of a particular journey, the distance value at the start must be known and subtracted from each subsequent recording. The brake and power notches selected by the driver can be derived from the relevant bitcodes. Other driver actions captured include: gear selected (forward, neutral, reverse), headlight selection (none, dipped beam, full beam), use of the horn, door openings (left and right side), and emergency brake activations.

The Irish Rail system, on which the current research has been conducted, features a Continuous Automatic Warning System (CAWS) over parts of the network. The CAWS system provides the driver with an in-cab display of the last signal aspect (colour) encountered. Signal aspects dictate the speeds at which the train can safely travel while still being able to stop at red aspects. The CAWS system also draws the drivers' attention to restrictive aspects, and the driver must acknowledge this auditory warning. The signal aspect according to CAWS is recorded in the OTDR, along with any activations of the acknowledgement button by the driver.

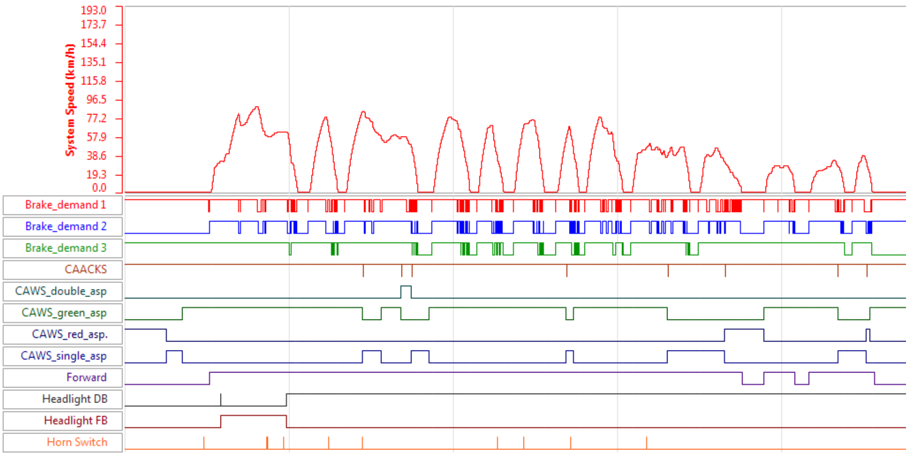

Figure 1. Sample of OTDR Data

Figure 1 illustrates a graphical view of some of the available data showing the train speed, the bitcodes for the brake demand, CAWS aspects and acknowledgements, gear selection and headlight and horn use. The data can be exported to Excel for further analysis.

Analysis Framework

Investigating the potential uses of this data for train and/or infrastructure performance monitoring requires the use of a framework to structure the data analysis. The proposed framework is shown in Figure 2. The overall journey is broken into segments between station stops. Station stops can be reliably identified by those occasions when the recorded train speed is zero and the train doors are opened. The framework then defines six distinct phases within each segment for analysis:

a) Station duties – defined as the time between the doors opening at the station stop (t_1) and the power being applied to leave the station (t_2). Possible metrics of interest in this phase include station dwell times, boarding times, time between doors closing and brake release/power applied, neutral gear selection, and the application of the brakes throughout the stop.
b) Station departures – defined as the time between the power being applied (t_2) and the first speed peak achieved on leaving the station (t_3). Possible metrics of interest in this phase include the power profile (i.e. the specific power notches used when applying power), acceleration rates (several different metrics may be possible), and maximum speed achieved.

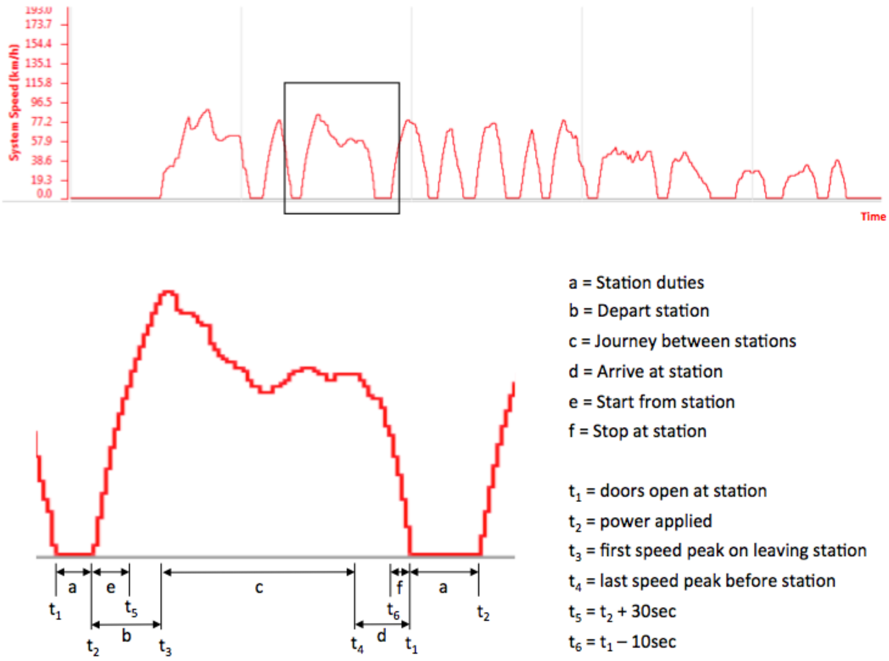

Figure 2. Analysis framework

c) Journey between stations – defined as the time between the first speed peak on departing the station (t_3) and the application of the brakes on approach to the next station (t_4). The application of brakes is analogous to the last peak in the speed profile, and although the application of brakes is not clearly identifiable in the speed profile presented, it is easily identifiable in the raw data. This phase does not occur on short journey segments where braking for the approaching station occurs immediately after the first speed peak but may be quite prolonged on journeys with lengthy intervals between stations. Possible metrics of interest in this phase include speed adherence (where data on line speed is known) and the variation in speed achieved using the median speed and a measure of dispersion.

d) Station arrival – defined as the time from the start of the brake application (t_4) and the doors opening at the station (t_1). Possible metrics of interest in this phase include deceleration rate, braking profiles, maximum brake level applied, and final brake application. More advanced analysis of longitudinal dynamics may also be useful, as well as analysis of power consumption.

The timeframe for station departures and station arrivals may vary widely between different journey segments, due to differences in the distance and permitted speed between the stations. Two further phases are therefore defined to facilitate comparison of departures and arrivals between stations:

e) Station starts – defined as the time between power being applied to leave the station (t_2) and approximately 30 seconds following this (t_5). Metrics of interest are likely to be similar to (b)
f) Station stops – defined as the last ten seconds before (t_6) the doors open at the station (t_1). Metrics of interest are likely to be similar to (d)

Both timeframes (30 seconds and 10 seconds) are arbitrary points, proposed to capture the initial acceleration away from the station and final deceleration towards the station. However, neither timeframe has yet been validated as the most appropriate.

In addition to each of the journey phases, some overall journey statistics may be useful. Speed adherence may be an obvious method to monitor driver performance, but to achieve this with OTDR data would require a model of the network populated with the permitted speeds. This is not currently easily available and speed adherence can be only manually assessed directly from the data. Other metrics from the overall journey include use of the horn, headlight usage, speeds at downgrades to red signals, use of emergency brake, the percentage of time running on restrictive signals, and overall power and brake profiles.

Analysis Potential

The analysis of OTDR data within the defined framework still presents a number of research challenges. First, the data generated is extremely large, with approximately 10,000 lines of data generated for each hour of travel with each signal being logged in each line of data. Big data techniques may be applied, but techniques such as map-reduce (Chu et al., 2007) are unlikely to give the required insights into performance and more complex processing of the data is likely to be required. This requires a computer programming and statistical analysis skill set that is likely beyond most HF practitioners. The data also requires extensive pre-processing to generate information. For example, the braking and power bitcodes must be combined and referenced to generate the actual brake or power selection. Such processing can be relatively easily achieved for a small number of journeys but must be automated in order to facilitate a more comprehensive analysis. The data also suffers from missing signals, such as door closing and wrong recordings, such as errors in the signal aspects generated by the CAWS system and incorrect power levels. These can complicate an automatic analysis and currently, in the research upon which this framework is based, must be manually identified and rectified. Such issues are likely to be temporary, and as the datasets are analysed and the possible human factors applications better identified and understood, more automated data analysis will be possible. But for initial exploration of the data, the processing is necessarily somewhat manual and time-consuming and may require collaboration with computer science and statistics disciplines.

Nevertheless, the data shows strong potential for generating useful knowledge of both train driver and infrastructure performance and should be further researched. Initial research questions include:

What is the variation between drivers and within drivers?

In order to serve as a useful element of a competence management system, the data should be able to differentiate between drivers who are performing well and those who require improvement. The first step in addressing this is to understand the level of variation within and between drivers, to determine whether it is possible to sensitively discriminate between drivers at any level. If it is not possible to discriminate between drivers, or types of drivers, then the use of the data in performance management is likely limited to the monitoring of rule infractions or specific events, e.g. overspeeds, use of emergency brakes, etc. If it is possible to discriminate between driver profiles, additional parameters may be available from the data. Research would therefore be required to identify what parameters can reliably be used to discriminate between good and poor performance.

What are the key parameters that indicate 'good' performance?

In order to facilitate improved performance, it must be possible to differentiate good performance from poor performance. At present there are no validated metrics that differentiate good drivers from poor drivers, or good infrastructure from poor infrastructure. Existing competence management systems may provide a starting point for driver performance to identify initial metrics for measurement and validation where train drivers are currently routinely assessed by supervisors travelling in cab or reviewing a single downloaded train journey to identify deviation from expected performance. More fundamental research is needed to determine metrics for assessing the infrastructure, but models such as the Route Drivability Tool (Hamilton & Clarke, 2005) may provide some insight. The data may facilitate monitoring beyond simple 'red-line' exceedances (e.g. speed adherence) and allow more detailed analysis of driver variability and resilience in line with the Safety II approach (Hollnagel, 2014). Future research should focus on identifying the metrics with the most potential to reveal insights on driver and/or infrastructure performance.

How to incorporate this data in competence management systems?

Once parameters for driver performance have been identified, a further important question is how to incorporate these into competence management systems in an ethical way. The data is collected at a far more granular, detailed and consistent level that is currently possible in driver management systems, and the prospect of continuous monitoring may not be desirable to drivers and may result in increased levels of stress and worry over relatively minor errors or variations in performance. Employee tracking has not been positively received in other organisations (e.g. Amazon) and there are genuine ethical questions about privacy and the use of this data. However, it is already in effect for similar professions (e.g. professional truck drivers). Finding the right balance of safety and performance monitoring is not a trivial task. One possible solution may be to direct much of the data towards the drivers themselves, as in road driving applications (e.g. Albert et al., 2011), to allow them to directly compare their own performance to an average of their peers rather than receiving feedback from supervisors.

Conclusions

The data available from OTDR appears very promising with potential applications in competence management and safety management. The detail available in the data sets may provide an approach to monitoring and improving railway safety performance, based on day-to-day operations. This paper has presented a framework within which to continue research in exploring this potential, and suggested some possible metrics that may be available. Further research is needed to validate these metrics and propose how they can be used in practice to improve railway safety and performance. Within this, there is a need to consider ethical use of the data in supporting drivers and not in 'big brother' style management.

Acknowledgements

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number 14/IFB/2717. The author additionally wishes to thank Iarnród Éireann - Irish Rail for their support.

References

Albert, G., Musicant, O., Lotan, T., Toledo, T., & Grimberg, E. (2011). Evaluating changes in the driving behaviour of young drivers a few years after licensure using in-vehicle data recorders. In Proceedings of the *6th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 337-343). Iowa: University of Iowa.

Branton, P. (1979). Investigations into the skills of train-driving. *Ergonomics, 22*, 155-164.

Buksh, A., Sharples, S., Wilson, J.R., Coplestone, A., & Morrisroe, G. (2013). A comparative cognitive task analysis of the different forms of driving in the UK rail system. In N. Dadashi, A. Scott, J.R. Wilson, & A. Mills (Eds.) *Rail Human Factors: Supporting Reliability, Safety, and Cost Reduction* (pp. 173-182). London: Taylor & Francis.

Chu, C., Kim, S.K., Lin, Y.A., Yu, Y., Bradski, G., Ng, A., & Olukotun, K. (2007). Map-reduce for machine learning on multicore. *Advances in neural information processing systems, 19*, 281 – 289.

De Fabris, S., Longo, G., & Medeossi, G. (2008). Automated analysis of train event recorder data to improve micro-simulation models. *WIT Transactions on the Built Environment, 103,* 575 – 583.

Doncaster, N. (2012). "By the seat of their pants" Cues and feedback used by train crew. In J.R. Wilson, A. Mills, T. Clarke, J. Rajan, & N. Dadashi (Eds.) *Rail Human Factors around the World: Impacts on and of People for Successful Rail Operations* (pp. 484-494). Boca Raton, USA: CRC Press.

Donmez, B. Ng Boyle, L., & Lee, J.D. (2008). Mitigating driver distraction with retrospect and current feedback. *Accident Analysis and Prevention, 40*, 776-786.

Dorrian, J., Hussey, F., & Dawson, D. (2007). Train driving efficiency and safety: Examining the cost of fatigue. *Journal of Sleep Research, 16*, 1-11.

Dorrian, J., Roach, G.D., Fletcher, A., & Dawson, D. (2006). The effects of fatigue on train handling during speed restrictions. *Transportation Research Part F, 9*, 243 -257.

Dunn, N., & Williamson, A. (2012). Driving monotonous routes in a train simulator: The effect of task demand on driving performance and subjective experience. *Ergonomics, 55*, 997-1008.

Elliot, A.C., Garner, S.D., & Grimes, E. (2007). The cognitive tasks of the driver: The approach and passage through diverging junctions. In J.R. Wilson, B. Norris, T. Clarke, & A. Mills (Eds.) *People and Rail Systems: Human Factors at the Heart of the Railway* (pp. 115-123). Aldershot: Ashgate.

Evans, A.W. (2011). Fatal train accidents on Europe's railways: 1980–2009. *Accident Analysis & Prevention, 43*, 391–401.

Gillis, I. (2007). Cognitive workload of train drivers. In J.R. Wilson, B. Norris, T. Clarke, & A. Mills (Eds.) People and Rail Systems: Human Factors at the Heart of the Railway (pp. 91-101). Aldershot: Ashgate.

Green, S.R., Barkby, S., Puttock, A., & Craggs, R. (2011). Automatically assessing driver performance using black box OTDR data. In Proceedings of the *5th IET Conference on Railway Condition Monitoring and Non-Destructive Testing* (RCM 2011) (pp. 1-5). New York: Institute of Electrical and Electronics Engineers (IEEE).

Hamilton, W.I. & Clarke, T. (2005). Driver performance modelling and its practical application to railway safety. *Applied Ergonomics*, *36*, 661-670.

Hollnagel, E. (2014). Safety-I and Safety-II: The past and future of safety management. Aldershot: Ashgate.

Jenson, M., Wagner, J., & Alexander, K. (2011). Analysis of in-vehicle driver behaviour data for improved safety. *International Journal of Vehicle Safety, 5*, 197 – 212.

Mosimann, C., & Rios, M. (2014). Reduce costs and increase safety by using information generated out of data from black boxes (OTMR, event recorder). In *CORE 2014: Rail Transport for a Vital Economy* (pp. 422-427). Adelaide: Railway Technical Society of Australasia.

Musicant, O., Bar-Gera, H., & Schechtman, E. (2010). Electronic records of undesirable driving events. *Transportation Research Part F, 13*, 71-79.

Naweed, A., & Aitken, J. (2014). Drive a mile in my seat: Signal design from a systems perspective. In Proceedings of *IRSE Australasia Technical Meeting* (pp. 1-7). Victoria: Institution of Railway Signal Engineers Australasia.

Rashidy, E.L., Ahmed, R., & van Gulijk, C. (2016). Driver competence performance indicators using OTMR. In Proceedings of *CIT2016 Congreso de Ingeniería del Transporte* (XII Congress of Transport Engineering) (pp. 354-361). Madrid: Foro de Ingeniería del Transporte.

Strathie, A., & Walker, G. (2015). Can link analysis be applied to identify behavioural patterns in train recorder data? *Human Factors, 58*, 205-217.

Walker, G., & Strathie, A. (2014). Combining human factors methods with transport data recordings. In N. Stanton, S. Landry, G. Di Bucchianic, & A. Vallicelli (Eds.) *Advances in Human Aspects of Transportation: Part 2* (pp. 236 – 245). Louisville: AHFE.

Wilson, J.R., & Norris, B.J. (2005). Rail human factors: Past, present and future. *Applied Ergonomics, 36*, 649-660.