

Open

Original Article

Safe, secure or punctual? A simulator study of train driver response to reports of explosives on a metro train

Kartikeya Tripathi and Hervé Borrión*

Department of Security and Crime Science, University College London, London WC1H 9EZ, UK

E-mail: h.borrión@ucl.ac.uk

*Corresponding author.

Abstract Many transport organizations now regard regular employees as important contributors to their security strategies. For this reason, it is essential to understand the interaction between service and security tasks. There is empirical evidence, for example, that workers under pressure to be punctual make more errors in the performance of safety procedures. It is therefore useful to determine whether punctuality goals have a similar effect on the performance of security procedures. This article reports a study conducted on a metropolitan rail system driving simulator to test whether train drivers also ‘take shortcuts’ in the performance of security procedures when placed under pressure of punctuality. Four hypotheses were tested using performance measures of two groups of 10 participants. The results show that there is a conflict not only between service and security goals, but also between safety and security goals.

Security Journal (2016) **29**, 87–105. doi:10.1057/sj.2015.46

Keywords: human error; railway security; task trade-off; railway simulator; terrorism

The online version of this article is available Open Access

Introduction

Following the derailment of a train that killed four passengers and injured more than 70 others in New York City, the Federal Railroad Administration (2014) carried out a 60-day safety assessment of the operator’s activities. The report submitted to the US Congress identified ‘over-emphasis of on-time performance’ as a key risk factor. This finding is consistent with current research on large and complex organizations, where workers who face a constant competition between following safety rules and working against the clock to fulfil production-related goals take shortcuts in the performance of certain tasks (Pate-Cornell and Murphy, 1996; Zohar, 2002; Zohar and Luria, 2005; Veltri *et al.*, 2013).

The study reported in this article was conducted to determine if these findings were directly applicable to security procedures – security scripts or crime control scripts (Borrión, 2013) – performed by railway employees working under a time constraint. Human error research in security procedures has been limited for two reasons. First, many of the security systems have only been recently employed following a surge in terrorist incidents in the last

few decades (White, 2011). Safety research on the other hand has a long tradition that can be traced back to the 1930s. Second, data from security procedures is not easily available to researchers because it is often sensitive or deemed classified. As a result empirical research on human and organizational factors in security procedures has been piecemeal and focused primarily on security screening operators' skills and cognitive processes in passenger air travel (Kraemer *et al*, 2009).

Of particular interest to this study were the security Standard Operating Procedures (SOPs) that train drivers are asked to implement when a suspicious item is reported on a moving train. The utility of this study was found in the fact that terrorist attacks using explosives form a continuing and serious threat to urban commuter rail networks (Wilson *et al*, 2007; Government Accountability Office, 2010; Jenkins *et al*, 2010; O'Neill *et al*, 2011), and several Metro Rail Systems (MRSs) worldwide rely on ordinary employees who have many other tasks to respond to security threats to the transport network.

Trade-offs between service and safety goals

Trade-offs between safety and production goals can take place for several reasons. Resource constraints and environmental factors of certain working conditions, in particular, make it impractical for workers to meet safety and production goals at the same time. For example, operators may be expected to work faster, but the nature of the task means that they cannot maintain safety standards at higher throughput rates (Marais and Saleh, 2008). Ford and Tetrick (2008) hypothesised that in such circumstances workers can either follow a procedure to fulfil safety goals or maximize production but they cannot do both.

The movement towards taking higher safety risk in working procedures because of increasing pressure from performance or operational goals has been studied by a number of researchers in the field of human error research. McLain (1995, p. 1731) suggested that workers can be safe only when there is slack on the shop floor because 'safety represents an additional task that can affect performance when the sum of required tasks exceeds attention or performance capacity'.

Therefore, in operational terms workers with high production pressures are unlikely to work safely because they start taking shortcuts on certain tasks to free their mental and physical resources to meet production tasks (Veltri *et al*, 2013).

Rasmussen (1999) observed that instructions and written procedures for working safely are almost never followed exactly when operators strive to become more efficient and more productive, and to deal with time and other sources of pressure. Rasmussen (1999, p. 42) noted that 'the stage for an accidental course of events very likely is prepared through time by the normal efforts of many actors in their respective daily work context, responding to the standing request to be cost-effective'.

Rasmussen's concept of movement towards the boundaries of safe behaviour under production pressure was referred to as 'practical drift' by Snook (2002). Hollnagel coined the metaphor 'Efficiency-Thoroughness Trade-Off' to explain that actors have to balance efficiency requirements of production goals and thoroughness in following safety procedures (Hollnagel, 2004, 2012; Rosness *et al*, 2012). In his noted work on human reliability, Hollnagel (2004) observed that operators genuinely try to do what is expected of them in following safety procedures, but at the same time they want to do it without what they perceive as unnecessary effort, or waste of time.



Workers also make trade-offs between safety and production goals in socio-technical fields when the work procedure gives them freedom to choose the tasks they perform, the order in which to perform them, and to put certain tasks on hold in order to complete others. Rasmussen (1999) explained that work situations give workers the flexibility to meet their objectives in multiple ways, and they use their subjective, situation-dependent judgment to choose the path of least resistance.

Therefore as part of this study, a questionnaire based on an 11-point Likert scale was added to reflect the current philosophy in human error research that an in-depth analysis of practitioners' own interpretation of tasks and errors is needed to comprehensively understand error in dynamic work procedures (Dekker and Pruchnicki, 2013). This is especially relevant when workers can make choices about which tasks they want to perform, in which order they want to perform them, and put certain tasks on hold for indefinite periods of time without affecting the basic integrity of the work process. In a widely cited study, Hollnagel and Amalberti (2001) asked observers to count human error in air traffic control (ATC) operations using a previously developed taxonomy. They found that a number of actions that were counted as errors by observers were not recorded as errors by ATC operators. The latter saw several departures from the taxonomy as deliberate strategies to manage an unforeseen problem during ATC operations.

Conceptual model: Safety, security and services

In order to assess whether the above findings apply to security procedures, it is essential to understand the place of security in the safety-versus-production models introduced above.

Infrastructures and other large urban systems are generally developed to provide users with extra capability. The primary goal of a system can therefore be regarded as the provision of the sought capability through the delivery of core business operations, for example, service tasks such as driving the train in the case of a MRS.

Operational procedures are designed with a number of assumptions about the characteristics of the transport system (including the infrastructure and staff), its users and its environment. In order to maintain these characteristics within their acceptable operational range, MRSs implement additional security and safety procedures.

From this point of view, safety and security goals can be regarded as secondary goals that are both concerned with protection of valuable elements from harm (Piètre-Cambacédès and Bouissou, 2013; Raspotnig and Opdahl, 2013). In MRSs the objectives of the stakeholders include the state of the transport system (for example, the equipment) and the employees, the state of those who (may) interact with it (for example, passengers and their property), and the effect of the system on its environment.

Safety and security measures are all aimed at reducing the likelihood and severity of harm (Raspotnig and Opdahl, 2013). From the employees' perspective, the main categories of measures are:

- Those aimed at controlling the state of the system using direct means (for example, replacing faulty parts with more reliable ones, hardening train carriages)
- Those aimed at controlling the state of the system indirectly by influencing the way employees operate it (for example, reducing the maximum speed limit when the environmental conditions are poor)

- Those aimed at controlling the influence of external components (for example, removing leaves on tracks, removing suspicious unattended bags onboard).

All three categories can include measures designed to reduce the probability of an unwanted event threatening the system, and the vulnerability of the system to this event. However, the essential difference between safety and security is that safety deals with *accidental* harm, while security deals with *malicious* harm. The adverse consequences of both safety and security incidents may be the same, but what distinguishes them is the absence or presence of malicious intent.

This difference has several implications. First, the likelihood of a malicious actor threatening the system is generally very low, whereas systemic and environmental factors of safety risks are everywhere. Second, the causal relationship between service tasks and safety risk is relatively evident. Safety becomes a concern as soon as employees operate the system, and safety risk generally increases as the level of performance (for example, velocity) increases (Veltri *et al*, 2013). In the contrary, the causal influence of service tasks on security risk is less obvious. Service tasks seem to have a limited impact on the security threats to MRSs, and their vulnerability to those. Furthermore, staff training focuses a lot more on the relationship between service tasks and safety than service tasks and security. For this reason, civilian staff may consider their involvement in security procedures as a separate role, disconnected from their main function in the organisation (Jenkins *et al*, 2010).

One similarity between safety and security that is relevant to the present study is that both are perceived as ‘eternal killjoys’ by workers where ‘security and safety share a common burden: they both imply guarding against high consequence events which are essentially negative (attacks, accidents) in contrast to desired outcomes, which are essentially positive (services delivered, goods produced)’ (Piètre-Cambacédès and Bouissou, 2013, p. 114).

In a MRS, employees at the sharp end of operations are expected to perform a range of tasks to deliver the core services of their organisation (for example, driving the train). We know from studies presented in the paper that staff take shortcuts in completing certain tasks when they work under time constraints. In the same vein as McLain’s (1995) work, a security procedure could be considered as an additional set of tasks that can affect performance when the sum of required tasks exceeds attention or performance capacity. This led to our first hypothesis:

Hypothesis 1 (H1): Drivers will commit errors in the conduct of security SOPs.

Inspired by the opening example, this study was conducted to understand how on-time performance might affect the performance of security procedures. The main question investigated was whether pressure of punctuality has a similar effect on performance of security procedures by ordinary employees of a MRS. This led to our second hypothesis:

Hypothesis 2 (H2): Emphasizing on-time performance will increase the number of errors made by drivers in the conduct of security SOPs.

Finally, we investigated whether the errors made by train drivers could be attributed to rational decisions. This led to two more hypotheses, Hypothesis 3 and Hypothesis 4, under

the assumption that security was considered as a secondary goal in comparison with service goals such as punctuality:

Hypothesis 3 (H3): Emphasizing on-time performance will increase the number of errors made by drivers in the conduct of the SOP tasks that hinder punctuality.

Hypothesis 4 (H4): Emphasizing on-time performance will reduce the number of errors made by drivers in the conduct of the SOP tasks that support punctuality.

There are several reasons why these hypotheses could be incorrect. In spite of many similarities, it could be inappropriate to consider security and safety as conceptually equivalent: malicious elements are far less likely to be present in MRSs than internal and external factors that could contribute to unintended failures and accidents; and service tasks have a lot less influence on security risk than on safety risk. However, these differences are limited in the particular case investigated in this study, as the latter concerns high-risk situations following the detection of an unattended bag onboard.

Context of the research

A MRS network in a large metropolitan city in South Asia cooperated in this study as part of a wider, long-term research into security procedures on MRS networks. Railway companies rely on front line railway staff – or ‘civilian staff’ as coined by Jenkins and Gersten (2001) – to carry out security procedures for two reasons. First, its members usually receive initial reports of suspicious activities from passengers on the network (Jenkins and Gersten, 2001). Second, it is a cost effective way for them to respond to a large number of reports of threats every year (Hartong *et al*, 2008).

In the network where the study took place, train drivers are expected to follow a specific security SOP whenever a passenger reports a suspicious item on the train. This 20-step procedure was designed for train drivers to assist the management of the item (usually its removal) through coordination with members of the Operations Control Centre (OCC) and the station team. The SOP was designed to ensure that suspicious items are dealt with in an acceptable time under operational circumstances. The drivers underwent classroom training to memorize the SOP and took written exams where the SOP was part of the coursework. Finally, they repeatedly rehearsed this procedure in a train driving simulator along with other technical, and safety SOPs.

Table 1 lists the 20 steps of the SOP, the description of the actions, and the rationale for their inclusion. The procedure is initiated after a passenger reports a suspicious item on the train to the driver using the passenger emergency communication unit (PECU) available in each coach. The driver then gathers specific pieces of information about the item from the passenger, and passes it on to the OCC. The latter conveys the same information to the members of the security team present at the next station. With this information, the security team can quickly locate the item and remove it from the train. If they feel the risk is too great, they ask the bomb disposal squad to deal with it.

Train punctuality is an important operational goal that train drivers are required to satisfy in their conventional role (Olsson and Haugland, 2004). If trains are delayed on the MRS then the whole network could be thrown out of schedule, which in turn would have negative commercial consequences for the railway company (Borrion *et al*, 2014). Under the train

Table 1: Details of actions and their purpose in the 20-step security procedure

<i>Task number</i>	<i>Action</i>	<i>Description of action</i>	<i>Purpose of action</i>
T1	Answer call	Driver receives call from passenger	Driver becomes aware that there is a suspicious item on the train and can initiate the procedure for its removal
T2	Ask coach	Driver asks passenger for the coach number in which the suspicious item is located	Driver has information that will assist station team to position itself at the right place on the platform to enter the coach with suspicious item
T3	Ask Position	Driver asks passenger where is the suspicious item located in the coach (in front, middle or rear of coach)	Driver has information that will assist station team to locate the suspicious item in the coach
T4	Ask description	Driver asks passenger for description of the suspicious item (for example, its size, colour, whether it is a bag or a box)	Driver has information that will assist station team to identify the suspicious item in the coach
T5	Make manual passenger announcement	Driver makes an announcement requesting passengers to move away from the suspicious item	Minimise chances of passengers triggering an explosive device by touching it, and minimise casualties in case it is detonated
T6	Keep PECU on 'Suspend' mode	Driver does not end his call with the passenger, and keeps communication unit in a suspended (on hold) mode.	Driver has the option of calling the passenger back to request more information about the suspicious item
T7	Call OCC	Driver calls the OCC to inform them that a suspicious item has been reported on the train	OCC becomes aware that there is a suspicious item on the train and can initiate the procedure for its removal
T8	Give coach to OCC	Driver gives OCC the coach number in which the suspicious item is located	OCC has information that will assist station team to position itself at the right place on the platform to enter the coach with suspicious item
T9	Give position to OCC	Driver gives OCC the location of the suspicious item in the coach (in front, middle or rear of coach)	OCC has information that will assist station team to locate the suspicious item in the coach
T10	Give description to OCC	Driver gives OCC description of the suspicious item (for example, its size, colour, whether it is a bag or box)	OCC has information that will assist station team to identify the suspicious item in the coach
T11	Request instructions from OCC	Driver requests instructions from OCC	Upon request from driver for further instructions, OCC informs driver that procedure to remove suspicious item will be coordinated with station team

T12	Follow new 'Target Speed' to station	Driver increases the speed of the train to a maximum of 80 km/h	Time for the train to reach the next station is minimized, so that the procedure to remove suspicious item at the station can start
T13	Call passenger back on PECU	Driver uses the PECU kept on suspend mode to speak to the passenger who reported the suspicious item	Driver can ask the passenger for any additional information about the suspicious item (for example, have the passengers moved away from the suspicious item?)
T14	Give new information to OCC	Driver gives additional information about the suspicious item to OCC	OCC has additional information that it can communicate to station team
T15	Stop at station (no overshoot/undershoot)	Driver stops the train at the station without overshooting or undershooting the stop point	Time is not wasted in stopping the train as train does not have to be reversed or taken forward to the stop point after coming to an initial halt.
T16	Open train doors	Driver opens the doors of the train	Station team can enter the coach with suspicious item, and passengers can move away from the suspicious item out of the train
T17	Make automatic passenger announcement	Driver makes a pre-recorded passenger announcement informing passengers that the station team will enter the train	Passengers are reassured and make way for the station team to enter the train
T18	Not close doors	Driver does not close the doors of the train	Station team can remove the suspicious item from the train
T19	Wait at station despite green signal and target speed	Driver does not move the train from the station despite the OCC giving an automated signal to do so	Train stays at the station while station team is locating and removing the suspicious item
T20	Move train on 'All Clear' from station team	Driver moves the train on getting a signal from the station team	Driver is aware that suspicious item has been removed by the station team and procedure for its removal is over

services agreement in place, the train drivers could face disciplinary action if they regularly fail to drive the train according to their schedule. In practice, train drivers could be asked for a written explanation if they drive the train behind schedule more than three times a month. The importance of this goal is also reflected in the presence of potential financial sanctions; drivers' managers could even lose a part of their monthly income in extreme cases.

The following four hypotheses were tested:

- H1:** Drivers will commit errors in the conduct of security SOPs.
- H2:** Emphasizing on-time performance will increase the number of errors made by drivers in the conduct of security SOPs.
- H3:** Emphasizing on-time performance will increase the number of errors made by drivers in the conduct of the SOP tasks that hinder punctuality.
- H4:** Emphasizing on-time performance will reduce the number of errors made by drivers in the conduct of the SOP tasks that support punctuality.

Method

Participants

Twenty male train drivers participated in the study (Range: 23–31 years, Median = 25, Mean = 25.25, SD = 2.29). They were assigned to two groups of 10 train drivers: the control group for whom on-time performance was not emphasised, and the treatment group for whom it was. At the time of the study the participants had undergone 12-month classroom training in driving MRS trains, and had spent another 6-month training on a simulator to gain train driving experience. (Range: 40–45 hours, Median = 42 hours, Mean = 42.25, SD = 1.72). Both during their classroom training and while driving a simulated train, the participants rehearsed various operations related to SOPs, including the procedure corresponding to the report of a suspicious item on the train.

The participants were offered no incentive to take part in the study. Their participation was entirely voluntary, based on informed consent, and the participants were told that they could withdraw from the study at any stage.

The test was similar in many ways to what they had been through during their training. Just like in their normal training schedule, the time they spent driving the simulated train for the study was also added to their professional track record. The study was conducted over 5 days. Each day four drivers recorded their performance, two in the morning and two in the afternoon.

Apparatus

All testing was carried out using a simulator at the driver training centre of the participating MRS. Hi-fidelity simulator studies enjoy several benefits over real-world studies. They are economical, versatile and have ethical advantages over real-world investigations by providing an inherently safe environment for participants (Jamson and Jamson, 2010).



Simulator studies have been used for experimental purposes in medicine (McFetrich, 2006; Bogenstätter *et al*, 2009), aviation (Nikolic and Sarter, 2007), road driving (Liu *et al*, 2009; Aksan *et al*, 2012) and railways (Dorrian *et al*, 2007; Tey *et al*, 2011).

The hi-fidelity train driving simulator used in this study had a simulator cab that was an exact replica of the train driver's cabin, with the same controls as in the MRS trains, for example. An in-built audio system generated sounds of rail-wheel interaction, train horn and other elements from the working environment.

The simulator cabin included a 3×3 m screen onto which track footage was projected. The software package used in the simulator replicated route length, number of stations and distance between stations according to the layout of the real network. The virtual train was a four coach commuter train.

The scoring sheet was in form of a checklist. The observer used it to record whether the participant performed (or did not perform) each individual task of the 20-step security procedure. The sheet also noted whether the participant was in the control group or the experimental group.

1. The performance of the participants was recorded on the scoring sheet represented in Figure 1.
2. The perception of participants on the importance of the security procedure to remove suspicious items from trains, and the likelihood that a reported item was effectively harmful was recorded after they finished driving the simulated train by means of a questionnaire and interview.

The scale followed a standard presentation format used to record responses of practitioners in health sciences (Carifio and Perla, 2007; Norman, 2010). An 11-point scale with no midpoint for neutral response was used to increase the sensitivity of the scale (Cummins and Gullone, 2000), and to reduce social desirability bias in responses of the participants (Garland, 1991).

Research design

The first hypothesis (H1) was tested by measuring the total number of errors made by the participants in the performance of the 20-step SOP for two levels of on-time pressure on employees. The manipulated condition (independent variable) was the presence or absence of a verbal message emphasising on-time punctuality. The other three hypotheses (H2, H3, H4) were tested using a 2×2 within-subjects experimental design. The manipulated condition (independent variable) was also the presence or absence of a verbal message emphasising on-time punctuality. For H2, the dependent variable was the total number of errors made by participants in the performance of the 20-step SOP. For H3, the dependent variable was the number of errors made in the conduct of Task 19. In the SOP, Task 19 appears to be the only task that directly hinders train punctuality; it requires train drivers to wait at the station after the traditional green light, until they receive the all clear signal from the station team. For H4, the dependent variable was the number of errors made in the conduct of Task 12. In the SOP, only Task 12 appears to be directly supporting train punctuality. It requires participants to increase the speed of the train from 35 km/h to 80 km/h, which would, in principle, help drivers respect their schedule.

Driver Number:		Group: Experimental/Control
T1	Receive passenger call on PECU	
T2	Check which coach number the passenger is in	
T3	Check where in coach the suspicious object (SO) is – Front, Middle or Rear	
T4	Get description of SO from passenger (bag? suitcase? size? colour? etc.)	
T5	Make passenger announcement 1	
T6	Keep PECU on suspend mode to call passenger back	
T7	Call OCC to report SO	
T8	Give coach number to OCC	
T9	Give position of SO in coach to OCC	
T10	Give description of SO to OCC	
T11	Get instructions from OCC	
T12	Follow new 'Target Speed' to next station	
T13	Call passenger back on PACU to get further details	
T14	Communicate new details to OCC	
T15	Stop at station – not overshoot or undershoot train	
T16	Open doors	
T17	Make Passenger announcement 2	
T18	Do not close doors	
T19	Wait at station despite getting signal and 'Target Speed'	
T20	Wait for 'all clear' from station team	

Figure 1: Checklist used for scoring train driver's performance on a 20-step procedure.

Procedure

One male train driver instructor with 10 years of experience in driving urban commuter trains acted as observer, and recorded the performance of the participants using the aforementioned scoring sheet during the experiment.

Employing peer train drivers as observers is considered to have a positive effect on participants in experiments as they feel more at ease when their performance is being recorded by someone who understands the demands of the job. In order to have uniformity in scoring, practice sessions were conducted with the observer and three train drivers who did not subsequently participate in the experiment. Any ambiguity in the mind of the observer over the scoring system was cleared during this pilot phase.

The scoring sheets used by the observer had no identifying details of the participants, and the observer assured them that their individual performance scores or characteristics would not be shared with the management. After the study the researchers took possession of all the scoring sheets for data analysis purposes and to ensure that the results of the study could not be used for other purposes.

A colleague acted as a passenger to report the presence of a suspicious item on the train to the participants. Figure 2 shows the script used to report the suspicious item in the study.

One of the researchers sat next to the observer, and recorded the time spent by each participant in talking to the colleague acting as a passenger, and the time they spent making the first passenger announcement (Task 5).

All participants were asked to drive the simulated train on a round trip of the MRS line.

The members of the experimental group were told that they were being tested for driving the train on schedule and were required to complete the journey within a stipulated time of 46 min. This corresponded to the expected journey time for the route on the MRS.

It must be noted that drivers in both groups had been through the same training, and were therefore familiar with this objective. This was designed to replicate the pressure in actual working conditions. However, in the experiment only the participants in the experimental group received these instructions that emphasised on-time performance.

The driving scenario in the simulator was kept uniform for all participants in both groups. They drove the train in conditions of good visibility, with no rain or obstructions on the track, and no technical faults.

Passenger (P) (if Train Driver answers call): 'Sir there is an item lying here. I do not know whose it is. Please do something about it.'

P (if TD asks for Coach): 'Coach number 3'

P (if TD asks for Position): 'It is lying in the middle of the coach'

P (if TD asks for Description): 'It is a blue bag'

P (if TD calls back for additional information): 'We have moved away from the bag'

(Translated from local vernacular language)

Figure 2: Script given to passenger for informing train driver of suspicious item on the train.

When the train left the third station on the virtual route, the colleague acting as a train passenger called the participant using the PECU to report a suspicious item on the train. This specific location, about 10 min from departure, was selected because the distance between the third and fourth stations is the longest between two consecutive stations on the route. In this way, the distance that had to be covered while implementing the SOP was the same for all participants. The observer started recording the performance of each participant on a scoring sheet after the call was made.

In keeping with industry standards, every participant was debriefed about his performance, and informed about the errors he made. The participant was then invited to complete the perception questionnaire, and to give his opinion about the causes of any errors. In addition, the observer asked him not to divulge details of the experiment to his colleagues as this could interfere with the study (as researchers would not be able to record natural performance of other drivers).

Results

Error rates

All 20 participants completed the test. Two drivers in the control group made no errors. All others made at least three errors. The largest number of errors (12) was made by a participant in the experimental group. Out of a maximum possible of 400 errors (20 tasks×20 drivers), the participants made 129 errors. This represents a rate of one error every three tasks.

In total, the 10 participants in the control group made 36 errors (18 per cent) in performing the 20 tasks of the SOP. Those in the experimental group made 93 out of a maximum possible of 200 errors (46 per cent) (Table 2).

In total, none of the participants in the control group made an error in the conduct of Task 19. In comparison, four participants (40 per cent) in the experimental group made an error in this task.

In total, five participants (50 per cent) in the control group made an error in the conduct of Task 12. In comparison, all 10 participants (100 per cent) in the experimental group made an error in this task.

H1: Drivers will commit errors in the conduct of security SOPs.

The 20 participating train drivers in the study committed 129 out of maximum possible 400 errors.

H2: Emphasizing on-time performance will increase the number of errors made by drivers in the conduct of security SOPs.

A two by two contingency table was constructed. The number of errors committed in the SOP was significantly larger for the experimental group (46 versus 18 per cent, $P=0.001$, Fisher's exact test, one tailed).

H3: Emphasizing on-time performance will increase the number of errors made by drivers in the conduct of the SOP tasks that hinder punctuality.

Table 2: Number of errors made by train drivers in control group and experimental group on each task in security procedure

<i>Task</i>	<i>Control group</i>	<i>Experimental group</i>
T1	0	0
T2	0	2
T3	4	8
T4	4	8
T5	0	3
T6	5	10
T7	0	0
T8	0	2
T9	4	8
T10	4	8
T11	0	0
T12	5	10
T13	5	10
T14	5	10
T15	0	2
T16	0	0
T17	0	4
T18	0	0
T19	0	4
T20	0	4
Total	36	93

A two by two contingency table was constructed. It was not found that the number of errors committed in the SOP was significantly larger for the experimental group (40 versus 0 per cent, $P=0.086$, Fisher's exact test, one tailed).

H4: Emphasizing on-time performance will reduce the number of errors made by drivers in the conduct of the SOP tasks that support punctuality.

A two by two contingency table was constructed. The number of errors committed in the SOP was significantly larger for the experimental group (100 versus 50 per cent, $P=0.032$, Fisher's exact test, one tailed).

Discussion

Summary of the results

The results of both the experimental and control groups show that errors are made in the performance of security procedures, which confirms that security procedures are not immune to shortcuts by employees (Hypothesis 1). The difference between the error rates of the two groups also support Hypothesis 2 that emphasising on-time performance increases the likelihood that drivers make errors in the conduct of security SOP. These results suggest that the effectiveness of transport security procedures could be jeopardised if those implementing them are placed under too much on-time pressure. Regarding Hypothesis 3, the error rate of

the experimental group is substantial: 40 points above that of the control group. However, the result is not statistically significant, possibly because of the small sample size. A sensitivity analysis revealed that one more error in the experimental condition would have been enough to make the result statistically significant.

Finally, a statistically significant difference was found between the two groups in performance of Task 12 but in the direction opposite to the one that Hypothesis 4 predicted. The original hypothesis stated that the experimental group under pressure of punctuality would make fewer errors in Task 12 than the control group. However, the results of the study showed that the opposite happened – the experimental group made more errors than the control group.

Service, safety and security

Hypothesis 4 was not validated, and the test results for performance of Task 12 were unexpected. The task required train drivers to increase the speed of the train from 35 km/h to 80 km/h. Seemingly, drivers under pressure of maintaining punctuality should have made fewer errors than the control group in this task as it encouraged them to drive the train faster, which supported their objective of punctuality. Moreover, the task also supported the organisation's security goals by reducing the time during which passengers were in the vicinity of a potential explosive device. Given the synergy of Task 12 with both service and security goals, the authors expected very few errors to be made in this task. In the contrary, none of the drivers in the experimental group increased the speed of the train in the simulator. It was found that drivers placed under greater pressure of punctuality were twice more likely not to increase the train speed on this task. The possible reasons were found in post-experiment interviews, and questionnaire responses.

In the interviews drivers from both groups said they were worried that they would not be able to stop the train properly, and overshoot the platform at the station if they increased the speed. They were also concerned that increasing the speed of the train changed the calculations of braking, and overall handling of the train which meant that the train could derail. Task 12 could therefore not be regarded as a task that unconditionally supports punctuality (service goals) since overshooting would disrupt the MRS. Furthermore, Task 12 also affected safety a lot more than initially envisaged.

This study was conducted to examine whether the findings about the conflict between safety and service goals also apply to security procedures. It is apparent that the problem could not be simplified to a conflict between security and service, but that safety goals had to be considered too, and that conflict between safety and security goals in dynamic work environments (Nikolic and Sarter, 2007) should also be examined.

With this in mind, three possible explanations may be invoked to explain the difference between the results of the two groups. First, drivers placed under greater on-time pressure felt their ability to control the train was diminished to such a point that it was preferable not to take the risk of carrying out this task. Second, they felt their ability was the same but became more risk averse under pressure. Third, the workload that would be additionally introduced by driving the train in such conditions was not welcome. Drivers under too much pressure may have preferred to implement a familiar task rather than a new and possibly more challenging one.



It should also be noted that security SOPs are often performed with no certainty about the reality of the threat. With hindsight, many procedures are unnecessarily performed because of the large number of false alarms (for example, very few unattended bags conceal an explosive device). This thinking was reflected in drivers' responses to the perception questionnaire, where all the participants rated the probability of a false alarm between 'high' and 'extremely high'. Belief in credibility of a threat has been known to influence the commitment of employees to follow procedures to mitigate it. Therefore, the perception of the drivers that the reported suspicious object was harmless forms a part of the explanation for their performance on the security SOP. They may have prioritized the mitigation of what they felt was a more likely risk from increasing the speed of the train over the remote possibility that there was an actual explosive on the train.

Limitations of the study

The primary limitation of this study relates to the use of a simulator environment for data collection. For ethical reasons, this study was initially conducted in a simulator as putting drivers under a false sense of urgency to be punctual in an actual train driving environment where a suspicious item had been reported on the train, would have posed safety and security risks both for the train drivers and passengers.

The validity of the results obtained in a simulator is always arguable since the complexity of the real world can never be replicated in its entirety. In order to artificially recreate a realistic driving environment, a simulator-based experiment must have emotional validity, physical validity, face validity, perceptual validity and behavioural validity (Jamson and Jamson, 2010). While the simulation environment looks very convincing (from the researcher's perspective at least), it may not be sufficiently realistic to generate the same level of stress and motivational pressure (Neale and Liebert, 1986) that would exist with a real-explosive threat. This may explain the relatively high error rate observed with both groups.

Nevertheless, it must be pointed though that even in real-life conditions only a fraction of suspicious items that are reported constitute a genuine threat. A majority of reports are about innocuous items forgotten by passengers on the train (Riley, 2004; Peterman, 2006). Therefore, even in real-life conditions the report of a suspicious item by a passenger may not cause an extremely high level of stress in the mind of the driver. For this reason, it is possible that the simulator adopted for this study did not significantly affect the ecological validity of the findings.

The small sample size of 20 train drivers was also a limitation of the study. This limited number was because of a decision to prioritize ecological validity. In this study, the selected participants were all individuals who were training to drive MRS trains in a context where explosives posed a serious and continuous threat to the railway network, and had undergone a special training to respond to reports of suspicious items from passengers on the train.

Studies with small sample sizes are, however, not uncommon in this field as the specialist nature of participants means that a large population is often not available.

Dorrian *et al* (2006, 2007) conducted two simulator studies to test the effect of fatigue on train drivers and in both studies the sample also consisted of 20 male train drivers. Jamson and Jamson (2010) recruited 18 drivers to test the validity of a low-cost driving simulator.

Earl *et al* (2012) developed a single-pilot line operations safety audit by collecting flight data of 14 pilots from a mid-sized aviation company. Huttunen *et al* (2011) studied the effect of mental workload with a sample size of 13 military pilots who performed tasks in a flight simulator. Nikolic and Sarter (2007) relied on a sample of 12 Boeing 747–400 pilots to study their response to flight disturbance in a simulator.

Another limitation of this study relates to the instrument. The procedure that served as a basis to measure error rate comprises 20 tasks. These were not designed for research purposes, and the experimental design was therefore inherently limited compared with a factorial design. The effects of performing (or not performing) the tasks vary across tasks. For example, performing Task 2 had a limited effect on safety compared with Task 12. Moreover, each task may have multiple effects. For example, Task 12 contributes to the organisation's security goals but could conflict with its safety goals. In addition, a given task may both contribute to or conflict with a given goal depending on the conditions and the way the task is performed. For example, increasing the speed of the train may increase the likelihood of being punctual unless the conditions are such that the driver is likely to overshoot the train, in which case the service will be delayed. As a result, the drivers' accounts elicited through the post-test interviews were used to make a number of inferences, and the validity of the findings are therefore limited by the degree of completeness and reliability of these accounts.

The fact that all drivers were new recruits and did not have a lot of real-world experience of driving trains is also a valid criticism of the study. The strategy to address the conflict between operational and security tasks could change in actual working conditions. The drivers could make a different set of mistakes in the SOP under pressure from specific strains of driving a real train, than the ones they made in the simulator without any practical experience.

Implications for railway security

Dörner (1989, p. 65) observed that 'contradictory goals are the rule, not the exception in complex organisations'. In such contexts managing conflicts between multiple goals is typically transferred to local operating units that resolve them in the form of daily decisions and trade-offs (Dekker and Pruchnicki, 2013). MRSs are dynamic transportation systems that cater to diverse needs of millions of passengers every day (Dwyer, 2010). Ordinary railway employees are expected to meet a number of service goals (Borrion *et al*, 2014) of which security is only one.

This research provides an initial investigation into the trade-offs made by MRS train drivers that has implications for the design and management of security procedures.

The study confirmed that civilian staff working under pressure are likely to make shortcuts in the performance of security procedures. Moreover, it was shown that as work pressure increases, the error rate was likely to increase as a result of trade-offs made by employees. It is therefore recommended that designers of security procedures consider background levels of pressure that civilian staff face as a result of their other responsibilities (Borrion *et al*, 2014).

Railway companies and researchers need to understand that not all tasks can be easily classified as service, security or safety tasks. For example, driving a train is a service task, but the introduction of a maximum speed in an SOP to deal with a suspicious object on the train



is motivated by security considerations. As a result, security procedures should be analysed by considering the tasks and their contribution to all important goals, but also their mutual dependencies and the conditions in which they are performed.

Performance of a task by front line staff depends on their perception of the task's contribution to meeting service, safety and security goals, and the priority that they give to those. In the case of safety and security, it would be expected that the degree of risks (in particular service, safety and security risks), as perceived by front line staff, has a significant effect on the performance of security procedures. This study has shown that this leads to a gap between the design of security procedures, the expectations the management team places upon front line staff, and how the procedures are conducted by the staff itself. Further studies will be needed to examine the nuances of conflict between safety, security and service goals, and how this conflict is played out in actual train driving conditions.

Conclusion

This study has shown that findings on conflict between service and safety goals are transferrable to security procedures in a MRS environment. A novel finding of the study was that researchers need to consider not just the conflict between service and safety goals, but also between safety and security goals. This adds another layer of complexity to understand trade-offs made by front line staff in meeting irreconcilable goals through practical strategies (Dekker and Pruchnicki, 2013), but should provide additional information to reduce operational risks.

Acknowledgements

This work supported by the Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/G037264/1]. The authors are grateful to the MRS organisation that provided access to their infrastructure, and allowed data collection to be performed. They are also indebted to the drivers and instructor who participated in this study. Finally, the authors thank the reviewers whose comments and suggestions greatly contributed to the quality of their manuscript.

References

- Aksan, N. *et al* (2012) Naturalistic distraction and driving safety in older drivers. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 55(4): 841–885.
- Bogenstätter, Y., Tschan, F., Semmer, N.K., Spychiger, M., Breuer, M. and Marsch, S. (2009) How accurate is information transmitted to medical professionals joining a medical emergency? A simulator study. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51(2): 115–125.
- Borrion, H. (2013) Quality assurance in crime scripting. *Crime Science* 2(1): 6.
- Borrion, H., Tripathi, K., Chen, P. and Moon, S. (2014) Threat detection: A framework for security architects and designers of metropolitan rail systems. *Urban Planning and Transport Research* 2(1): 173–194.
- Carifio, J. and Perla, R.J. (2007) Ten common misunderstandings, misconceptions, persistent myths and urban legends about likert scales and likert response formats and their antidotes. *Journal of Social Sciences* 3(3): 106.

- Cummins, R.A. and Gullone, E. (2000) Why We Should Not Use 5-Point Likert scales: The Case for Subjective Quality of Life Measurement. *Proceedings of the Second International Conference on Quality of Life in Cities*. Singapore: National University of Singapore, pp. 74–93.
- Dekker, S. and Pruchnicki, S. (2013) Drifting into failure: Theorising the dynamics of disaster incubation. *Theoretical Issues in Ergonomics Science* 15(6): 1–11.
- Dörner, D. (1989) *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations*. Cambridge, MA: Perseus Books.
- Dorrian, J., Roach, G.D., Fletcher, A. and Dawson, D. (2006) The effects of fatigue on train handling during speed restrictions. *Transportation Research Part F: Traffic Psychology and Behaviour* 9(4): 243–257.
- Dorrian, J., Roach, G.D., Fletcher, A. and Dawson, D. (2007) Simulated train driving: Fatigue, self-awareness and cognitive disengagement. *Applied Ergonomics* 38(2): 155–166.
- Dwyer, A.S. (2010) Intending to Travel. Benefit, Cost, and Active Management of Terrorism-related Risk in the Railway Environment. British Transport Police. Unpublished.
- Earl, L., Bates, P.R., Murray, P.S., Ian Glendon, A. and Creed, P.A. (2012) Developing a single-pilot line operations safety audit. *Aviation Psychology and Applied Human Factors* 2(2): 49–61.
- Federal Railroad Administration. (2014) Operation deep dive metro-north commuter railroad safety assessment, <http://www.fra.dot.gov/Elib/Document/3586>, accessed 16 September 2014.
- Ford, M.T. and Tetrick, L.E. (2008) Safety motivation and human resource management in North America. *The International Journal of Human Resource Management* 19(8): 1472–1485.
- Garland, R. (1991) The mid-point on a rating scale: Is it desirable. *Marketing Bulletin* 2(1): 66–70.
- Government Accountability Office. (2010) Technology assessment: Explosives detection technology to protect passenger rail. United States Government Accountability Office, <http://www.gao.gov/new.items/d10898.pdf>.
- Hartong, M., Goel, R. and Wijesekera, D. (2008) Security and the US rail infrastructure. *International Journal of Critical Infrastructure Protection* 1: 15–28.
- Hollnagel, E. and Amalberti, R. (2001) *The Emperor's New Clothes, or Whatever Happened to 'human error'?* The Fourth International Workshop on Human Error, Safety and System Development, 11 June, Linköping, Sweden.
- Hollnagel, E. (2004) *Barriers and Accident Prevention*. Aldershot, UK: Ashgate.
- Hollnagel, E. (2012) *The ETTO Principle: Efficiency-Thoroughness Trade-off: Why Things That Go Right Sometimes Go Wrong*. Aldershot, UK: Ashgate.
- Huttunen, K., Keränen, H., Väyrynen, E., Pääkkönen, R. and Leino, T. (2011) Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. *Applied Ergonomics* 42(2): 348–357.
- Jamson, S.L. and Jamson, A.H. (2010) The validity of a low-cost simulator for the assessment of the effects of in-vehicle information systems. *Safety Science* 48(10): 1477–1483.
- Jenkins, B. and Gersten, L. (2001) *Protecting Public Surface Transportation Against Terrorism and Serious Crime: Continuing Research of Best Security Practices*. San Jose, CA: Mineta Transportation Institute.
- Jenkins, B.M., Butterworth, B.R. and Gerston, L.N. (2010) Supplement to MTI study on selective passenger screening in the mass transit rail environment. San Jose, CA: Mineta Transportation Institute. <http://www.transweb.sjsu.edu/MTIportal/research/publications/documents/Selective%20Screening%20Supplement%20%28with%20Covers%29.pdf>, accessed 15 July 2015.
- Kraemer, S., Carayon, P. and Sanquist, T.F. (2009) Human and organizational factors in security screening and inspection systems: Conceptual framework and key research needs. *Cognition, Technology & Work* 11(1): 29–41.
- Liu, C.C., Hosking, S.G. and Lenné, M.G. (2009) Predicting driver drowsiness using vehicle measures: Recent insights and future challenges. *Journal of Safety Research* 40(4): 239–245.
- Marais, K.B. and Saleh, J.H. (2008) Conceptualizing and communicating organizational risk dynamics in the thoroughness – efficiency space. *Reliability Engineering System Safety* 93(11): 1710–1719.
- McFetrich, J. (2006) A structured literature review on the use of high fidelity patient simulators for teaching in emergency medicine. *Emergency Medicine Journal* 23(7): 509–511.
- McLain, D.L. (1995) Responses to health and safety risk in the work environment. *Academy of Management Journal* 38(6): 1726–1743.
- Neale, J.M. and Liebert, R.M. (1986) *Science and Behavior: An Introduction to Methods of Research*. Englewood Cliffs, NJ: Prentice-Hall.
- Nikolic, M.I. and Sarter, N.B. (2007) Flight deck disturbance management: A simulator study of diagnosis and recovery from breakdowns in pilot-automation coordination. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49(4): 553–563.



- Norman, G. (2010) Likert scales, levels of measurement and the 'laws' of statistics. *Advances in Health Sciences Education* 15(5): 625–632.
- Olsson, N.O. and Haugland, H. (2004) Influencing factors on train punctuality – Results from some Norwegian studies. *Transport Policy* 11(4): 387–397.
- O'Neill, C., Robinson, A.M. and Ingleton, S. (2011) Mitigating the effects of firebomb and blast attacks on metro systems. *Procedia – Social and Behavioral Sciences* 48: 3518–3527.
- Pate-Cornell, M.E. and Murphy, D.M. (1996) Human and management factors in probabilistic risk analysis: The SAM approach and observations from recent applications. *Reliability Engineering System Safety* 53(2): 115–126.
- Peterman, D.R. (2006) Passenger rail security: Overview of issues. Congressional Research Service, <http://www.au.af.mil/au/awc/awcgate/crs/rl32625.pdf>, accessed 15 July 2015.
- Piètre-Cambacédès, L. and Bouissou, M. (2013) Cross-fertilization between safety and security engineering. *Reliability Engineering & System Safety* 110: 110–126.
- Rasmussen, J. (1999) The concept of human error: Is it useful for the design of safe systems in healthcare. In: C. Vincent and B. deMol (eds.) *Safety in Medicine*. London: Elsevier, pp. 31–47.
- Rasputnig, C. and Opdahl, A. (2013) Comparing risk identification techniques for safety and security requirements. *Journal of Systems and Software* 86(4): 1124–1151.
- Riley, J. (2004) Terrorism and rail security. RAND Corporation. Santa Monica, CA, http://www.rand.org/pubs/testimonies/2005/RAND_CT224.pdf, accessed 15 July 2015.
- Rosness, R., Blakstad, H.C., Forseth, U., Dahle, I.B. and Wiig, S. (2012) Environmental conditions for safety work. Theoretical foundations. *Safety Science* 50(10): 1967–1976.
- Snook, S.A. (2002) *Friendly Fire: The Accidental Shootdown of US Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press.
- Tey, L.S., Wallis, G., Ferreira, L. and Hojati, A.T. (2011) *Using a Driving Simulator to Assess Driver Compliance at Railway Level Crossings*. Proceedings of the 34th Australasian Transport Research Forum; 28–30 September, Adelaide, Australia, pp. 1–13.
- Veltri, A. *et al* (2013) Understanding safety in the context of business operations: An exploratory study using case studies. *Safety Science* 55: 119–134.
- White, J. (2011) *Terrorism and Homeland Security*. Belmont, CA: Cengage Learning.
- Wilson, J.M., Jackson, B.A., Eisman, M. and Riley, K.J. (2007) Securing America's passenger rail systems. RAND Corporation. Santa Monica, CA, http://www.rand.org/pubs/monographs/2007/RAND_MG705.pdf, accessed 15 July 2015.
- Zohar, D. (2002) The effects of leadership dimensions, safety climate, and assigned priorities on minor injuries in work groups. *Journal of Organizational Behavior* 23(1): 75–92.
- Zohar, D. and Luria, G. (2005) A multilevel model of safety climate: Cross-level relationships between organization and group-level climates. *Journal of Applied Psychology* 90(4): 616.



This work is licensed under a Creative Commons Attribution 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>