AUTOMATION EFFECTS IN TRAIN DRIVING WITH TRAIN PROTECTION SYSTEMS – ASSESSING PERSON- AND TASK-RELATED FACTORS

Sonja Giesemann

German Aerospace Center (DLR e.V.)

Signals passed at danger (SPADs) are amongst the most common incidents in railway operations and are largely attributed to human errors. Train protection systems have therefore been introduced to mitigate and prevent SPADs. Overrunning and slipping past stops, however, remains an issue and very little empirical research has looked into the effects of such partial automation systems thereby, especially on the train driver's cognition, behaviour and activation state. To target this issue, an attention related integrative model of two automation phenomena as proposed by Parasuraman and Manzey (2010) was adapted to train driving with a German train protection system. Preliminary results of the derived questionnaire study on person and task related factors presented here revealed some effects regarding task-related competence and control expectations in train drivers that are deemed unfavourable in view of the emergence of attention deficits and automation errors.

Introduction

Even though seldom associated with severe consequences, events where a signal is passed at danger (SPAD) are of great concern within the railway industry. One reason for this concern is that SPADs are amongst the most common incidents in railway operations and are largely attributed to errors (in judgment or reaction) on part of the train driver. The German railway accident examination center (EUB, 2012) for example, has registered 462 non-technical failure SPADs out of a total 752 reported incidents[1] in 2011.

---

[1] An incident according to EUB is „an event in railway operations which compromises the safe operation of a train without eliciting direct damage of persons, property or environment"

Retrospective investigations of such events additionally uncover complex and highly individual interactions between various influencing factors, such as type of signal, time of day/ year, weather conditions, driver experience and other person related aspects. Therefore causes of SPADs have been deemed *"random human errors"* from a safety point of view (Nikandros & Tombs, 2007).

Great effort has consequently been put into the prevention of passing signals at warning and stopping aspects by technical solutions with assistance or automation character.
Such train protection systems serve, for example, to signalise braking curves, monitor speed restrictions and stops at signals and initiate automatic train stops whenever a train driver misbehaves.
However, although these systems have helped reducing SPAD risks and mitigating severity of consequences, they have not fully eliminated the issue (RSSB, 2012).

Catastrophic accidents, like the Hordorf train collision in Germany last year (EUB, 2011), can still occur if a train driver fails to attend to rail-side signals where such technologies fail or have not yet been equipped. And even with well operating systems in place, inadequate interpretations of and reactions to signal aspects can still cause critical incidents, for instance by "slipping past" a stop and its associated safety overlap or by unjustified releases of train protection braking.

It thus seems essential investigating the human factors in modern train driving with train protection systems.

Issues

Driving a train requires extensive knowledge, for example of operational rules and vehicle behaviour, and the ability to integrate various static and dynamic sources of information, such as schedules and track-side signals. The actual driving task is at the same time limited to one dimension, which is accelerating and decelerating (Branton, 1979).

These contrasting aspects of the train driver task have been widely acknowledged to encourage "adverse mental states" that lead to critical events in the railway system (Baysari et al., 2008). Fatigue, monotony and boredom seem to be issues related to normal and thus routine operation conditions (Edkins & Pollock, 1997; Wilde & Stinson, 1983; Buck & Lamonde, 1993; Dorrian et al., 2006). A high cognitive workload, on the other hand, is also being discussed in the context of driver information systems (Tappan, 2011; Kecklund et al., 2001; Roth & Multer, 2009).

What is striking is that very little empirical research has looked into the role of automation, especially for the train driver's cognition, behaviour and activation state, and how this could be related to critical incidents.
Even though it is very well known from the flight and process control domain for

example, that introducing (partial) automation leads to fundamental changes in task structure and related demands (i.e. Bainbridge, 1983; Sarter et al., 1997; Wickens & Hollands, 2000), it seems that automation effects with train protection systems have not yet been taken into consideration when it comes to SPAD events.

Approach

The German train protection system "PZB 90" serves to punctually monitor speed and attentiveness to warnings and to cover overruns of stop signals by eliciting emergency braking whenever a restriction is disregarded. These interventions are based on track-side and on-board magnets of different hertz frequencies.

The corresponding on-board interface consists of several indicator lamps, conveying status information and respective implicit calls for action (see figure 1), as well as three push-buttons ("order", "alert" and "clear") to acknowledge signal aspects and also to override restrictions under certain conditions.
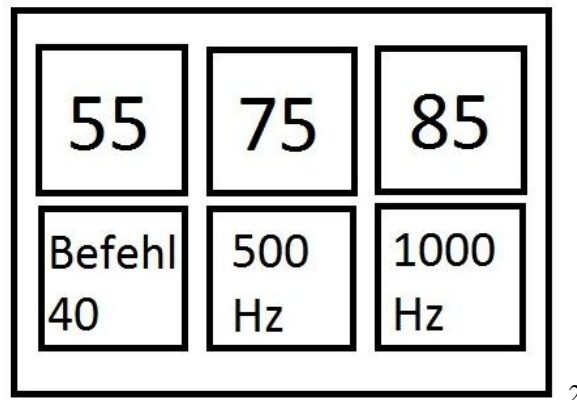

2

Figure 1: Schematic View of PZB Indicator Lamps (indicating train specifications/ target speeds on the top and supervision status on the bottom)

In the context of rail-bound transport, where the driving task is limited to longitudinal control, such a system can assumedly be characterised as partial automation as defined i.e. by the German Federal Highway Research Institute (BASt), meaning that "*the system takes over lateral and longitudinal control, the driver shall permanently monitor the system and shall be prepared to take over control at any time*".

---

[2] "Befehl 40" means a signal is being passed by order with a maximum speed of 40 km/h

This is assumed, since the train driver has to permanently monitor the interface (rail-side signals and corresponding changes of indicator lights) and act upon its implications for speed adjustments. Even though these adjustments remain manually controlled, they are based on feedback (indication to "take over control") and safeguarded against erroneous behaviour by the system.

Confidence in this safeguarding of manoeuvres appears to be high, since the monitoring of brake was occasionally termed a "life insurance" by train drivers.

Potential effects of such (partial) automation have been described in relation to attention by a model of Parasuraman and Manzey (2010). It comprises "complacency", describing insufficient monitoring of system state (Parasuraman, et al., 1993) and "automation bias", describing the insufficient verification of system recommendations for action (Mosier & Skitka, 1996). According to the model, errors occur when the automation fails, due to a loss of situation awareness by the supervisory controller which traces back to inappropriate attention allocation.

Since the model proposes various person and task related factors to influence the development of such attentional deficits, it was adapted to train driving with PZB 90 train protection and a questionnaire study for assessing these factors regarding the train driving task was designed.

Methodology

In accordance to the original model and additional empirical evidence (e.g. Prinzel et al., 2001), it is assumed that the train driving task is acted out by train drivers with individual traits and attitudes that can abet automation errors.

The survey was thus constructed to investigate relevant constructs both from a general and task-specific perspective. It was administered on-line with restricted access in order to ensure that participants were in fact professionally trained train drivers.

Attitudes

Trust was measured specifically as an attitude towards the PZB 90 system. Three of the five subscales of the "Human Computer Trust Scale" (HCT - Madsen & Gregor, 2000) were therefore adapted and translated to German. The HCT scale was based on a trust definition by McAllister (1995) which includes both a user's confidence in the system as well as his willingness to act upon its recommendations. As this willingness can be linked directly to the definition of automation bias (above) and is supposedly related to the more affective "believing in a system" (Madsen & Gregor,

2000), the respective adapted subscales "personal attachment", "faith" and "perceived reliability"[3] were administered here.

Five-point Likert-scale ratings of trust in particular aspects of the railway system were additionally obtained, ranging from "do not trust at all" (1) to "trust absolutely" (5).

In order to assess tendencies towards overreliance on technical systems in train drivers in general, the "Complacency-Potential Rating Scale" (CPRS – Singh et al., 1993a) was additionally used. By obtaining ratings of general statements on benefits and costs of automation related to trust, confidence, reliance and safety, this scale aims at assessing whether train drivers might be prone to complacency as the *"psychological state characterized by a low index of suspicion"* that Wiener (1981) defines.

This is underpinned by some empirical evidence, demonstrating that scoring high on complacency-potential does correlate with poor monitoring behaviour, at least when dividing experimental groups using the median split (Singh et al., 1993b; Prinzel et al., 2001). These studies additionally suggest that complacency should be regarded as a set of attitudes, rather than a trait, as it showed no correlation with personality inventories (ibid.).

Traits
Nevertheless, it was chosen to also investigate whether personality traits of train drivers would indicate effects that could foster the development of complacency and automation errors.

Self-efficacy was assessed globally via the "General Self-Efficacy Scale" (GSE) by Schwarzer and Jerusalem (1999). The scale was developed according to Banduras concept that *"perceived self-efficacy facilitates goal-setting, effort investment, persistence in face of barriers and recovery from setbacks"* (quoted after ibid.). This in return suggests that people scoring low on self-efficacy might exhibit stronger passivity in everyday life.

Krampen's (1991) questionnaire of competence- and control-beliefs (FKK) was used complementary to assess the train drivers' work-related characteristics. This tool was developed based on the author's action theory personality model, which builds on self-efficacy (according to Bandura, 1977) and also social-learning theory (i.a. according to Rotter, 1955, 1982).

---

[3] Reliability was originally thought to be a cognitive component of trust, but showed strong overlaps with the attachment scale (Madsen & Gregor, 2000)

It consists of four conceptually independent subscales, all of which were modified from a generalised to a task-context specific instruction and wording:

- Self-concept of abilities (SK-scale)
- Internal locus of control (I-scale)
- Socially induced external locus of control/ "powerful others control" (P-scale)
- Fatalistic external locus of control/ "chance control" (C-scale)

Standardised comparison values (T-values) from exhaustive reference samples are available for both personality tools.

Task-context
Participants were further asked to state their main information requirements and sources as well as non-satisfied additional information needs for an optimal driving performance.

Findings

Results
The results presented in this section have been obtained from a pilot sample of N= 7 male train drivers. Participants were on average 31.57 years old (SD=9.89) and had three to 24 years of train driving experience. Four participants worked in passenger and three in freight train operations. Out of these, two passenger and two freight train drivers worked in regional transport, the rest operated in intercity transport. Even though this seems to be a fairly representative sample, results have to be considered preliminary and can only be analysed descriptively, due to the small sample size.

Trust-ratings in the PZB 90 system were slightly above average on the 5 to 25 range for all three HCT subscales, ranging from M=17.29 for reliability and faith (SD_{reli}= 2.87; SD_{faith}= 3.04) to M=18.57 for personal attachment (SD_{pers}= 2.15).

Accordingly, trust in the train protection system as an aspect of the railway system was rated medium. The same holds for trust in track-side signalling. Trust in human operators on the other hand was rated fairly well -for the train driver himself as well as other train drivers and the control center operators (see table 1).

Table 1: Trust-ratings for Railway System Aspects

	Mean rating	Standard deviation
Train protection system	3.29	0.49
Track-side signalling	3.29	0.49
Myself	4	1
Other train drivers	4	0.58
Control center operators	4.29	1.11

Scores regarding the general complacency-potential of train drivers were comparably low. With a mean of 50.71 (SD=4.96) and a median of 51, results are well below median split scores previously obtained by Prinzel et al. (Md=58; 2001) or Singh et al. (Md=56; 1993b).

General self-efficacy perception (GSE) scores also displayed no exceptional results. As figure 2 shows, only one participant exhibited a normalised T-value considerably below the average reference sample range, as indicated by the grey bar.
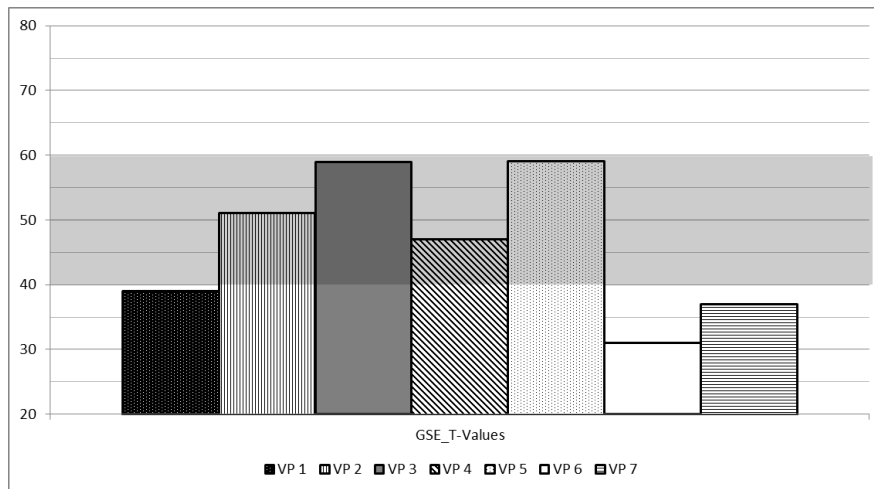

Figure 2: T-values of General Self-Efficacy Scores

The results for the task-specifically modified competence- and control expectation questionnaire (FKK), however, revealed some effects on the subscales assessing "self-concept of abilities" (SK) and "powerful others control" (P). Four participants scored well below the normal range for the SK scale and three highly above it for the P scale, with participant six exhibiting both deviations (see figure 3).
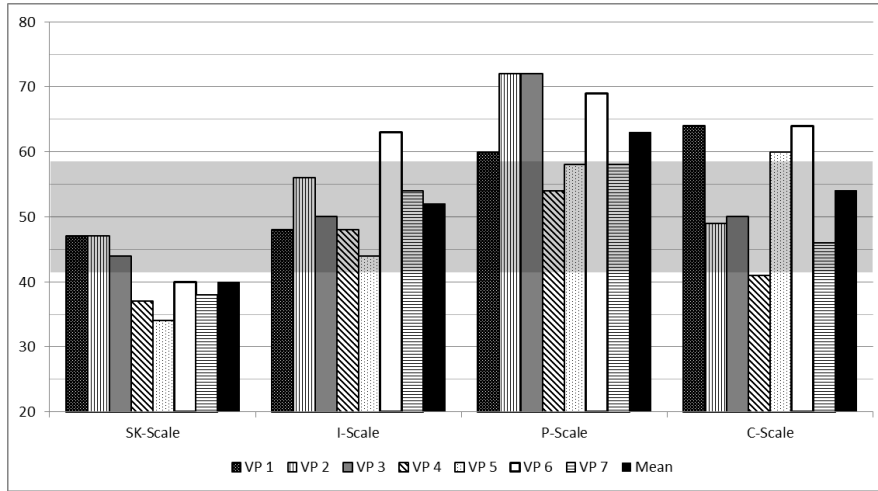
Figure 3: T-values of Competence- and Control-Beliefs Subscales

This is of particular interest, since the FKK questionnaire basically comprises the same construct as the GSE scale, but is extended by action-theory aspects and was instructed towards the participants' beliefs with reference to their working-context.

Pearson correlations were subsequently computed to investigate whether these task-specific train driver characteristics can be related to the attitudes examined earlier. Significant correlations occurred only for the SK scale.

It was positively correlated with the faith aspect of the HCT scale and negatively correlated with the trust-related complacency-potential of the CPRS as well as the trust-ratings for track-side signaling.

Table 2: Significant Correlations with Self-Concept of Ability Scale

	r	$F_{1,5}$	α-level
HCT Faith	0.764	6.990	.95
CPRS Trust	- 0.904	22.440	.95
Track-side Signals	- 0.737	5.939	.90

Discussion

According to Krampen (1991), people who score low on the "self-concept of ability" (SK) scale are rather passive and observant, unsure of themselves and insecure in novel situations; they tend to have low self-esteem, lack ideas, know little options for action and see few scopes for action in difficult situations.

Train drivers in the reported preliminary sample exhibited these tendencies in relation to their working life, even though their general expectations towards scope

of action and problem-solving resources in everyday life were unobtrusive. This indicates that some aspects of the train driving task and working-environment do in fact promote passivity. Passivity, in turn, is a key factor in the loss of situation awareness and subsequent development of automation errors as described by Parasuraman & Manzey (2010).

Since having an unfavorable concept of personal abilities is furthermore negatively correlated with trust in track-side signaling, it could be argued that when train drivers are rather passive and observant, they rely more on external system feedback to initiate task execution, rather than following anticipatory, self-determined strategies.
This could be critical, since anticipation is a necessary element for higher level situation awareness and a lack thereof could promote automation errors at system failure (see Parasuraman & Manzey, 2010).

A similar relation can be found with trust-based complacency-potential scores, indicating that the lower a train driver's expectations regarding own competence are, the stronger automated systems are trusted to function more accurately than a human operator. Taking into consideration that PZB 90 was called a "life insurance" during previous workplace interviews, insufficient attention to the correct functioning of the system ("complacent behaviour" according to Parasuraman et al., 1993) seems a likely consequence.

The train drivers' "self-concept of abilities" score is, however, also positively correlated with the HCT subscale of faith, meaning the conviction of a systems future ability to perform well, even in unknown situations.
Considering that participants were instructed to rate the modified HCT items in reference to "train driving with PZB 90 train protection", possibly intertwining the rating of the train protection system and the driving task, it seems likely that train drivers who are insecure in novel situations (low SK score) would also have little faith in the correct functioning of the entire system (driver-machine interaction) in unknown situations due to expectations of their own of shortcomings.

Together with the high work-related scores on the "powerful others control" scale, identifying people who are usually not assertive at all, frequency experience feelings of powerlessness and helplessness, feel inferior to powerful others, see events in life to be caused by others and perceive themselves and their lives as strongly dependent on other people (Krampen, 1991), these relations indicate an overall (and possibly undesired) passive and dependent role of the train driver. This is also congruent with the fairly high ratings of trust in control center operators, since they steer train movements with access to anticipatory information that train drivers lack.

Initial inspections of the qualitative data regarding the informational task-context support these assumptions. Almost every participant stated that he needed

information about specifics of the route ahead as well as surrounding train movements to drive safely and efficiently, but rated availability as insufficient. Also the rare and deficient communication with control center operators was perceived as a nuisance by most participants.

Conclusion
The preliminary results obtained by this small pilot sample questionnaire study indicate that the train driving task elicits effects on task-related competence and control expectations which are unfavourable in view of the emergence of attention deficits and automation errors.

While attitudes towards the train protection system do not seem to be a reason for concern, the lacking feasibility of anticipation and a proactive driving style might force train drivers into passivity and condition the inability to verify system behaviour with environmental cues. This could abet loss of situation awareness and thus errors upon automation failure.

Future work is needed to determine whether the findings hold true for a larger sample. If this is the case, experimental work is additionally needed to assess how the findings have an impact on behavioural data of train drivers and whether context information could be presented to train drivers to allow for development of adequate situation awareness and an active driving style.

References

Bandura, A. 1977, *Social Learning Theory,* (Prentice Hall, Englewood Cliffs)

Bainbridge, L. 1983, Ironies of Automation, *Automatica*, 19, 775–779

Grasser: T. M. and Westhoff, D. 2012, BASt-study: Definitions of Automation and Legal Issues in Germany, *Transportation Research Board - Road Vehicle Automation Workshop*, 25th July 2012, Irvine, California

Baysari, M. T., McIntosh, A. S. and Wilson, A. R. 2008, Understanding the Human Factors Contribution to Railway Accidents and Incidents in Australia, *Accident Analysis and Prevention*, 40, 1750–1757

Branton, P. 1979, Investigations into the Skills of Train-Driving, *Ergonomics*, 22 (2), 155-164

Buck, L. and Lamonde, F. 1993. Critical Incidents and Fatigue Among Locomotive Engineers, *Safety Science*, 16, l-18

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 9*, 243–257

Edkins, G. D. and Pollock, C.M. 1997, The Influence of Sustained Attention on Railway Accidents, *Accident Analyses and Prevention*, 29 (4), 533-539

Eisenbahn-Unfalluntersuchungsstelle des Bundes (EUB) 2011, *Unfallbericht Zugkollision Überleitstelle Hordorf / Strecke Magdeburg Hbf – Halberstadt am 29.01.2011*, (Bundesministerium für Verkehr, Bau und Stadtentwicklung, Bonn)

Eisenbahn-Unfalluntersuchungsstelle des Bundes (EUB) 2012, *Eisenbahn-Unfalluntersuchung Jahresbericht 2011*, (Bundesministerium für Verkehr, Bau und Stadtentwicklung, Bonn)

Kecklund, L. and the TRAIN project group 2001, *Final Report on the TRAIN Project*, (Banverket, Lund)

Krampen, G. 1991, *Fragebogen zu Kompetenz- und Kontrollüberzeugungen* (FKK), (Hogrefe, Göttingen)

Madsen, M. and Gregor, S. 2000, Measuring Human-Computer Trust, *Proceedings of Eleventh Australasian Conference on Information Systems*, 6th-8th December, Brisbane, Australia

McAllister, D. 1995, Affect- and Cognition-based Trust as Foundations for Interpersonal Cooperation in Organizations, *Academy of Management Journal*, 38, 24-59

Mosier, K. L., and Skitka, L. J. 1996, Human Decision Makers and Automated Decision Aids: Made for Each Other? In R. Parasuraman and M. Mouloua (eds.), *Automation and Human Performance: Theory and Application*, (Erlbaum, Mahwah, NJ), 201–220

Nikandros, G. and Tombs, D. 2007, Measuring Railway Signals Passed at Danger. *Proceedings of the twelfth Australian workshop on safety critical systems and software and safety-related programmable systems,* 86, 41-46

Parasuraman, R., Molloy, R. and Singh, L. 1993, Performance Consequences of Automation Induced Complacency, *The International Journal of Aviation Psychology*, 3 (1), 1-23

Parasuraman, R. and Manzey, D. H. 2010, Complacency and Bias in Human Use of Automation: An Attentional Integration, *The Journal of the Human Factors and Ergonomics Society*, 52, 381-410

Prinzel, L. J., De Vries, H., Freeman, F. G., and Mikulka, P. 2001, *Examination of Automation-induced Complacency and Individual Difference Variates*, (National Aeronautics and Space Administration Langley Research Center, Hampton)

Roth, E. M., and Multer, J. 2009, *Technology Implications of a Cognitive Task Analysis for Locomotive Engineers*, (U.S. Department of Transportation/ Federal Railroad Administration, Washington, D.C.)

Rotter, J. B. 1955, The Role of the Psychological Situation in Determining the Direction of Human Behaviour, *Nebraska Symposium on Motivation*, 3, 245-268

Rotter, J. B. 1982, *The Development and Application of Social Learning Theory*, (Praeger, New York)

Rail Safety and Standards Board –RSSB 2012. *RSSB Annual Safety Performance Report, 2011/12 -A reference guide to safety trends on GB railways*, (Rail Safety and Standards Board, London)

Sarter, N., Woods, D. D. and Billings, C. 1997, Automation Surprises. In G. Salvendy, (ed.) *Handbook of Human Factors/Ergonomics*, Second Edition, (Wiley, New York)

Schwarzer, R. and Jerusalem, M. 1999, *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (Freie Universität Berlin, Berlin)

Singh, L., Molloy, R., and Parasuraman, R. 1993a, Automation-induced "Complacency": Development of a Complacency-potential Scale, *International Journal of Aviation Psychology*, 3, 111-122

Singh, I. L., Molloy, R., and Parasuraman, R. 1993b, Individual Differences in Monitoring Failures of Automation, *Journal of General Psychology*, 120, 357–373

Tappan, J. M., Pitman, D.J., Cummings, M. L. and Miglianico, D. 2011, Display Requirements for an Interactive Rail Scheduling Display, *Proceedings of the 14th International Conference on Human-Computer Interaction*, 9-14 July, Orlando, Florida

Wickens, C. D., and Hollands, J. G. 2000, *Engineering Psychology and Human Performance*, (Prentice Hall, New York)

Wiener, E. L. 1981, Complacency: Is the Term Useful for Air Safety? *Proceedings of the 26th Corporate Aviation Safety Seminar*, Flight Safety Foundation, Denver, 116–125

Wilde & Stinson, 1983 The monitoring of Vigilance in Locomotive Engineers, *Accident Analysis and Prevention*, 15 (2), 87-93