Individual markers of resilience in train traffic control: The role of operators’ goals and strategic mental models and their implications for variation, expertise and performance.

Lo, J.C.; Pluyter, K.R.; Meijer, S.A.

DOI
10.1177/0018720815622357

Publication date
2016

Published in
Human Factors: the journal of the human factors and ergonomics society

Citation (APA)
Lo, J. C., Pluyter, K. R., & Meijer, S. A. (2016). Individual markers of resilience in train traffic control: The role of operators’ goals and strategic mental models and their implications for variation, expertise and performance. *Human Factors: the journal of the human factors and ergonomics society*, *58*(1), 80-91. https://doi.org/10.1177/0018720815622357

Individual Markers of Resilience in Train Traffic Control: The Role of Operators' Goals and Strategic Mental Models and Implications for Variation, Expertise and Performance

Julia C. Lo[1], Kari R. Pluyter & Sebastiaan A. Meijer[2]

[1]Delft University of Technology, Delft, The Netherlands

[2]KTH Royal Institute of Technology, Stockholm, Sweden

Running head

Individual Markers of Resilience

Manuscript type

Research article

Word count

4383

Acknowledgments

This study was funded through the Railway Gaming Suite program – a joint project by ProRail and Delft University of Technology. We thank Guido Band from the University of Leiden and Emdzad Sehic from ProRail for their involvement. We also thank the editor and the reviewers for their involvement and constructive feedback.

Abstract

Objective: To examine individual markers of resilience and obtain quantitative insights into the understanding and the implications of variation and expertise levels in train traffic operators' goals and strategic mental models, and their impact on performance.

Background: The Dutch railways are one of the world's most heavy utilized railway networks and have been identified to be weak in system and organizational resilience.

Methods: Twenty-two train traffic controllers enacted two scenarios in a human-in-the-loop simulator. Their experience, goals, strategic mental models and performance were assessed through questionnaires and simulator logs. Goals were operationalized through performance indicators, and strategic mental models through train completion strategies.

Results: A variation was found between operators for both self-reported primary performance indicators and completion strategies. Further, the primary goal of only 14% of the operators reflected the primary organizational goal (i.e. arrival punctuality). An incongruence was also found between train traffic controllers' self-reported performance indicators and objective performance in a more disrupted condition. The level of experience tends to impact performance differently.

Conclusion: There is a gap between primary organizational goals and preferred individual goals. Further, the relative strong diversity in primary operator goals and strategic mental models indicates weak resilience at the individual level.

Application: With recent and upcoming large-scale changes throughout the socio-technical space of the railway infrastructure organization, the findings are useful to facilitate future railway traffic control and the development of a resilient system.

Keywords

goal competition, diversity, organizational resilience, railway, socio-technical system

Précis

This study investigated the variation in goals and strategic mental models of train traffic controllers as individual markers of resilience. Individual differences in experience within goals and strategic mental models and their link with performance were also investigated.

INTRODUCTION

Resilience engineering studies are relevant in multiple domains, especially those that are highly complex and known for their hazards (Nemeth, Wears, Patel, Rosen, & Cook, 2011). Domains that are most heavily investigated are aviation (22%), healthcare (19%), the chemical and petrochemical industry (16%), nuclear power plants (10%) and railway (8%) (Righi, Saurin, & Wachs, 2015).

For the Dutch railway infrastructure managing organization ProRail, the notion of resilience and robustness strongly resonates in the organization to improve the system along these concepts (Meijer, 2012). The idea is that when the system cannot maintain the regular way of working, resilience is required to respond through the adaptation of strategies (Burnard & Bhamra, 2011; Hollnagel & Woods, 2005; Hollnagel, 2014).

However, resilience is linked across different levels with influencing mechanisms on the industry at the highest level, followed by plant and operations (organization), and teams and individuals at the lowest level (Back, Furniss, Hildebrandt, & Blandford, 2008; Sheridan, 2002). Research often focuses on a specific unit of analysis, as it is yet not well understood how resilience is linked across these different levels (Righi et al., 2015). The study of cross-level interactions inside the system is, however, crucial to prevent brittleness in the overall system, which can be facilitated through proactive safety management (Gomes, Woods, Carvalho, Huber, & Borges, 2009).

An analysis of railway safety operations in the Netherlands revealed poor to mixed resilience levels (Hale & Heijer, 2006). The debundling and privatization of the railway system that was widely introduced across Europe in the 1990s, causing extensive institutional fragmentation of the system, is a possible reason for the low resilience levels

(Hale & Heijer, 2006; Knieps, 2013). The debundling of the railway system inextricably led to more brittle operational processes for railway traffic operators, resulting in, for example, unclear and conflicting goals and the development of multiple coping strategies (Steenhuisen & De Bruijne, 2009; Veeneman, 2006). This phenomenon can be also be labelled as a gap between the system as designed or imagined and the system as it is actually operated, which results in a distance between the various levels (Dekker, 2006).

At an individual level, resilience engineering can help operators to develop robust yet flexible responses to disturbances inside or outside the organization (Chialastri & Pozzi, 2008; Lengnick-Hall & Beck, 2005). As such, performance variability is normal, though it needs to be controlled. Performance variability that leads to positive outcomes should be promoted (Hollnagel, 2008; 2014). Having shared goals, experiences, robust responses to simple problems and flexible responses to complex problems is essential to the development of a resilient organization (Lengnick-Hall, Beck, & Lengnick-Hall, 2011).

Departing from resilience studies in the Dutch railways at a system and organizational level, this study focused on the individual level of railway traffic operators, in order to provide recent and quantitative insights to further the understanding of variations in their cognition and behavior and the implications thereof. The central research questions were: to what extent do organizational and individual goals correspond? What is the level of diversity in the goals and strategic mental models of train traffic operators given operators' work experience, and how does this relate to their performance?

The following section briefly introduces the Dutch railway system from a number of perspectives. This is followed by a brief presentation of the theoretical background to goals and strategic mental models. The subsequent sections present the method, results, and discussion and conclusion.

A MULTILEVEL OVERVIEW: RAILWAY TRANSPORT IN THE NETHERLANDS

The Dutch railways transport more than one million passengers and operate about 350 freight trains per day in a relatively small country, making it one of the busiest railway systems in Europe and even the world (Meijer, 2012; ProRail, 2013; Ramaekers, De Wit, & Pouwels, 2009). The debundling and privatization of the Dutch railway sector in 1995 initially led to a decrease in the performance of the system: between 2000 and 2001, arrival punctuality dropped from about 87% to 80% (Algemene Rekenkamer 2012; Steenhuisen, 2009). Although the punctuality of trains has recovered over the years, reaching 92% in 2014, both the principle passenger transport manager (Nederlandse Spoorwegen; NS) and the infrastructure manager (ProRail) were penalized for performing insufficiently on the agreed performance indicators – NS for passenger dissatisfaction in terms of punctuality and quality of service (e.g. number of available seats, crowdedness during peak hours), and ProRail for the insufficient availability of the infrastructure due to malfunctions (Rijksoverheid, 2015).

Organizational performance indicators

Safety, reliability, service and capacity use can be seen as key public values in the railway domain (Wilson, Farrington-Darby, Cox, Bye, & Hockey, 2007). The general public values that are held in the governance of railway transport are rather stable over time, unlike the operationalization and quantification of these values into goals or performance indicators (Veeneman & Van de Velde, 2006). For instance, reliability can be conceptualized in a number of ways, such as punctuality, which can be further operationalized in terms of, for example, arrival, departure or overall (arrival and departure) punctuality. Departure punctuality was a performance indicator until 2006, when arrival punctuality became the indicator (Veeneman, 2006). However, both railway infrastructure and passenger transport managers set different thresholds in arrival punctuality, namely < 3 minutes and < 5 minutes, respectively (NS, 2015; ProRail, 2015a). The formalization of performance indicators is an annual iterative process with occasionally ad hoc organizational reactions throughout the year in the case of unexpected large-scale disruptions that are subject to media scrutiny.

Train traffic control

Railway traffic operations differ between European countries in a number of ways, such as organization, roles and responsibilities, and level of automation (Golightly et al., 2013). In the Netherlands, a train traffic management system is used to execute the timetables, which are operated by train traffic controllers. The primary responsibility of these controllers is to execute train timetables in an accurate and punctual manner (Sulmann, 2000). Maintaining the operational safety of the rail system and recovering after disruptions and accidents is an essential part of their job (Crawford, Toft, & Kift, 2014). Train traffic controllers do not perceive their primary task as challenging as long as routes are already scheduled (Roth & Patterson, 2005). However, a more active role is needed in unsafe situations that cannot be controlled by the automated safety system or when there is a system malfunction (Sulmann, 2000).

Future developments

In terms of future developments, ProRail and the government stated their intention to double the railway track capacity between 2008 and 2020 (now extended to 2028), which should lead to a timetable that supports both an intercity and a local train service six times per hour in both directions between major cities (Meijer, Van der Kracht, Van Luipen, & Schaafsma, 2009; ProRail, 2015b). Given the restriction of a capacity increase through the mere addition of tracks, a change in the organizational processes is also required. As such, process optimization programs are being implemented that focus on, for instance, increasing the centralization of decision making to the national control center (operational control center rail; OCCR) for disruption mitigation procedures and restructuring the roles and responsibilities of operators. Switches are increasingly being removed at major stations (e.g. 110 of the 170 switches are being removed at Utrecht Central station) in order to, for example, facilitate corridor management, shorter travel times and more reliable traffic control, while bottleneck areas in the infrastructure are being expanded and upgraded. Finally, the replacement of the current traffic management system is being explored.

GOALS AND MENTAL MODELS

Goals

Goals are states or ends that someone wants to achieve (Latham & Locke 1991; Mohammed, Klimoski, & Rentsch, 2000; Popava & Sharpanskykh, 2011). Operators' goals influence their mental model selection and therefore their decision making and performance (Endsley, 1995). In a dynamic environment, individuals focus on elements in the environment that are goal related. Deriving the meaning of the elements and the projection to the future is done in light of the goal and the active mental models (Endsley, 1995). Goals influence the valuation of multiple options during decision making (Mohammed et al., 2000). In order to achieve resilience, operators need to have a common set of goals (Lengnick- Hall, Beck, & Lengnick Hall, 2011).

Mental models and expertise

Mental models are mental representations of humans, systems, artifacts and situations formed by experience, observation and training (Endsley, 1995; Schaffernicht & Groesser, 2011; Wilson, 2000). Mental models store knowledge that is necessary for human–environment interaction (Klimoski & Mohammed, 1994; Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000). This knowledge is crucial for solving problems effectively, such as those faced by train traffic controllers when confronted with multiple disruptions to the train schedule. Visual attention and evaluation of relevant information in complex problem situations improves when mental models are well developed.

The degree of development of mental models differs between novices and experts. Experts with extensive domain knowledge have developed the ability to perceive important patterns and features that are not seen by novices (Glaser & Chi, 1988; Bransford, Brown, & Cocking, 2000; Bogard, Liu, & Chiang, 2013). Experts also have the capacity to better recognize meaningful patterns due to their superior knowledge organization and extensive domain knowledge (Glaser & Chi, 1988). In contrast, novices' knowledge consists of facts, procedures and formulas that are not as well organized, as they do not have integrated mental models. Novices are therefore oriented towards surface characteristics in problem solving (Glaser & Chi, 1988; Bogard, Liu, & Chiang, 2013). Furthermore, experts have developed a condition–action ability through practice. Experts have conditioned knowledge: the recognition of specific patterns triggers an appropriate response that is useful for problem solving (Glaser & Chi, 1988; Bransford, Brown, & Cocking, 2000). Different levels of expertise may influence the performance of train traffic controllers and therefore resilience at an individual level (Lengick-Hall, Beck, & Lengnick-Hall, 2011).

METHOD

Experimental setting

A simulator session was used to familiarize train traffic controllers with the new infrastructure that would result from the removal of 66 switches in three months' time. The simulator was strongly focused on the logistical aspects of train traffic control and much less on technical safety-related aspects. The infrastructure that was simulated was the train traffic area around Utrecht Central Station. This area is operated by two train traffic workstations. One controller was responsible for the trains that belong to the ‘turn’ (in Dutch: *keer*) area, and a second controller was responsible for the ‘through’ (in Dutch: *door*) area. The role allocation was reversed in the second round.

Two scenarios were designed for the participants: scenario 1 consisted of a light disruption in the train traffic flow caused by minor train delays, whereas scenario 2 represented a moderately to severely disrupted flow. In the first round, 22 controllers participated in scenario 1. In the second round, 10 participants participated in scenario 1 and 10 participated in scenario 2. Both scenarios were designed in collaboration with two senior train traffic controllers. Train traffic controllers were asked to perform their job as they typically would at their actual workstation. No interaction between the train traffic controllers was needed to conduct their tasks.

Participants

All 22 train traffic controllers (18 males, 4 females) worked at Utrecht Central Station. They took part in a 2 (workstation area: turn or through) x 2 (severity of disruption: high versus low level of train delays) within-subject experimental design.

Materials

Work experience and *job role* were assessed using questionnaires. Participants were assigned to a high or a low experience group based on their work experience as train traffic controllers. The cut-off point was set at 10 years, as a new traffic management system had been implemented 10 years earlier (Bary, 2015).

Operator goals were operationalized through performance indicators (Popova & Sharpanskykh, 2011). A list of performance indicators for train traffic controllers was created prior to this session by six senior train traffic controllers. Participants ranked these performance indicators on a scale of 1 to 7 (1 = most important, 7 = least important).

Speed of acquaintance was included to find out how fast the participants were able to get accustomed to the new infrastructure. This item was measured on a 5-point Likert scale, ranging from fully disagree to fully agree. Participants could also opt for 'I do not know' as an answer.

Performance was measured using five performance indicators, namely arrival punctuality, departure punctuality, amount of arrival delay, amount of departure delay, and platform consistency. Arrival and departure punctuality was operationalized through trains that arrive at (or depart from) Utrecht Central Station on time or with less than a 3-minute delay. These trains were counted, summed up and divided by the total number of arrived/departed trains. For the arrival and departure delay in minutes, the amount of delays in minutes was summed up and divided by the total number of arrived/departed trains. With regards to platform consistency, all trains that did not arrive at the planned track were counted and summed up, and the same was done for all trains that did not arrive at the planned platform. Secondly, the total number of trains that did not arrive at the planned platform and at the planned track were summed up and divided by the total number of arrived trains for each train traffic controller.

Strategic mental models. Mental models can be conceptualized as declarative (knowledge of what), procedural (knowledge of how) or strategic (knowledge of what and how, and applied to the context) (Mohammed, Ferzandi, & Hamilton, 2010; Salas, Stout, & Cannon-Bowers, 1994). Strategic mental models can also be operationalized by generating lists of actions with subject matter experts (Webber, Chen, Payne, & Zaccaro, 2000). As such, the completion strategies of a train traffic controller could be an indicator of the controller's strategic mental model. Simulator logs were used to analyze the completion strategies where different ways of dealing with the train delays (i.e. the different order of departure of trains that were handled given their delay) were expected to be possible. Given the length of scenario 1, three conflict points for completion strategies for the through workstation and one conflict point for the turn workstation were identified; for scenario 2, one and two completion strategies were identified for the through workstation and the turn workstation, respectively. Different completion strategies were subsequently assessed by analyzing whether the completion strategies were followed according to the preferred completion strategy (as was scheduled) and the different strategies applied, to assess the variability per operator and per conflict point. Analyses were done based on participants who enacted scenario 1 in both rounds and those who enacted scenario 1 and subsequently scenario 2, in order to obtain four conflict points per individual.

Simulator validity was measured through three components: structural validity (the degree of similarity in structure between the simulated and the reference system), processes validity (the degree of similarity in processes between the simulated and the reference system) and psychological reality (the degree to which the participants perceived the simulated system as realistic) in line with Raser (1969), using a questionnaire designed by Lo, Sehic and Meijer (2014). An example of a structural validity item is: ‘I can apply the information from the information sources in the simulator in a similar way as in the real world’ (α = .65 with the removal of one item). The item ‘The train traffic flow in the simulator is similar in their processes to the real world train traffic flow’ represented process validity (α = .60). An example of psychological reality (α = .67) is ‘The simulation environment feels more or less like my own work environment’. These items were measured on a Likert scale.

Procedure

The participants completed a questionnaire before the start of the session. They then enacted the two 40-minute scenarios. At the end of each round, they completed another questionnaire. During the second round, knowledge probes were administered for the purpose of another study. Video recordings were made throughout both sessions.

RESULTS

Six of the 22 participants were excluded from the analysis because they had known about the train delays. Two of the 16 participants included in the analysis had enacted scenario 2 twice. As there were a few problems with the simulator, not all train traffic controllers received the same number of trains. This issue was controlled for by using an average score of the objective performance measures and reviewing the severity of issues through video recordings for events that hindered participants in their options or decision making.

The average score of the participants’ work experience in their current function was 10.3 years (*SD* = 9.24).

Simulator validity

The findings show that the participants tended to be slightly positive about the validity of the simulator considering the task they were given (see Table 1). The participants also indicated that they had quickly got used to the simulator.

TABLE 1: Descriptive statistics of the validity of the simulator on the three validity components and speed of acquaintance with the simulator and the two workstations, measured on a 5-point Likert scale.

	N	**M**	**SD**
Structural validity	21	3.5	.92
Process validity	20	3.6	.66
Psychological reality	22	3.7	.71
Speed of acquaintance with the simulator	20	4.2	.83
Speed of acquaintance with the turn workstation	21	3.9	.62
Speed of acquaintance with the through workstation	21	4.0	.59

Regarding learning effects between scenarios, the participants indicated that they had got used to both workstations relatively quickly.

Goals

Figure 1 shows a relative moderate goal consistency among the train traffic controllers. Three controllers added two more performance indicators, but these were not included in the analysis.

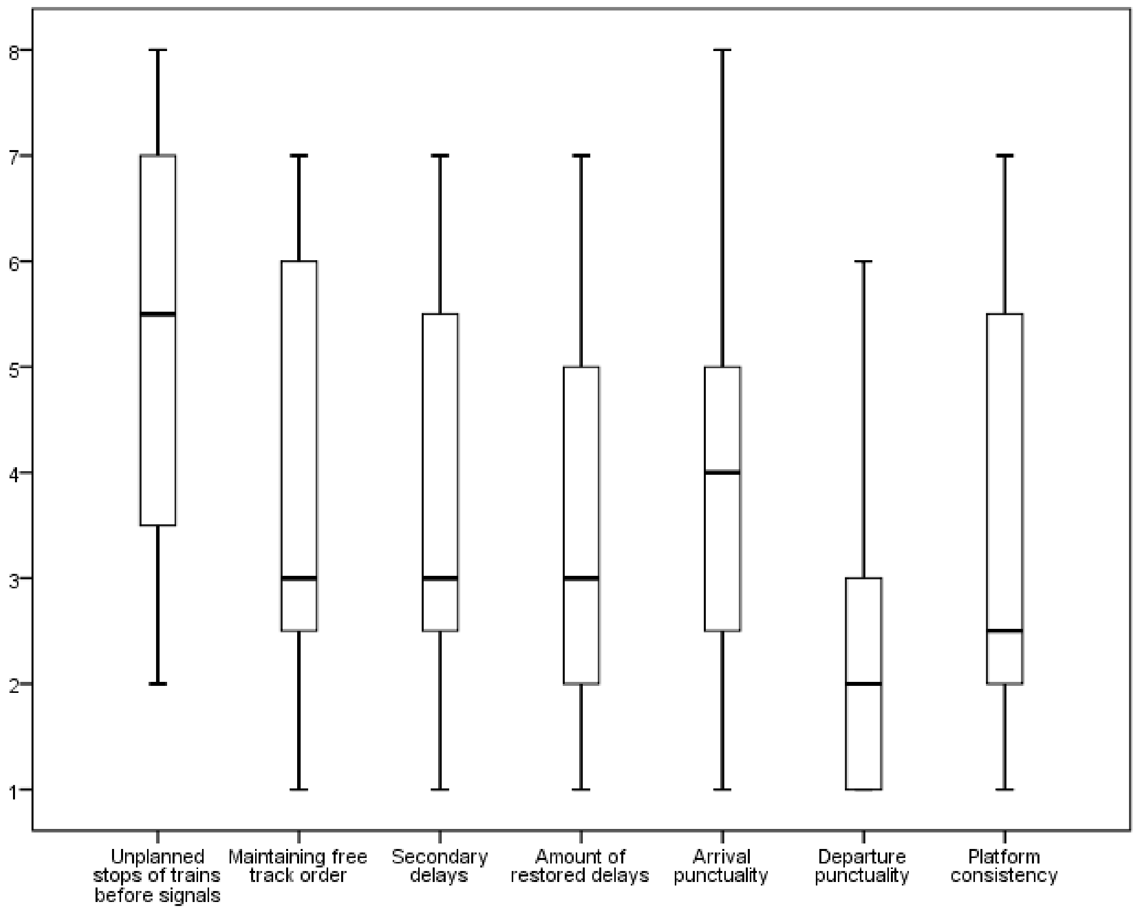

Figure 1. Median distribution of self-reported performance indicators (x-axis) with the ranking scale depicted on the y-axis (N=20).

In the assessment of primary preferred performance indicators, however, departure punctuality was consistently perceived as most important (36%). This was followed by achieving high platform consistency (18%), arrival punctuality (14%), maintaining free track order (i.e. track use between stations in the planned order) (9%), the number of restored delays and secondary delays (both 5%) and the avoidance of unplanned stops of trains before signals (0%). As such, these results show a very fragmented preference with regards to primary key performance indicators.

Strategic mental models

The operators' strategic mental models were analyzed to obtain insights into the diversity of their individual completion strategies. The overall findings show that participants handled on average 61% of the completion strategies in the preferred manner (*SD* = 31.5). Those who enacted scenario 1 twice handled 53% of the completion strategies in a deviating manner (*SD* = 21.1). Participants who enacted scenarios 1 and 2 handled on average 37% of the completion strategies in the preferred manner (*SD* = 14.2), and 65% in an alternative manner (*SD* = 24.2) (see Figure 2). Based on Figure 2, a qualitative assessment supports the variation in completion strategies with regards to the operators' individual completion strategies.

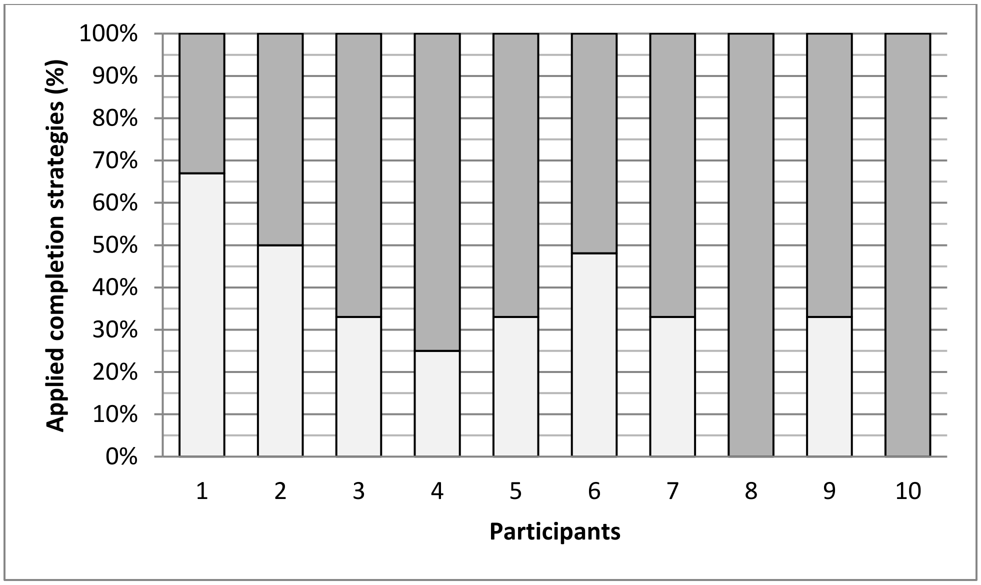

Figure 2. Applied completion strategies per participant for operators who enacted scenarios 1 and 2. A white band indicates a preferred completion strategy being followed, and a grey band indicates alternative completion strategies. Even numbers represent participants from the through workstation, odd numbers those from the turn workstation.

An analysis of the level of variation in completion strategies for each conflict point revealed diversity based on between one and three different completion strategies for four conflicting points in scenario 1, and on five different variations of completion strategies for three conflicting points in scenario 2 (see Figure 3). A qualitative assessment would show that there is a level of variation in the completion strategies with regards to different conflict points and that this differs between scenarios: operators dealt with these conflict points with more diverse completion strategies in the moderately disrupted scenario than in the lightly disrupted scenario. Further, it is notable that preferred completion strategies were implemented more frequently in scenario 1.

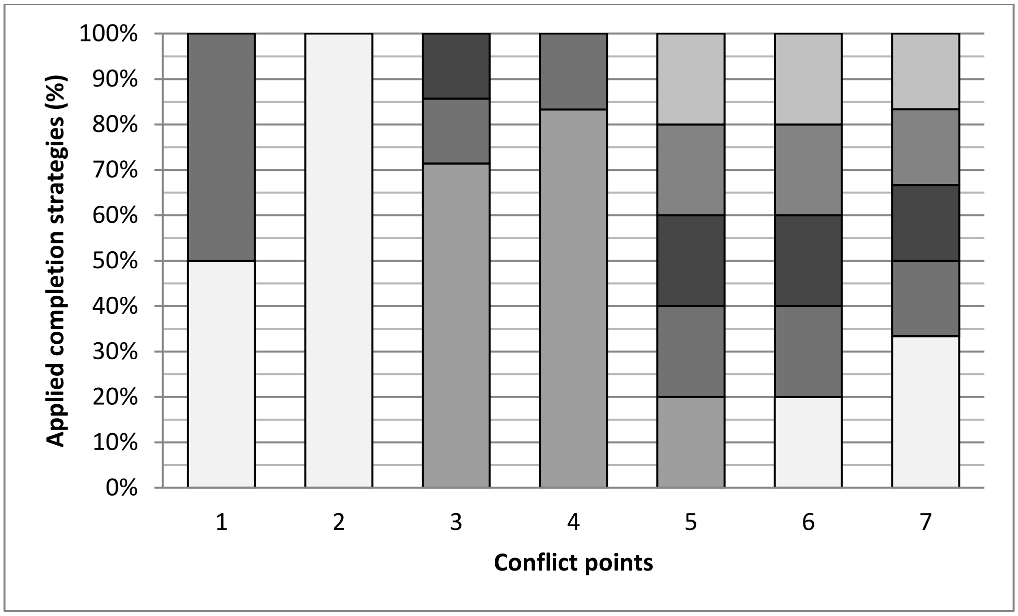

Figure 3. Applied completion strategies per conflict point for scenario 1 (1–4) with N =14 and scenario 2 (5–7) with N = 10. A white color indicates a preferred completion strategy being followed, while different shades of grey indicate different completion strategies. Numbers 1, 2, 5 and 6 represent conflict points from the through workstation and numbers 3, 4 and 7 represent conflict points for the turn workstation.

Performance

Spearman correlation tests were performed to test whether there is a congruence between the self-reported relative importance of performance indicators and objective performance (see Table 2). Although scenario 1 does not reveal any significant correlations, scenario 2 does, namely a strong positive correlation between self-reported departure punctuality and objective arrival delay. Also a strong negative correlation was found between self-reported arrival punctuality and objective departure delay. A trend for a negative correlation between self-reported platform consistency and objective arrival delay was also found.

TABLE 2: Correlation between self-reported performance indicators and objective performance indicators for scenario 2.

Self-reported performance indicator	**Objective performance indicator**	**N**	**r**	**p**
Departure punctuality	Arrival delay	10	.79	.007**
Arrival punctuality	Departure delay	10	-.73	.018*
Platform consistency	Arrival delay	9	-.59	.097

*$p \leq .01$, **$p \leq .05$

Although unexpected, these results provide interesting insights into goal competition, as they suggest that arrival punctuality and departure delay, departure punctuality and arrival delay, and platform consistency and arrival delay are competing goals.

A Spearman correlation test was also performed between the applied preferred and alternative completion strategies and performance. No significant relations were found.

Experience

It was expected that the more experienced controllers would outperform the less experienced controllers due to their better organized mental models. The analyses showed a significant tendency in scenario 1 for controllers with less experience in their current function to have a higher arrival punctuality score than the more experienced controllers (see Table 3). An opposite tendency was found in scenario 2: the controllers with more experience in their current function have a higher arrival punctuality score than the controllers with less experience in their current function.

TABLE 3: Differences in objective performance between more and less experienced train traffic controllers in their current function.

Scenario	**Objective performance indicator**	**Experience**	**Mean Rank**	**N**	**U**	**p**
1	Arrival punctuality	Low	13.8	11	13.0	.005*
		High	6.4	9		
2	Arrival punctuality	Low	3.9	5	4.5	.09
		High	7.1	5		

*$p \leq .01$, **$p \leq .05$

It was also investigated whether the applied completion strategies and performance indicators preference differed between the high and the low experience group. No significant difference was found for the variation in applied completion strategies, indicating that both more and less experienced operators show diversity in their completion strategies. For the different primary performance indicators, a trend was found for a difference in the importance of maintaining free track order (U = 17.0, p = .073) and unplanned stops of trains before signals (U = 20.5.0, p = .095). More experienced operators indicated these goals as being more important compared to the less experience operators. This is possibly because both goals are felt to be of importance to achieve a good performance in their train traffic operations, whereas less experienced controllers do not yet feel this.

DISCUSSION AND CONCLUSION

This study investigated the following research questions: to what extent do organizational and individual goals correspond? And what is the level of diversity in the goals and strategic mental models of train traffic operators given operators' work experience, and how does this relate to their performance?

First, the level of correspondence between organizational and individual goals was explored. This correspondence appeared to be moderate when looking at the median distribution. However, when assessing the preference for arrival punctuality, this goal ranked in the third position with 14% of the controllers adhering to the primary organizational goal. Operators indicated that they valuate departure punctuality (36%) and platform consistency (18%) as more important than the primary organizational goal arrival punctuality. The low absolute percentages spread over multiple goals revealed a strong diversity in operators' goal preference. A diversity between operators in completion orders was also found: as many as five different completion strategies were identified in the moderately to severely disrupted scenario. It is notable that in this study, the level of diversity in strategic mental models could not be related to worse or better performance.

The valuation of the controllers' goals was not reflected in their performance. The results show that in a moderate to severe traffic condition, controllers who highly value arrival punctuality showed more departure delay. Controllers who focused on departure punctuality had less arrival delay, and those who focused on a high level of platform consistency had less departure delay. Although these results do not confirm the expectations, they are in line with the fact that individual goals do not always lead to the system performance that corresponds to their personal goals. In fact, the presence of multiple and competing goals can be seen as characteristics of complex, ill-structured environments, as they have to be weighed and prioritized and compromises have to be made (Amelung & Funke, 2013; Funke, 1991; Hong, 1998). To obtain resilience, performance requires certain goals to take precedence over other goals (Woods, 2006). The moderate to severe traffic condition was a more complex situation and the controllers possibly had to make more compromises. These goals were probably not as conflicting in the less complex situation because the scenario did not cause a conflict between arrival and departure goals.

This study also revealed a difference in the valuation of the goals ‘maintaining free track order’ and ‘unplanned stops of trains before signals’ between the more and the less experienced operators: the former considered these goals to be more important. As such, more experienced operators appear to be more comfortable about satisfying lower prioritized organization goals. A trend was found for the level of experience impacting performance: less experienced controllers showed better arrival punctuality than experienced controllers when no complex disruptions were introduced (scenario 1). In contrast, the opposite trend was found when more train delays were introduced (scenario 2). The results of scenario 2 are in agreement with previous studies, following the line that more experienced controllers perform better in complex situations because of their well-developed mental models (Bogard, Liu, & Chiang, 2013).

Some limitations to this study should be mentioned. Although the simulator problems were controlled for, they nonetheless necessitated a small sample size. Also, given the length of each scenario, the number of conflict points per workstation was rather low. A limitation of this study in terms of goals trade-off consequences is that the level of violations was not assessed; we did not assess when a certain goal was violated during the simulator study or what the implications were of prioritizing one goal over another in these scenarios. These points should be taken into consideration in future studies.

Further, in line with the measured individual markers of resilience in this study, future research could investigate the diversity of strategic mental models in the actual work environment. The level of diversity of completion strategies between workstations could also be further investigated.

In sum, the primary organizational goal was not reflected at the operational level. An explanation for this might be the difference in the realization of operator’s goal versus the evaluation of operator’s performance. In an exemplary case, train traffic controllers may recognize arrival punctuality as both a primary organizational goal and an individual goal; however, due to external factors influencing the train traffic flow, a high arrival punctuality cannot be guaranteed by the operator alone. As such, operators may develop different preferences and coping mechanisms to better reflect their performance.

Although variability in cognition and behavior is both healthy and allowed, it can be argued that the revealed goals and strategic mental models of operators are too diverse and therefore unpredictable, and most probably weaken the resilience at the system level. These results could be used an indicator of brittle points that prevent the creation of a resilient organization (Gomes et al., 2009). It is observable that there are gaps between the work that is expected and the work that is done. Especially with the upcoming and planned large-scale changes in the railway system, it could be undesirable to continue with the redesign without involving the operational layer. Participatory design could be used as a joint approach to shape these changes (Falzon, 2008), enabling a new generation to work in a restructured work environment and to resonate these changes throughout all the levels.

Key points

- There is an incongruence between organizational and individual goals, indicating a gap between the work that is expected and the work that is done.
- The resilience of the Dutch railway system is low due to rather strong variations in the goals and strategic mental models as the behavior of operators becomes more unpredictable.
- In a more complex state of the traffic system, there is an incongruence between train traffic controllers' self-reported performance indicators and objective performance, possibly indicating goal competition.

References

Algemene Rekenkamer (2012). *Aanbesteden door NS-Railinfrabeheer* [Procurement by NS-Railinfrabeheer]. Retrieved from http://www.rekenkamer.nl/Publicaties/Onderzoeksrapporten/Introducties/2001/06/Aanbestedingen_door_NS_Railinfrabeheer_1995_2000

Amelung, D., & Funke, J. (2013). Dealing with the uncertainties of climate engineering: Warnings from a psychological complex problem solving perspective. *Technology in Society*, *35*(1), 32-40.

Back, J., Furniss, D., Hildebrandt, M., & Blandford, A. (2008). Resilience markers for safer systems and organisations. In M. Harrison & M.-A. Sujan (Eds), *Computer Safety, Reliability, and Security* (Vol. 5219, pp. 99-112): Springer Berlin Heidelberg.

Bary, E. (2015). *Procesleiding op het Spoor* [Traffic management in the railways] [Blog]. Retrieved from http://www.nicospilt.com/index_procesleiding.htm

Bogard, T., Liu, M., & Chiang, Y. H. V. (2013). Thresholds of knowledge development in complex problem solving: a multiple-case study of advanced learners' cognitive processes. *Educational Technology Research and Development*, *61*(3), 465-503.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience and School*. Washington, DC: National Academy Press.

Burnard, K., & Bhamra, R. (2011). Organisational resilience: Development of a conceptual framework for organisational responses. *International Journal of Production Research*, *49*(18), 5581-5599.

Chialastri, A., & Pozzi, S. (2008). Resilience in the Aviation System. In M. Harrison & M.-A. Sujan (Eds.), *Computer Safety, Reliability, and Security* (Vol. 5219, pp. 86-98): Springer Berlin Heidelberg.

Crawford, E., Toft, Y., & Kift, R. (2014). Attending to technology adoption in railway control rooms to increase functional resilience. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics* (Vol. 8532, pp. 447-457): Springer International Publishing.

Golightly, D., Sandblad, B., Dadashi, N., Andersson, A. W., Tschirner, S., & Sharples, S. (2013). A socio-technical comparison of rail traffic control between GB and Sweden. In N. Dadashi, A. Scott, J. R. Wilson, & A. Mills (Eds.), *Rail Human Factors: Supporting Reliability, Safety and Cost Reduction* (pp. 367–376). London: Taylor & Francis.

Dekker, S. (2006). Resilience engineering: Chronicling the emergence of confused consensus. In E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts* (pp. 77-92). Aldershot: Ashgate Publishing.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(1), 32-64.

Falzon, P. (2008). Enabling safety: issues in design and continuous design. *Cognition, Technology & Work*, *10*(1), 7-14.

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg, & P. A. Frensch (Eds.), *Complex Problem Solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Erlbaum.

Glaser, R., & Chi, M. (1988). Overview. In M. Chi, R. Glaser, & M. Farr (Eds.), *The Nature of Expertise* (pp. xv–xxxvi). Hillsdale, NJ: Lawrence Erlbaum.

Gomes, J. O., Woods, D. D., Carvalho, P. V. R., Huber, G. J., & Borges, M. R. S. (2009). Resilience and brittleness in the offshore helicopter transportation system: The identification of constraints and sacrifice decisions in pilots' work. Reliability *Engineering & System Safety, 94*(2), 311-319.

Hale, A., & Heijer, T. (2006). Is resilience really necessary? The case of railways. In E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts* (pp. 125-147). Aldershot: Ashgate Publishing.

Hollnagel, E. (2008). Preface: Resilience engineering in a nutshell. In E. Hollnagel, C. P. Nemeth, & S. Dekker (Eds.), *Resilience Engineering Perspectives: Remaining Sensitive to the Possibility of Failure* (pp. xi-xiv). Ashgate Publishing.

Hollnagel, E. (2014). Resilience engineering and the built environment. *Building Research & Information, 42*(2), 221-228.

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press.

Hong, N. S. (1998). *The relationship between well-structured and ill-structured problem solving in multimedia simulation* (Doctoral dissertation). Retrieved from: http://www.cet.edu/pdf/structure.pdf

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of management, 20*(2), 403-437.

Knieps, G. (2013). Competition and the railroads: A European perspective. *Journal of Competition Law and Economics, 9*(1), 153-169.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, *50*(2), 212-247.

Lengnick-Hall, C. A., & Beck, T. E. (2005). Adaptive fit versus robust transformation: How organizations respond to environmental change. *Journal of Management, 31*(5), 738- 757.

Lengnick-Hall, C. A., Beck, T. E., & Lengnick-Hall, M. L. (2011). Developing a capacity for organizational resilience through strategic human resource management. *Human Resource Management Review, 21*(3), 243-255.

Lo, J. C., Sehic, E., & Meijer, S. A. (2014). Explicit or implicit situation awareness? Situation awareness measurements of train traffic controllers in a monitoring mode. In D. Harris (Ed.). *Engineering Psychology and Cognitive Ergonomics*, 511-521.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology, 85*(2), 273-283.

Meijer, S. A. (2012). Introducing gaming simulation in the Dutch railways. *Procedia: Social and Behavioral Sciences, 48,* 41-51.

Meijer, S. A., Van der Kracht, P., Van Luipen, J. J. W., & Schaafsma, A. A. M. (2009). Studying a control concept for high-frequency train transport. *Infrastructure Systems and Services: Developing 21st Century Infrastructure Networks*, 1-6.

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010), Metaphor no more: A 15-year review of the team mental model construct, *Journal of Management, 36*(4), 876-910.

Mohammed, S., Klimoski, R., & Rentsch, J. R. (2000). The measurement of team mental models: We have no shared schema. *Organizational Research Methods, 3*(2), 123- 165.

Nemeth, C., Wears, R., Patel, S., Rosen, G., & Cook, R. (2011). Resilience is not control: Healthcare, crisis management, and ICT. *Cognition, Technology & Work*, *13*(3), 189-202.

NS (2015). *Jaarverslag 2014* [Annual report 2014]. Retrieved from http://nsjaarverslag.nl/FbContent.ashx/downloads/NS-jaarverslag.pdf

Popova, V., & Sharpanskykh, A. (2011). Formal modelling of organizational goals based on performance indicators. *Data & Knowledge Engineering, 70*(4), 335-364.

ProRail (2013). Laatste sporen aangesloten op verkeersleiding ProRail [Final tracks connected to ProRail control center]. Retrieved from http://www.prorail.nl/pers/persberichten/laatste-sporen-aangesloten-op-verkeersleiding-prorail

ProRail (2015a). *Jaarverslag 2014* [Annual report 2014]. Retrieved from http://www.jaarverslagprorail.nl/FbContent.ashx/downloads/ProRail-jaarverslag-2014.pdf

ProRail (2015b). *Programma Hoogfrequent Spoorvervoer* [High frequency railway transport program]. Retrieved from http://www.prorail.nl/projecten/goederenroute-oost-nederland/programma-hoogfrequent-spoorvervoer

Ramaekers, P, De Wit, T., & Pouwels, M. (2009). *Hoe druk is het nu werkelijk op het Nederlandse spoor? Het Nederlandse spoorgebruik in vergelijking met de rest* *van de EU-27* [How busy is it really on the Dutch railway tracks?]. Centraal Bureau voor de Statistiek.

Raser, J.C. (1969). *Simulations and Society: An Exploration of Scientific Gaming*. Boston: Allyn & Bacon.

Righi, A. W., Saurin, T. A., & Wachs, P. (2015). A systematic literature review of resilience engineering: Research areas and a research agenda proposal. *Reliability Engineering & System Safety, 141*(0), 142-152.

Rijksoverheid (2015). Mansveld: boetes voor NS en ProRail [Mansveld: Penalties for NS and ProRail]. Retrieved from http://www.rijksoverheid.nl/nieuws/2015/02/12/mansveld-boetes-voor-ns-en-prorail.html

Roth, E. M., & Patterson, E. S. (2005). Using observational study as a tool for discovery: Uncovering cognitive and collaborative demands and adaptive strategies. *How Professionals Make Decisions*, 379-393.

Salas, E., Stout, R. J., & Cannon-Bowers, J. A. (1994). The role of shared mental models in developing shared situational awareness. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.), *Situational Awareness in Complex Systems*, 297-304. Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Schaffernicht, M., & Groesser, S. N. (2011). A comprehensive method for comparing mental models of dynamic systems. *European Journal of Operational Research, 210*(1), 57- 67.

Sheridan, T. B. (2002). *Humans and Automation: System Design and Research Issues.* Santa Monica, CA: Wiley.

Steenhuisen, B., & De Bruijne, M. (2009). The brittleness of unbundled train systems: Crumbling operational coping strategies. *Second International Symposium on Engineering Systems*, 1–12.

Steenhuisen, B. (2009). *Competing Public Values: Coping Strategies in Heavily Regulated Utility Industries*. Enschede: Gildeprint Drukkerijen.

Sulmann, H. (2000). *Railverkeersleiding: van sein tot sein* [Railway traffic control: From signal to signal]. (Unpublished manuscript).

Veeneman, W. W. (2006). Punctualiteit, wat heet. [Punctuality: what does it mean?]. *Verkeerskunde: Vaktijdschrift over Verkeer en Vervoer, 57*(5), 58-59.

Veeneman, W. W., & Van de Velde, D. M. (2006). The value of bus and train: Public values in public transport. In Herman Maier (Ed.), *Proceedings of the European Transport Conference* (pp. 1-12). London: Association for European Transport.

Webber, S. S., Chen, G., Payne, S. C., Marsh, S. M., & Zaccaro, S. J. (2000). Enhancing team mental model measurement with performance appraisal practices. *Organizational Research Methods, 3*(4), 307-322.

Wilson, J. R. (2000). The place and value of mental models. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 44*(1), 49-52.

Wilson, J. R., Farrington-Darby, T., Cox, G., Bye, R., & Hockey, G. R. J. (2007). The railway as a socio-technical system: Human factors at the heart of successful rail engineering. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 221*(1), 101-115.

Woods, D. D. (2006). Essential characteristics of resilience. In E. Hollnagel, D. D. Woods & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts* (pp. 205-233). Aldershot: Ashgate Publishing.

Biographies

Julia C. Lo is a PhD candidate at the Faculty of Technology, Policy and Management of Delft University of Technology. She obtained her Master degrees in Social Psychology (2007) and Organizational Communication (2008) at the University of Amsterdam.

Kari R. Pluyter obtained her Master degree in Applied Cognitive Psychology at the University of Leiden in 2015.

Sebastiaan Meijer is a full professor at KTH Royal Institute of Technology, School of Technology and Health, and is associated with Delft University of Technology, Faculty of Technology, Policy and Management. He received his PhD in Gaming and Supply Chain Management in 2009 from Wageningen University.