

Explicit or Implicit Situation Awareness? Situation Awareness Measurements of Train Traffic Controllers in a Monitoring Mode

Julia C. Lo¹, Emdzad Sehic^{1,2}, and Sebastiaan A. Meijer^{1,3}

¹ Faculty of Technology, Policy and Management, Delft University of Technology,
The Netherlands

j.c.lo@tudelft.nl

² ProRail, The Netherlands

emdzaad.sehic@prorail.nl

³ Department of Transport Science, KTH Royal Institute of Technology, Sweden
smeijer@kth.se

Abstract. Railway traffic control faces the challenge of ensuring a high infrastructure capacity to maintain a constant train traffic flow. The current study assesses the situation awareness (SA), as a predictor of decision-making, of train traffic controllers to gain novel insights in their cognition. This study puts emphasis on levels of implicit and explicit situation awareness in a monitoring mode, through measures of SAGAT, MARS and performance. A human-in-the-loop simulator, called the PRL game is used to simulate the workspace of train traffic controllers. Initial findings indicate rather low levels of explicit SA, on the contrary to higher subjective SA scores through MARS and observer ratings, and a high performance on the punctuality and unplanned stops of trains.

Keywords: Situation Awareness, Implicit, Explicit, Train Traffic Control, Gaming Simulation.

1 Introduction

The railways in the Netherlands is characterized as the most dense and heavy utilized railway infrastructures in Europe [1]. The growing demand on diversified and higher frequency train schedules puts challenges on railway infrastructure innovations as well as on the implications of the work environment for train traffic and (regional and national) network controllers.

Railway infrastructure innovations are sought in processes as the increase of infrastructure capacity cannot be achieved by technical solutions alone [2]. Therefore, gaming simulations have been predominantly used so far as a research environment to test the viability of these process innovations. A gaming simulation can be seen as a simulation of a system, in which humans take part through game design methods and concepts such as immersion [3]. They comprise of different types, varying from

human-in-the-loop alike simulators to table top games. For the Dutch railways, human-in-the-loop simulators are currently used for individual operators, whereas board or table top gaming simulations are applied to study a larger part of the railway traffic operating systems (for examples see [2]).

One of the key elements to determine the viability of the various innovation processes is the quality of decision-making of operators. Ultimately their decisions impact their performance and therefore might impact the performance of the railway traffic system. The cognitive concept of situation awareness is often addressed as a predictor of decision-making in complex socio-technical systems.

Situation awareness (SA) has been most widely defined as 1. the perception of elements in the environment, 2. the comprehension of these elements and 3. the projection of these elements in the near future [4]. Situation awareness can be seen as a dynamic mental model of the situation, in which explicit and implicit levels of knowledge can be distinguished [5]. Explicit knowledge refers to active knowledge that resides in the working memory, while implicit knowledge refers to less active knowledge that cannot be inferred from queries or knowledge probes, being non-intentional, non-conscious and intuitive [6-8]. Following Croft, Banbury, Butler and Berry [6], implicit situation awareness can also be viewed as implicit processes in SA. Implicit processes are characterized as extremely durable, more robust in the light of competing attentional demands, and related to an increase in expertise. Measurements examples of implicit SA are described as comparisons of recalling probes (explicit SA) with performance-based measures or speed/accuracy measurements of elements (e.g. indicating hostile or friendly aircrafts) [6], [8-9].

The implications of cognitive concepts, such as situation awareness for decision-making, have been studied in a number of domains, e.g. aviation, nuclear power plants, military and firefighting [6]. However, previous studies identified different cognitive strategies across the domains. For example, air traffic controllers focus 90% of their time on processing information [10], while 95% of the tactical commanders in the military domain use a recognition decision strategy [11]. These different strategies can have implications on how the SA of operators is formed. Railway traffic controllers spend a significant amount of time in monitoring the train traffic flow. They operate in a fashion that requires only active involvement from the human operators when the delays of trains are affecting the train traffic flow, which might be caused by a failure in material, infrastructure or incidents. Findings from an ethnographical study on the decision-making of railway traffic controllers (i.e. train traffic and regional network controllers) reveal that operators do not always look in-depth into different consequences and have difficulties in making their reasoning explicit [12]. However, as far as known, no studies exist that investigate the SA of railway traffic controllers that make use of a traffic management system.

The findings from the ethnographic study raise the question how relevant cognitive constructs are formed, i.e. situation awareness, thus leading to the research question: *to what extent do train traffic controllers exhibit explicit or implicit levels of situation awareness?* The formulated question provides novel insights in the awareness levels of train traffic controllers and how this affects their performance in a monitoring mode of operations.

2 Method

2.1 Experimental Setting

As part of a suite of railway games, the human-in-the-loop simulator – PRL game – is used in this study. The inclusion of ‘game’ in its name, originated from previous design versions, which had abstract interfaces compared to the current version.

Aside from the current research focus, the general purpose of the study was formulated to investigate the impact of a gaming simulation session on the quality control processes of a new train timetable. Therefore, the session was also focused on the experience of the current human-in-the-loop simulator and to provide feedback regarding the new train timetable. Based on a set of gaming simulation components [2-3], Table 1 describes the characteristics of the gaming simulation NTTZ (new train timetable Zwolle).

Table 1. Characteristics of the PRL game NTTZ

<i>Core aspect</i>	<i>Description</i>
Purpose	Studying the impact of a game session on the quality control processes of a new train timetable
Scenarios	Two for each participant: 1. 2013 train timetable, 2. 2014 train timetable
Simulated world	Detailed infrastructure; detailed timetable; limited options in number of actions; larger area of train traffic operations (merged workspace of Zwolle station east-side and Zwolle station west-side)
# of participants	1 per session
Roles	Train traffic controller
Type of role	Similar to their own roles
Objectives	Execution of tasks – similar as to in their daily work
Constraints	Exclusion of roles outside the defined infrastructure area, exclusion of train driver, no large disruption
Load	Average train delays
Situation (external influencing factors)	Presence of individual observers seated next or near the participant, facilitators
Time model	Continuous

Both scenarios were designed together with subject matter experts to simulate a light disruption by mildly delayed trains, which was based on the realization data, i.e. average delays of trains in a month of that year with an average amount of disruptions. Further on, the first scenario focused on the 2013 timetable for participants to familiarize with the simulated environment and to obtain a base rating of SA and performance of train traffic controllers. Similarly, SA and performance were measured as well in the 2014 scenario.

As indicated in Table 1, the simulated workspace was represented by two merged workspaces surrounding the station of Zwolle with the borders of each workspace including smaller stations in their vicinity. In most of the cases, one train traffic controller is responsible for monitoring and controlling the train traffic flow at one workspace. Together, these workspaces form with two other workspaces (Hengelo and Deventer) the regional traffic control center in Zwolle.

Participants. Eleven train traffic controllers from one regional traffic control center in Zwolle took part in both scenarios. Train traffic controllers were selected based on their competence to operate at the simulated workspace.

2.2 Materials

A number of background questions were presented before each session: *work experience in the railway sector*, *work experience in the current job function*, *perceived experience of the workspace*, *perceived competences in comparison to peers*, *motivation in participation of the PRL game*. The latter three items were measured on a five-point Likert scale, varying from ‘fully disagree/strongly unexperienced’ to ‘fully agree/strongly experienced’.

Multiple situation awareness measurement methods were used to triangulate measures of SA. Firstly, at the end and two times during each scenario, the gaming simulation was frozen, in which participants receive 22 queries in total, in accordance with the *Situation Awareness Global Assessment Technique (SAGAT)*. In total three pauses were introduced, where seven to eight SAGAT questions were presented with a multiple-choice answering format (e.g. [13]). The pauses were planned after possible conflicting choices in the train traffic flow. The SAGAT queries were based on a goal-directed task analysis (GDTA) (e.g. [9]) from a national network controller. A subject matter expert translated the relevant SA requirements to a number of queries for a train traffic controller. Examples of the SAGAT questions are displayed in Table 2.

Table 2. Examples of the SAGAT queries

<i>SA level</i>	<i>Query example</i>
1	At which track does train 13828 arrive in Zwolle? Track 14, track 15, track 16
2	Which train leaves first according to planning from station Zwolle? [Train number] 12522, 3629, 9119
3	How is the track capacity at 7:46 in station Zwolle? 6 tracks free, 5 tracks free, 4 tracks free, other, namely: ..

Four SA probes were removed during the analysis, which reduced the number of SAGAT queries to 19 for each scenario. For the analysis of the results, the items were firstly scored as correct/incorrect, and then the percentage of correct answers was calculated.

Subjective SA ratings were collected through the *Mission Awareness Rating Scale* (MARS) [3], [14] and presented at the end of each scenario. The selected MARS questions can be seen as the *perceived own situation awareness* and are related to the three SA levels as identified by Endsley [4], and respectively scored on a four-point scale, varying from ‘fully disagree’ to ‘fully agree’. The average of the three levels of SA was calculated for the results.

Additionally, MARS-based questions were used to measure the *observed situation awareness*. One subject matter expert was present during all sessions, and rated the observed situation awareness with similar questions and a similar scale as the *perceived own situation awareness*. An observation sheet, which is also used by instructors during training session was provided as a guideline to rate the observed situation awareness of the participants. Similarly, the average of the three levels of SA was computed for the results.

In the railway sector, performance is measured on system level. That is, no official objective measurements for the individual performance of train traffic controllers exist. The performance indicators ‘punctuality’ and ‘unplanned stops’ were identified in consultation with the railway performance & analytics department. These results were retrieved from log files of the PRL game. Punctuality can be defined by the entry and exit times of trains within a specific region for a specific workspace. The punctuality of trains is often measured in percentages over a certain amount of time and for a certain level of delay. As the performance indicator for the Dutch railway infrastructure organization is set at three minutes, this is also used in the current setup. Secondly, unplanned stops are defined as the number of times that train drivers encounter an unplanned red signal. Unplanned stops can be seen as hazardous for safety, therefore the reduction of unexpected red signals is wished to be achieved.

The human-in-the-loop experience of train traffic controllers was captured by measurements of *gaming simulation validity* [15] at the end of each session. In line with Raser [15], three out of four components of gaming simulation validity were included, namely structural validity (similarity in structure between the simulated and reference systems), processes validity (similarity in processes between the simulated and reference system), and psychological reality (the degree to which the participants/players perceive the simulated system as realistic). Three items measured structural validity ($\alpha = .60$). An example of an item was: ‘I can apply the information from the information sources in the simulator in a similar way as in the real world’. Process validity was measured as well by three items ($\alpha = .90$), e.g. ‘the train traffic flow in the simulator is similar in their processes to the real world train traffic flow’. Thirdly, psychological reality was measured by seven items ($\alpha = .84$), e.g. ‘the train model is sufficiently realistic for the current task’. All items were measured on a five-point Likert scale, varying from ‘fully disagree’ to ‘fully agree’.

Five *workload* items from the NASA-TLX [16] were presented after each scenario ($\alpha = .64$). One item related to the physical demand was removed as this was unnecessary to measure given the task at hand. In line with the scales used at other items, a five-point Likert scale was applied.

Lastly, the difficulty between the two scenarios was measured on a similar five-point Likert scale in the post-session questionnaire, in which the statement was provided: 'I was able to quickly get accustomed to the new timetable'.

2.3 Procedure

The sessions took place over three days. All train traffic controllers conducted two scenarios: one scenario with the current 2013 train timetable and one scenario with the new 2014 train timetable, in which the length of each scenario was 35 minutes. At the start of the session, participants firstly received instructions on the possibilities and limitations of the PRL game. Additionally, they were also asked for permission to record the session by video. Subsequently, participants received the pre-questionnaire. 15 minutes after the start of the scenario, the PRL game was paused, in which participants were asked to turn to the desk behind them and answer the SA probes. Two more freezes of the PRL game followed every 10 minutes, in which the third pause marked the end of the scenario. Similar halts in the scenarios were introduced in the second scenario, which ended with a post-questionnaire. In both scenarios, observers asked questions and feedback about the usability of the PRL game, their decision-making and their preferences related to details in the timetables, during segments of the scenarios where train traffic controller was monitoring the train traffic flow. Conversations during the operator's task in a non-severe disruption were planned in the procedure as they are seen as consistent with the displayed behavior of train traffic controllers in their work environment.



Fig. 1. Two PRL game sets: in the foreground the PRL game set with the new timetable, on the background the set with the current timetable

3 Results

Ten male train traffic controllers and one female operator took part in the sessions. Their work experience within the current job function was 16.5 years ($SD = 8.9$), however the average of their overall work experience in the railway sector was higher, $M = 20.8$, $SD = 9.8$. Participants indicated that they perceived their competence on the current workspace as high ($M = 4.3$, $SD = .65$), as well in comparison to colleagues within the regional traffic control center ($M = 3.8$, $SD = .75$). A high interest was indicated to participate in the PRL game ($M = 4.4$, $SD = .92$).

3.1 Perceived Difficulty of the Scenarios and Workload

As expected, the second scenario was not perceived as challenging to the train traffic operators, as they were able to quickly get accustomed to the new timetable ($M = 4.2$, $SD = 2.1$). Additionally, participants indicated that they perceived a low workload ($M = 1.7$, $SD = .47$). Qualitative data obtained during the session supported both results.

It should be remarked that the scenario was initially designed to simulate a light disruption. However, train traffic controllers perceived the delays of the trains as not sufficiently problematic to interfere with the train traffic flow, possibly due to the fact that the automation function of the train traffic system ('ARI') was not triggered by the current train delays to indicate red highlights on the timetable screen (i.e. ARI is not able to manage to execute the train path assignment of a train). Therefore, there was less interaction between the operators and the PRL game than expected, making it a monitoring mode instead.

3.2 Gaming Simulation Validity

The PRL game is validated for the use of the current task through the gaming simulation validity dimensions, structural validity, process validity and psychological reality. The results indicate a rather positive perception of the PRL game by the train traffic controllers (see Table 3). Qualitative data support this notion as well. The rather positive scores on the three validity types also assert the assumption that levels of situation awareness and performance in the simulated environment should be comparable to a similar task in their real work environment.

Table 3. Gaming simulation validity dimensions of the PRL game for the current task

	<i>N</i>	<i>M</i>	<i>SD</i>
Structural validity	11	3.6	.53
Process validity	11	3.7	.71
Psychological reality	11	3.8	.53

3.3 Situation Awareness and Performance

Results from measurements on situation awareness and performance are depicted in Table 4. The lower number of participants in the scenarios can be subscribed to missing, unclear or discarded data (e.g. by deviation from the probes instructions).

Table 4. Measurements of situation awareness and performance

	<i>Scenario 1</i>			<i>Scenario 2</i>		
	N	M	SD	N	M	SD
SAGAT (%)	9	44.4	17.68	9	37.11	11.07
Perceived SA (1-4)	8	3.1	.59	11	3.3	.49
Observed SA (1-4)	11	3.5	.43	10	3.7	.41
Punctuality (%)	11	99.3	1.20	11	98.6	1.64
Unplanned stops	11	2.2	.87	11	2.4	1.29

As the variables punctuality, perceived situation awareness, observed situation awareness (second scenario) and unplanned stops (first scenario) were significant on the Kolmogorov-Smirnov statistic for normality, a Wilcoxon test was conducted to analyze differences between scenarios in scores. A significant difference was found for the observed SA scores ($Z = -2.33, p = .02$).

The values in the table indicate rather low explicit SA levels measured through SAGAT. However, subjective ratings by participants themselves and observers show high to very high levels of situation awareness. Also for punctuality, the results indicate near optimal achievements by the train traffic controllers.

Correlations were drawn between the variables to investigate the relation between situation awareness and performance. Relevant trends were found for a relation between explicit situation awareness and punctuality in the first scenario ($\rho = .64, p = .06$); a higher explicit situation awareness leads to a higher punctuality of trains. Additionally, trends were found for the perceived situation awareness and both performance indicators punctuality and unplanned stops in the second scenario (respectively ($\rho = .53, p = .09$; $\rho = -.52, p = .10$)). A higher perceived SA leads to a higher punctuality, and a higher perceived SA leads to less unplanned stops of trains.

The findings reveal that train traffic controllers perceived their SA as high and show a high performance in the current monitoring mode. However, the train traffic controllers score reasonably low on objective (explicit) measures of SA, which is in line with the implication for the presence of implicit situation awareness. This finding might be supported by the results on the SA probes that are categorized by SA level (see Table 5).

The results show that SA level 1 items were fairly low and beneath level 2 scores, although absolute values of the level 2 items remained rather low. Theoretically, SA would drop with each SA level, i.e. operators set a baseline for their SA by the elements they perceive in the situation. Following this, their comprehension (SA level 2) is equal to lower than their perception, and similarly their projection (SA level 3) is equal or lower than their comprehension.

Table 5. SA probes per SA level

	<i>Scenario 1</i>		<i>Scenario 2</i>	
	# total items (N=9)	% correct	# total items (N=9)	% correct
SA level 1	99	37	108	39
SA level 2	54	65	45	42
SA level 3	18	39	18	17

Additionally, implicit SA could also be measured by using explicit SA measures and relating this to certain conflicting choices in the scenario. However, since the operators did not show much interference with the expected conflicts in the scenario and the train traffic flow in general, no implications could be drawn with regards to their SA queries and behavior in the scenarios.

Explorative analyses were conducted to investigate the potential role of individual differences between train traffic controllers. A trend was found for a negative correlation between the work experience in the railway domain and percentage of correct SAGAT answers; $r = -.65$, $p = .06$; a higher experience in the railway domain lead to a lower explicit situation awareness.

4 Discussion and Conclusion

The current study attempts to investigate the level of explicit and implicit situation awareness at train traffic controllers. The findings show rather low levels of explicit SA, on the contrary to (very) high subjective SA scores through MARS and observer ratings, and high performance on the punctuality and unplanned stops of trains. It is possible that the low explicit SA scores are influenced by the fact that the light disruption was not seen as that problematic and operators therefore portrayed monitoring modes of operation. This changed mode of operation might also have been affected by the role of automation as train traffic controllers rather would rely on and are triggered by the automation function of the train traffic system, causing a low explicit SA. Similar findings with low SAGAT scores have been found in earlier studies where no active decision-making was taking place [17]. Another possible explanation for the low explicit SA might be related to relevance of the presented SA queries in the unexpected changed mode of operation. It is possible that certain information is less relevant as different goals are achieved in certain circumstances, e.g. operators search for deviations and irregularities in the train traffic flow.

Further on, a trend is found for the negative relation between work experience and explicit SA. This result might be in line with the notion from previous studies that an increased implicit SA is more common for expert operators, e.g. [6]. Additionally, trends were found for the relation between objective (explicit) ratings of situation awareness and performance, but not found for both scenarios however, limiting the generalizability of these results.

Another limitation of the study is that no official performance indicators could be used, as these do not exist within the railway infrastructure organization. As critical remarks can be drawn for both performance indicators, more investigation is needed

to determine the theoretical and computational implications of individual performance indicators in the railway domain.

Nonetheless, these findings reveal novel descriptions on the situation awareness of train traffic operators in a monitoring mode. Further research is needed to identify the levels of explicit situation awareness in high disrupted conditions and to explore its relationship with levels of expertise.

Acknowledgements. This research was funded through the Railway Gaming Suite program, a joint project by ProRail and Delft University of Technology.

References

1. Ramaekers, P., De Wit, T., Pouwels, M.: 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 (2009)
2. Meijer, S.A.: Introducing gaming simulation in the Dutch railways. *Procedia - Social and Behavioral Sciences* 48, 41–51 (2012)
3. Lo, J.C., Meijer, S.A.: Measuring group situation awareness in a multiactor gaming simulation: A pilot study of railway and passenger traffic operators. In: *Proceedings of the 57th Human Factors and Ergonomics Society Annual Meeting*, pp. 177–181 (2013)
4. Endsley, M.R.: Design and evaluation for situation awareness enhancement. In: *Proceedings of the 32nd Human Factors Society Annual Meeting*, pp. 97–101. SAGE Publications, Boston (1988)
5. Adams, M.J., Tenney, Y.J., Pew, R.W.: Situation awareness and the cognitive management of complex systems. *Human Factors* 37, 85–104 (1995)
6. Croft, D.G., Banbury, S.P., Butler, L.T., Berry, D.C.: The role of awareness in situation awareness. In: Banbury, S., Tremblay, S. (eds.) *A Cognitive Approach to Situation Awareness: Theory and Application*, pp. 82–103. MPG Books Ltd., Bodmin (2004)
7. Endsley, M.R.: The role of situation awareness in naturalistic decision making. In: Zsombok, C.E., Klein, G. (eds.) *Naturalistic Decision Making*, pp. 269–283. Lawrence Erlbaum Associates, Mahwah (1997)
8. Gugerty, L.J.: Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied* 3, 42–66 (1997)
9. Endsley, M.R.: Direct measurement of situation awareness: Validity and use of SAGAT. In: Endsley, M.R., Garland, D.J. (eds.) *Situation Awareness Analysis and Measurement*, pp. 147–174. LEA, Mahwah (2000)
10. Kaempff, G.L., Orsanu, J.: Current and future applications of naturalistic decision making in aviation. In: Zsombok, C.E., Klein, G. (eds.) *Naturalistic Decision Making*, pp. 81–90. Lawrence Erlbaum Associates, Mahwah (1997)
11. Kaempff, G.L., Klein, G., Thordsen, M.L., Wolf, S.: Decision making in complex naval command-and-control environments. *Human Factors* 38, 220–231 (1996)
12. Steenhuisen, B.: *Competing Public Values: Coping Strategies in Heavily Regulated Utility Industries*. Gildeprint Drukkerijen, Enschede (2009)
13. Strater, L.D., Endsley, M.R., Pleban, R.J., Matthews, M.D.: Measures of platoon leader situation awareness in virtual decision-making exercises. U.S. Army Research Institute for the Behavioral and Social Sciences (2001)

14. Matthews, M.D., Beal, S.A.: Assessing situation awareness in field training exercises, U.S. Army Research Institute for the Behavioral and Social Sciences (2002)
15. Raser, J.C.: *Simulations and Society: An Exploration of Scientific Gaming*. Allyn & Bacon, Boston (1969)
16. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) *Human Mental Workload*, pp. 139–183. North-Holland, Amsterdam (1988)
17. Endsley, M.R., Rodgers, M.D.: Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errors and automation (No. DOT/FAA/AM-97/13). Federal Aviation Administration Office of Aviation Medicine (1997)