Investigating Strategies in Rail Signalling: Comparison of Simulation Methods

Nora BALFE[1], Robert J. HOUGHTON[2], Jockman CHEUNG[2] and Sarah SHARPLES[2]

[1]*Centre for Innovative Human Systems, Trinity College Dublin, Ireland*
[2]*Human Factors Research Group, University of Nottingham, Nottingham, UK*

Abstract. Two simulation methods used to study signaller decision-making strategies are described and compared through experimentation. The first method was a 'static scenario' method in which expert participants were presented with printouts from a simulated scenario and asked to give their strategy for routing the trains in the area. The second was a more traditional dynamic simulation undertaken on the same high fidelity simulator. The first method allowed greater control over the experiment and presents interesting opportunities for collecting additional qualitative data from participants, but the second method was more realistic and featured improved participant performance.

Keywords. Simulation, decision-making, rail human factors, expert strategies

1. Introduction

As increasing levels of automation are proposed for control systems such as rail signalling control, there is a need to understand how signallers currently manage networks in order to facilitate the design of future collaborative systems that support operators in appropriately managing their network and improving performance. The research reported here used two different simulation methods to elicit strategies from railway signallers in support of this goal.

Previous research in this area has examined signaller interaction with the current generation of automation on the GB network (Balfe, Sharples, Wilson, 2015; Balfe, Wilson, Sharples & Clarke, 2012) and qualitative research has been undertaken in the area of signaller routing strategies (e.g. Patrick, Balfe, Wilson & Houghton, 2013; Kauppi et al., 1996). The study by Patrick et al. (2013) highlighted the wide range of information required to generate strategies, including train, infrastructure, situation and service knowledge. Other research in the rail domain has identified that experienced signallers develop strategies to manage trains during instances of disruption, e.g. late running (e.g. Haavisto et al, 2010), but research examining the factors involved in these strategies has tended to focus on particular aspects, such as location prediction or decision support (e.g. Lenoir, 1993; Andersson et al., 1997). The research presented in this paper attempted to use more quantitative methods to understand signaller strategies more broadly. Due to the constraints of researching in a live, safety-critical environment, it was not possible to conduct in-depth quantitative research in the live signal box, so simulation methods were deemed more appropriate. The use of simulation and scenarios in rail knowledge/strategy elicitation is not currently widespread, with interview and observation methods being more common approaches (Cooke, 1994). Simulation offers a more controllable method to collect empirical data on expert task performance.

Simulation types vary, from full mission simulators that accurately replicate the environment, system behaviour, and even sound and motion through to prototypes or desktop simulations that replicate only the most salient cognitive or procedural elements of a task. Two simulation methods were used in the research reported here: one dynamic with participants interacting with a high fidelity simulator and the other static using paper scenarios. The static scenario method presented in this paper borrows elements from a limited information approach to knowledge elicitation (Shadbolt, 2005), in that information was deliberately added during the course of the experiment to observe the effect on strategies. Each of the two experiments will be discussed in sequence and the paper will conclude with a comparison of the techniques and the results obtained from each. The results of each experiment are not discussed in detail, but simply enough to compare the two methods.

2. Experiment 1: Static Simulation

This experiment was based around six static scenarios that were prepared on a high fidelity simulator. Screen prints were taken from the simulator, and supporting information in terms of additional train running information was mocked up. The participants were shown the printouts and asked to generate a strategy for the trains present in each scenario. Information was slowly added, and changes to the signallers' original strategy were noted in order to understand the effect of the additional information on the strategies.

2.1 Method

Nine experienced, male signallers took part in the experiment. All currently operated the workstation used in the scenarios and their experience ranged from 4 to 38 years, with a mean of 15 years. A high fidelity signalling simulator was used to create the scenarios for the experiment that were presented to the participants on paper. The printouts were annotated with the participants' strategy and different colours were used as each additional piece of information was presented. The scenarios were created with the assistance of an experienced supervisor at the control centre where the experiment was held. They were based on real situations that had occurred and caused significant delays, and several incorporated 'regulation statements' or procedures for the order of trains through a junction in particular circumstances. The participants signed a consent and demographics form and were assured of anonymity and right to withdraw. The printouts were then presented and the participant given 60 seconds to generate a strategy for routing trains in that situation, which they then gave verbally and drew on a blank schematic provided. Additional information on the scenario was then provided in the order: time of day, timetable extract, delay information. After each piece of additional information, the participant was given the opportunity to change their strategy, but not feedback was given on the quality of each strategy presented. The results were analysed by comparing the strategies proposed by participants to pre-prepared 'recommended' strategies. These were developed during the scenario selection by an experienced supervisor and were designed to give the correct priority to the trains, minimise delay and obey regulation statements. Each proposed strategy was rated against the scheme described in Table 1.

Table 1: Strategy rating scheme

Good Strategy	Closely follows the recommended strategy Minimal delay incurred if strategy is implemented Justified reasoning for using the strategy
Acceptable strategy	Parts of the implemented strategy follow the recommended strategy Some delay incurred if strategy is implemented Obeys basic signalling standards/reasoning

Poor Strategy	Does not follow recommended strategy Significant delay incurred if strategy is implemented Unjustified/poor reasoning for using the strategy

2.2 Results

The results are discussed in terms of the variation in number of strategies and the quality of the proposed strategies. Figure 1 describes how the quality of the proposed strategies improved across all six scenarios with the addition of information. Only a small improvement can be seen with the addition of time information. A much larger improvement is seen with the addition of timetable information, and again a small improvement can be seen with the addition of train running information.

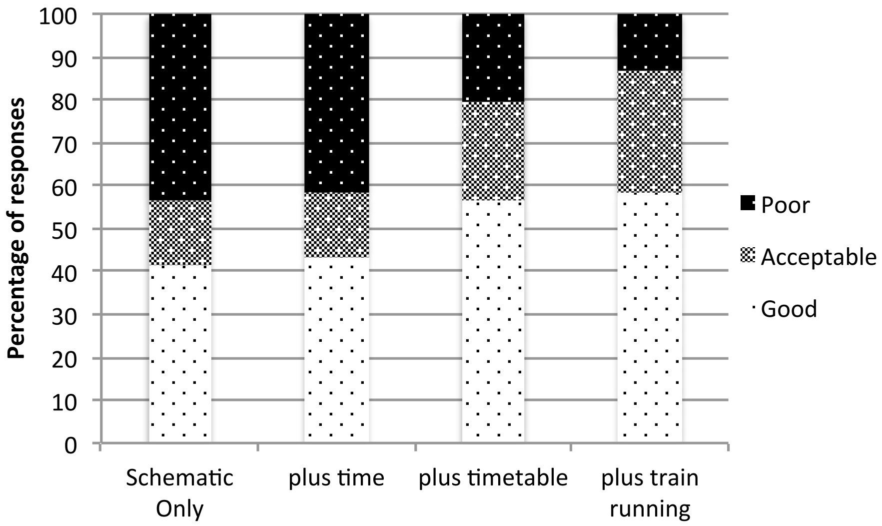

Figure 1: Response Quality (Exp.1)

Table 2 describes the variation in strategies, with a maximum of seven different strategies proposed for Scenario 6. With the exception of Scenario 1, all scenarios showed a consistent or decreasing number (i.e. less variation) of strategies as more information was added, suggesting that the increasing information helped participants coalesce towards similar strategies.

Table 2: Strategies observed (Exp. 1)

	Scen.1	Scen.2	Scen.3	Scen.4	Scen.5	Scen.6
Total Strategies	*5*	*4*	*4*	*2*	*5*	*7*
Screenshots only	1	4	4	2	4	6
Plus time	1	3	4	2	4	6
Plus timetable	3	2	3	2	4	4
Plus delays	5	2	3	2	4	4

3. Experiment 2: Dynamic Simulation

This experiment was a more traditional full simulation on a high fidelity signalling simulator. A 40-minute simulation, on the same geographic area as the first experiment, was created but featuring different conflicts. This simulation contained six different conflicts between trains.

Each participant ran the experiment and was then debriefed while viewing a recording of their simulation in order to gather qualitative data on the reasons for their routing decisions.

3.1 Method

A new cohort of nine male, experienced participants took part in this experiment. All currently operated the workstation used in the scenarios and their experience ranged from 4 to 23 years, with a mean of 12 years. A high fidelity signalling simulator based in the control centre was used for the study. Participants signed a consent form and were assured of anonymity and right to withdraw before filling in a demographic questionnaire. The participant was then given a short time to familiarise himself with the simulator before starting the simulation. During the simulation, the experimenter noted the order of trains through junctions and the prioritisation given to trains. After the 40-minute simulation was complete, the participants were given a brief break before watching a replay of their simulation. The participant was asked to explain each of the decisions made on train priority and the reasons behind them. The decisions were analysed using the same approach and criteria as Experiment 1.

3.2 Results

Table 3: Strategies Observed (Exp. 2)

	No. Strategies
Conflict 1	6
Conflict 2	2
Conflict 3	2
Conflict 4	2
Conflict 5	2
Conflict 6	2

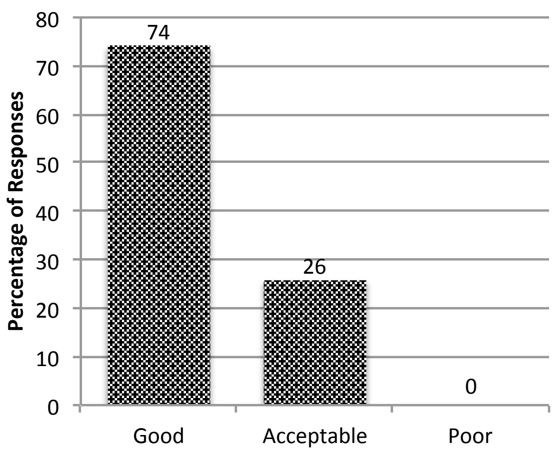

Figure 2: Strategy Quality (Exp. 2)

The results of Experiment 2 are again discussed in terms of variation and quality of the strategies used by participants. The variation in strategies was much lower in this experiment (Table 3). Apart from the first conflict, all conflicts elicited only two different strategies. This reflects the fact that individual conflicts were assessed in this experiment, as opposed to the full scenarios evaluated in Experiment 1. The quality of the strategies elicited was much higher in the dynamic experiment, with almost 75% of strategies being graded as good. No poor strategies emerged from this experiment.

4. General Discussion

There were noteworthy differences in the results obtained between the two methods, both in terms of variation in strategies utilised and in terms of the quality of the strategy. Two reasons are proposed for these differences; first, the reduced fidelity of the static scenarios may have led to a reduced performance. If this is the case, this approach may not be appropriate for realistic, in-depth study of operator strategies, due to the potentially significant effect on performance. However, the exaggerated differences between signallers may be beneficial when looking for small differences in the way operators generate strategies, or the relative importance of different elements of the task.

The second possible explanation is that the static scenarios asked the participants to project their decision-making further into the future, whereas in the dynamic experiment they were able to change or adjust their decisions as the simulation progressed and they received feedback on the impact of their previous decisions. This suggests that signallers do not accurately predict the situation ahead of time, but adjust their plan as the scenario unfolds, i.e. cognition distributed over time (Hollan, Hutchins, & Kirsh, 2000).

From a research perspective, the static scenarios allowed the total strategy to be noted and analysed, while additional decisions over time are more hidden in the dynamic experiment and are thus harder to identify. The post experiment debrief can be used to elicit some of this information, however it is very dependent on memory and highly skilled operators using expert strategies may not actively process and encode their decisions, and therefore may be unlikely to accurately recall or verbalise the reasons for each decision. The comparative advantages and disadvantages of the two methods are shown in Table 4.

Table 4: Comparison of the two methods

	Advantages	**Disadvantages**
Static	More targeted analysis Easy to set up and administer Portable Easier to elicit further information Repeatable	Less realistic Inaccurate measurement of performance
Dynamic	Higher ecological validity More integrated scenarios possible Fuller data	Difficult to analyse More resource intensive Less accurately repeatable

The static scenario experiment proved an interesting approach to this type of research, and offered huge potential to control the experiment and draw out strategies. However, the experiment reported here was the first application of this type of methodology by this team, and improvements on the methodology could be made. For example, a more structured approach to collecting supporting qualitative data on the reasons for strategies would have added richer data, and additional measures including time taken to develop a strategy would have added more useful information.

In both methods we attempted to move away from protocol analysis and detailed interview techniques. Although some qualitative data was collected, this was to support the rating of strategies, rather than detailed investigation of the reasons behind the strategies. This approach has left us with many questions, particularly around the variance in strategies exhibited in the static scenario experiment. It is therefore our conclusion that qualitative data is necessary in this field and future studies should incorporate a higher level of qualitative data collection.

5. Conclusions

Both methods provided useful results and insights into the signaller decision-making process. The paper-based scenario method shows great potential for eliciting information on decision making as it allows the researcher to better control the scenario and the information presented and facilitates further elicitation of information through probe techniques immediately following the scenario presentation, which would be disruptive in a full simulation. It may also be more practical for early exploratory studies, since it is less resource intensive and transportable. However, it is not suitable for accurately assessing performance and is therefore recommended as an exploratory method. Conversely, the dynamic simulation had higher ecological validity and the ability to capture real world dynamic decision making, making it more representative of real performance and issues but difficult to elicit detailed information on individual steps in the process. Both methods would benefit from strongly integrated qualitative data collection to provide a rich description in support of the quantitative metrics.

Acknowledgements

The authors wish to acknowledge Network Rail for the access granted to participants, assistance with scenario design and running, and the simulator itself.

References

Andersson, A.W., Sandblad, B., Hellstrom, P., Frej, I., & Gideon, A. (1997). A systems analysis approach to modelling train traffic control. In Proceeding of WCRR 97, pp. 16-19.

Balfe, N., Wilson, J.R., Sharples, S. & Clarke, T. (2008). Structured Observations of Automation Use. In P.D. Bust (Ed.), Contemporary Ergonomics 2008, (pp. 552-557). London: Taylor & Francis.

Balfe, N., Wilson, J.R., Sharples, S., & Clarke, T. (2012). Development of design principles for automated systems. Ergonomics, 55(1), pp. 37-54.

Cooke, N.J. (1994). Varieties of knowledge elicitation. International Journal of Human-Computer Studies, 41, pp. 801-849.

Haavisto, M., Ruuhilehto, K., Oedewald, P. (2010). Train traffic controller’s task demands during major railroad construction work. European conference on Human Centred Design for Intelligent Transport Systems, Berlin, Germany, April 29-30.

Hollins, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. ACM Transactions on Human-Computer Interaction (TOCHI), 7(2), 174-196.

Kauppi, A., Wikstrom, J., Sandblad, B., Andersson, A.W. (2006). Control strategies for managing train traffic – Difficulties today and solutions for the future. Technical report no. 2006-24. Department of Information Technology at Uppsala University.

Lenoir, T.M.J. (1993). Analyses of cognitive processes in train traffic control. Ergonomics, 26(11), 1361-1368.

Patrick, C., Balfe, N., Wilson, J.R., & Houghton, R. (2013). Signaller information use in traffic regulation decisions. In N. Dadashi, A. Scott, J.R. Wilson, & A. Mills (Eds.), Rail Human Factors: Supporting Reliability, Safety and Cost Reduction (pp. 409-418). London: Taylor & Francis.

Shadbolt, N. (2005). Eliciting expertise. In: Wilson, J.R. & Corlett, E.N. (Eds.), Evaluation of Human Work: A Practical Ergonomics Methodology. 3rd ed. CRC Press: Boca Raton, pp. 185-218.