IRSE News

Institution of Railway Signal Engineers

March 2021

Automation
in control centres

Driver fatigue
and automatic train operation

Rail and the IoT
intelligent railways

Effects of automatic train operation on regional train drivers


Lucie Pannecoucke and Roger Dällenbach

Within their Automatic Train Operation (ATO) project, Switzerland's Schweizerische Südostbahn AG (SOB) has performed a systematic human factor study of the effects of Grade of Automation 2 (GoA2) on regional train drivers. With this level of automation, tasks such as accelerating and braking are removed from the driver's workload and are performed by an automated system. New tasks are created such as additional visual supervision and taking over control in the event of issues arising. This human factors study aims at researching effects of the automation on train drivers' fatigue, workload and performance.

This study was performed in collaboration with AWK Group and the ZHAW (Zürcher Hochschule für Angewandte Wissenschaften – Zurich University of Applied Sciences).

State of research: fatigue in train drivers

Prolonged manual train driving has been significantly researched (Filtness and Naweed, 2017, Stein and Naumann, 2016). It can lead to cognitive underload (Filtness and Naweed, 2017). This underload leads to passive fatigue (theory from Desmond and Hancock, 2001), in which continuously low workload increases the fatigue of the train driver. The negative effects of fatigue on the performance of train drivers have also been established (Filtness and Naweed, 2017).

The automation of train driving leads to more vigilance tasks such as information acquisition and visual monitoring (Branderburger & Hörmann, 2016) whilst reducing the control tasks. It therefore contributes to a further underload of the train driver, as automation takes over handling the accelerating and braking tasks from the train driver. With automatic train driving, the train driver is required to perform monitoring tasks rather than directly driving the train.

It has been established that performance in monitoring tasks can decrease after 30 minutes (Mackworth, 1948). In the case of train drivers, it has been shown that the "performance in degraded operations decreases if train drivers execute their operational tasks with high level of automation compared to low levels of automation" (Brandenburger and Jipp, 2017).

In summary, the current state of research raises a risk that with increased automation, the train drivers might experience a cognitive underload that leads to passive fatigue and under-performance. Yet these studies do not fully answer the needs of train operation companies today, as two main operational characteristics of the SOB lines differ from the conditions used in the research environment:

- The degradation of performance is calculated between manual train operation and GoA2 level. In most modern trains, today's drivers are already using a certain level of automation with the use of a speed regulator.
- The degradation of performance is researched on high-speed trains, whereas the SOB is operating regional services. With urban trains, tasks such as opening and closing doors remain allocated to the drivers and occur every 4 minutes on average, compared to every 30 to 60 minutes for high-speed trains. Train driving in urban operation therefore represents a higher task load than for high-speed trains, the effects of which on the fatigue and performance have not yet been established.

The project's targets and method

The SOB study aims to verify the degradation of performance of train drivers by adding a comparison between driving with GoA2 and driving under the current operating conditions; GoA1 with speed regulator. In addition, this study focuses on train operation representing urban trains with regular stops.

For this study, 31 train drivers were divided into three groups as represented in Table 1.

In all three groups, automatic train protection ETCS L1 Limited Supervision (LS) was active.

A train simulator was used (Figure 1) and the track Biberbrugg – Arth Goldau (back and forth) was video-recorded and used for the simulated drives. Each driver underwent six or seven drives as follows:

Figure 1 – The simulator used for the research.

Group	Driving mode	Description	Participants
1	GoA1 without speed regulator	The drivers in this group drove without automation of the driving and braking system. Driving and braking were tasks manually performed by the drivers, as well as the opening and closing of the doors.	9
2	GoA1 with speed regulator	The drivers in this group drove with a speed regulator: the train driver chose a target speed and the automatic system takes over the task to accelerate or brake to reach and maintain target speed. The traction force was chosen by the driver. The driver performed the tasks of setting-up target speed, setting-up traction force, opening and closing the doors. In addition, the driver remained responsible for the train operation and monitored the speed regulator and was ready to take over control in case of deviations to the expected behaviour.	10
3	GoA2	The drivers in this group drove with the highest level of automation in the study: GoA2. In this level, the target speed was read directly from the ATO trackside and balises (simulated on the simulator). The train driver performed the tasks of authorising the ATO system to start and of opening and closing the doors. The traction force was calculated automatically by the ATO system. The train driver remained responsible for the train operation and monitored the GoA2 system, ready to take over control in the event of deviations from the expected behaviour. In this group, the drivers had the choice to drive with or without a commercial radio receiver. The radio programme was a free choice of music, news or other types of programme. The radio system was linked to the simulator and automatically attenuated in case of a system signal. All drivers driving with the radio on voluntarily agreed to do so.	12

Table 1 – The driver groups and their tasks.

Drivers of the groups 1 and 2 underwent six drives of approximately 25 minutes each:

- The first two drives were acclimatisation drives for all groups, no event was provoked and no measurements were made.
- During the third drive, a first event was provoked (defect of the catenary dropping to 0V) and the reaction time of the drivers was measured.
- A 20-minute break occurred.
- The fourth and fifth drives were uneventful.
- During the sixth drive, two events occurred in random order: a main signal became closed (after passing an open approach signal) and a defect of a crossing barrier was announced. Reaction time of the driver to brake, take over the system or acknowledge the defect was measured (whichever reaction occurred first).

Train drivers of the group 3 underwent a seventh drive in which the ATO system missed a stop due to wrong journey profile. Reaction time of the drivers was also measured.

During the break, multiple human factors were measured such as:

- Fatigue using the Karolinska Scale: self-evaluation by the participant with a measure between 1 (extremely alert) and 9 (extremely sleepy, fighting sleep).
- Workload using the NASA-Task Load Index (TLX) Scale. This scale uses self-evaluation by the participants and is based on six factors contributing to workload: mental demand, physical demand, temporal demand, overall performance, effort, and frustration level.

The simulation occurred in three shifts (morning shift from 7am to 11am, afternoon shift from 11:30am to 3:30pm and evening shift from 4pm to 8pm). The participants were randomly assigned to a driving group and to a shift, ensuring the shift would not interfere with the measurement of fatigue in the different groups.

The hypotheses to be verified

Using the measured factors, following hypotheses were verified:

- [H1]: The fatigue of the train drivers increases with time, in all automation levels.
- [H2]: The fatigue of the train drivers increases more with a higher automation grade (GoA1 with speed control and GoA2), compared to a lower automation grade (GoA1 without speed control).
- [H3]: The reduction of workload of the train drivers is higher between GoA1 without speed control and GoA1 with speed control, compared to GoA1 with speed control and GoA2.
- [H4]: The performance of the train drivers is similar under GoA1 with speed control and GoA2 driving modes.

Results

H1 The fatigue of the train drivers increases with time, in all automation levels.

Out of the 31 participants, the level of fatigue increased after the simulation for 28 participants. The participants for whom the fatigue decreased were in different automation groups, two of the three participants had been allocated to the morning shift and one to the evening shift. The results are shown in Table 2.

It can be observed that the fatigue of the drivers increases on average in all three groups by 1.23 points. Hypothesis 1 can be confirmed from this experience.

H2 The fatigue of the train drivers increases more with a higher automation grade (GoA1 with speed control and GoA2) than with a lower automation grade (GoA1 without speed control).

The average increase of fatigue in the first group shows the highest value (+1.43), whereas the increase of fatigue in the second and third group show similar values, respectively +1.10 and +1.08. This is shown in Figure 2.

It appears that driving under GoA1 with speed control and driving under GoA2 both lead to a lower increase of fatigue than driving under GoA1 without speed control. Therefore, the second hypothesis is rejected. This is a surprising result in this study, as considering the researched studies before designing our simulation, we had expected a greater increase of fatigue in higher automation groups. To deepen our understanding of this result, the workload of the drivers under each automation mode must be analysed (our third hypothesis).

H3 The reduction of workload of the drivers is higher between GoA1 without speed control and GoA1 with speed control, compared to between GoA1 with speed control and GoA2.

The workload of drivers was measured after three drives and after the overall simulation. The six dimensions of workload are self-evaluated by the drivers and results are combined to assess the overall workload, graded from zero (low workload) to 100 (high workload).

The dimensions of the NASA-TLX Index are defined as follows:

- Mental Demand: How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?
- Physical Demand: How much physical activity was required? Was the task easy or demanding, slack or strenuous?
- Temporal Demand: How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?
- Overall Performance: How successful were you in performing the task? How satisfied were you with your performance?

Figure 2 – Increase of fatigue before and after the simulation.

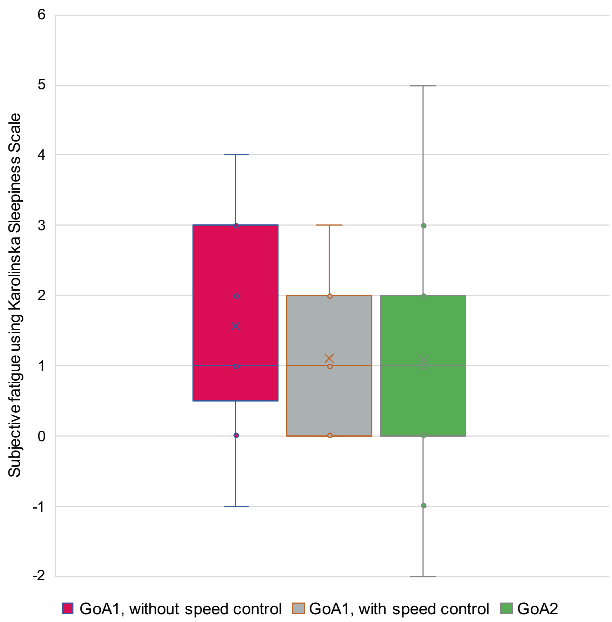

Table 2 – Fatigue of the drivers before, during and after the simulation.

Group	Average fatigue during the simulation	Average fatigue during break	Average fatigue after the simulation	Increase of fatigue between before and after simulation
1: GoA1 without speed control	2.67	3.56	4.22	+1.56
2: GoA1 with speed control	3.10	3.40	4.20	+1.10
3: GoA2 overall	3.58	4.08	4.67	+1.08
GoA2 with radio	4.28	4.28	4.71	+0.43
GoA2 without radio	2.60	3.80	4.60	+2

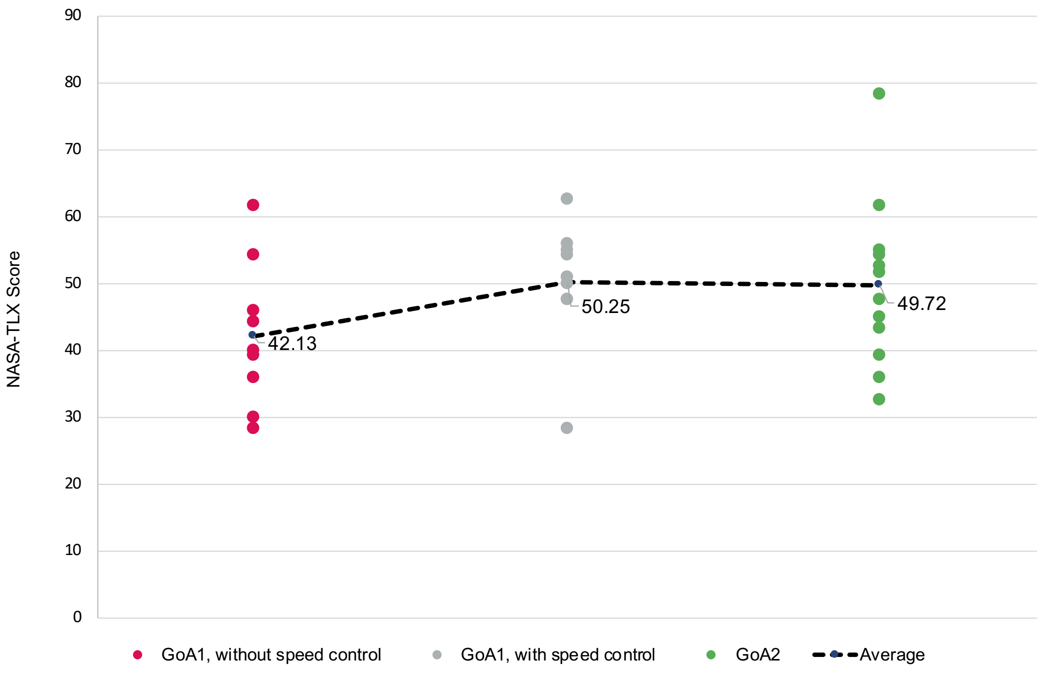

Figure 3 – Workload in all groups at the end of the simulation.

- Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?
- Frustration Level: How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?

The results in Figure 3 show that the workload of the drivers is lower in the first group (average 42.13 points) in comparison with the second group (average 50.25 points), despite a higher level of automation with use of the speed regulator in the second group. Between the second and third group (average 49.72 points), a similar level of workload can be observed. It seems that the increase of automation such as introduction of the speed regulator or GoA2 does not lead to a reduction of workload compared to GoA1 without speed regulator. To understand this result, the single dimensions of the NASA-TLX scores were analysed.

In comparing GoA1 without speed control to GoA1 with speed control, it can be observed that the introduction of speed control leads to:

- A significant increase in the dimension "Mental demand" (+4.96 points).
- An increase in the dimensions "Overall performance", "Temporal Demand" and "Effort" (respectively +2.16, +2.34 and +2.68 points).
- A decrease in the dimension "Physical Demand" (-1.92 points).

Although the overall index results are similar between GoA1 with speed control and GoA2, the detailed assessments of each index show significant differences. The introduction of GoA2 compared to GoA1 with speed control leads to:

- An increase of the dimensions "Effort" and "Frustration" (respectively +1.43 and +3.47 points).
- A decrease of the dimensions "Mental demand" (-2.40 points).
- A significant decrease of the dimension "Temporal Demand" (-3.15 points).

To assess our third hypothesis, the increase of workload occurring between the break time and the end of simulation were assessed, i.e., during the second part of the simulation. The results are shown in Table 3.

It can be observed that the workload provoked by driving is higher in the second and third group compared to the first group. Considering all three groups started their shift with the simulation, the influence of an earlier workload in the day can be disregarded. Hypothesis 3 is rejected.

Linking our results of the previous paragraph and this one, it appears that the group with the lower increase of fatigue is the group with the highest increase of workload, and vice-versa. The lower increase of fatigue in higher automation groups (GoA1 with speed control and GoA2) compared to GoA1 without speed control might be explained by the higher workload in these two groups, hence moving away from underload spectrum and passive fatigue.

Table 3 – Workload of the drivers at break time and at the end of simulation.

Group	Average workload during the break (after three drives)	Average workload after simulation	Raise of workload between break and end of simulation
1: GoA1 without speed control	39.35	42.13	+2.78
2: GoA1 with speed control	45.92	50.25	+4.33
3: GoA2 overall	44.79	49.72	+4.93
GoA2 with radio	42.14	48.10	+5.95
GoA2 without radio	48.5	52.00	+3.5

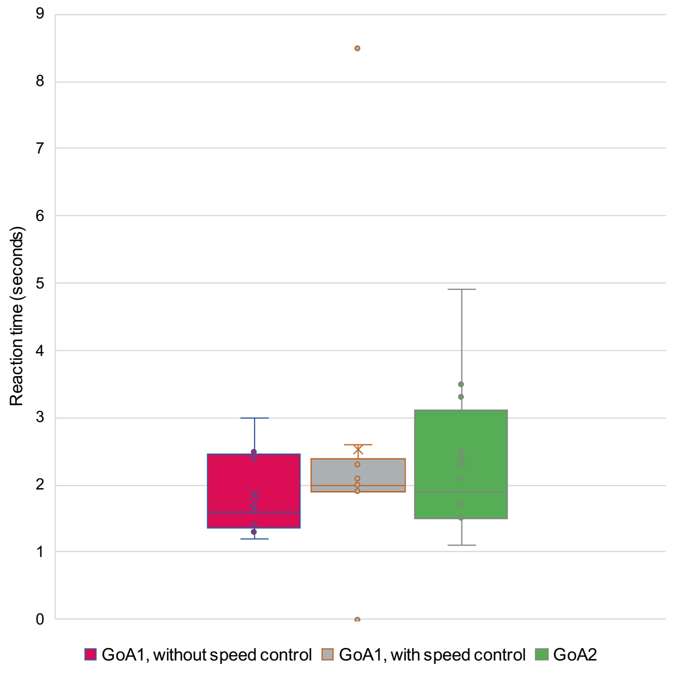

Figure 4 – Reaction times after second event.

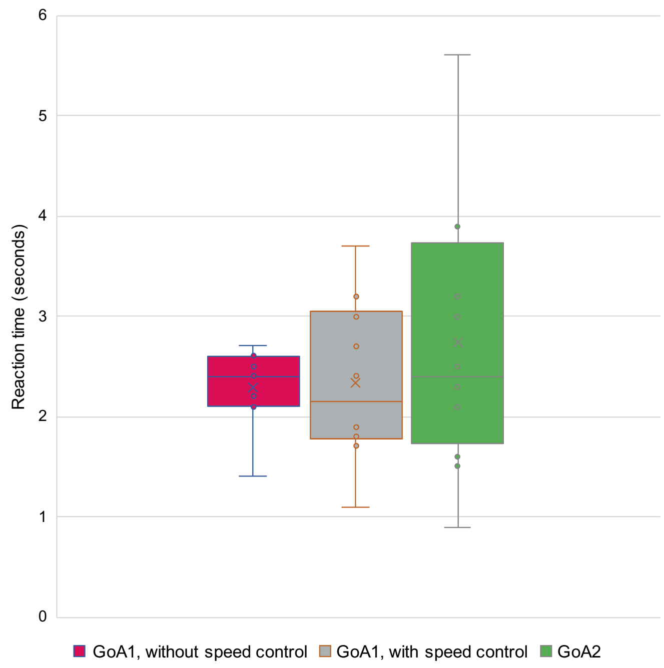

Figure 5 – Reaction times after third event.

H4 The performance of the train drivers is similar under GoA1 with speed control and GoA2 driving modes.

To assess this hypothesis, only the second (signal closure) and third events (crossing barrier defect) were used. This is due to results for the first event (catenary voltage dropping to 0V) being reported by locomotives are furthest away to the reality of train driving on a real track and results provide a wide range of possibilities with low statistical probability. Indeed, in real driving, in case of a catenary voltage dropping to 0V, the main switch of the train opens and the noise of the HVAC fans in the cab disappears. In the simulation, this change to the soundscape was not audible by the locomotive drivers, hence the event being missed or misinterpreted.

Reaction times of the drivers in all three automation groups for the second and third events are shown in Figures 4 and 5.

In the second event, drivers of the third group were slightly faster than drivers of the second group whereas in the third event, the contrary occurred. The differences of reaction times between the second and third group are -0.2 and +0.4 seconds respectively, which can be considered similar, hence a confirmation of Hypothesis 4.

The effect of having an activated commercial radio receiver can be observed: in both events, drivers of the third group who chose to listen to radio are faster than drivers who chose not to.

Applications and outlook

With this study, the influence of automation on the fatigue and performance of train drivers seems to differ from the results obtained in high-speed trains. Indeed, with hypothesis 2, we observed a lower increase of fatigue in the drivers who were performing in highly automated environments (GoA1 with speed control and GoA2) compared to drivers performing under GoA1 without speed control. According to hypothesis 3, this seems to be explained by the higher NASA-TLX workload when driving under GoA1 with speed control and GoA2, compared to GoA1 without speed control. Indeed, it seems the introduction of speed control has led to a higher workload for the train drivers, and the introduction of GoA2 will lead to an even higher workload as per the NASA-TLX index.

Therefore, the introduction of higher automation grades does not seem to negatively impact fatigue nor performance of the S-Bahn train drivers compared to the already widely introduced driving mode with speed regulation.

Yet, while the introduction of GoA2 significantly reduces time pressure for the train drivers, the component of the NASA-TLX index most raised by this new technology is the frustration. This finding must be carefully considered while introducing GoA2. In the frame of this study, an interview of S-Bahn train drivers was performed with 28 participants from Switzerland to assess the acceptance of introducing GoA2. While the initial answer without accompanying measures was rather low (rated on average with a 2.2 out of 10), the acceptance was raised significantly when the locomotive drivers were consulted to optimise the frequency of drives and widening the geographical area of driving under GoA2 (with an average of 5 out of 10). Additionally, the interviews showed that the introduction of a commercial radio receiver under GoA2 could potentially further raise the acceptance of train drivers. In our study, we could not show any negative effect of introducing radio under GoA2, but we observed a much lower increase of the frustration occurred while driving with radio.

We conclude that while the introduction of GoA2 compared to GoA1 with speed regulation for S-Bahn drivers does not lead to an increased fatigue nor decreased performance, it could lead to a dissatisfaction and disengagement of the train drivers. To support the train drivers in their acceptance of the new technology, adapting the frequencies of driving, widening the geographical area of driving and listening to radio. We are strongly convinced that for the industry to fully benefit from the introduction of GoA2, the personnel affected most, the train drivers, must be fully considered and supported by accompanying measures.

About the authors ...

Lucie Pannecoucke, senior consultant, AWK Group AG, Zürich, Switzerland

Lucie is an experienced project and people manager with more than 12 years of experience in the development of railway systems. She has managed complex industrial projects for delivering complete locomotives, and control and command system upgrades to various national operators and private customers. In this study, Lucie leads human factor studies to assess the effects of automation on the train drivers, collecting existing theoretical knowledge, writing and performing interviews, designing and evaluating the simulation and assessing overall results.

Roger Dällenbach, Schweizerische Südostbahn (SOB), Zürich, Switzerland

Roger has gained a broad range of experience in project management during his career. Over the past 14 years, he has managed several high-volume fleet procurement projects (newly designed passenger trains as well as complex maintenance vehicles). Before moving into the rail industry, he worked for 14 years in the field of aviation as metrology and avionics engineer and purchasing manager. This study is a part of the ATO pilot project currently in progress at SOB, in which Roger occupies the role of the senior project manager.

Contributor: Dr Céline Mühlethaler, ZHAW, Centre for Aviation, Winterthur, Switzerland

Céline is a senior lecturer and head of the human factors research team. She and her team have provided the methodology, instruments, and research advisory as well as consulting support.

What do you think?

Are you surprised by the findings of this research? Perhaps you have experience that entirely backs up what Lucie and her team have discovered? Maybe your railway or organisation has carried out its own research into the human impact of automation, or maybe you have experience in another field, perhaps aviation or road transport.

Why not share your experience with other IRSE members? Email us at **editor@irsenews.co.uk**.