Effects of automatic train protection on human factors and driving behaviour

Max Hely\(^a\), Todd Shardlow\(^b\), Bruce Butt\(^c\), Rena Friswell\(^d\), Andrew McIntosh\(^e\), Ann Williamson\(^d\)

\(^a\)Independent Transport Safety Regulator, NSW, AUSTRALIA; \(^b\)Chevron, Perth, WA, AUSTRALIA; \(^c\)Sydney Trains, NSW, AUSTRALIA; \(^d\)Transport & Road Safety (TARS), School of Aviation, UNSW, AUSTRALIA; \(^e\)McIntosh Consultancy & Research, Sydney, NSW, AUSTRALIA.

Introduction: The European Train Control System (ETCS) is recognised worldwide as the pre-eminent standard for automatic train protection (ATP) systems, particularly where interoperability between suppliers and countries is of concern. However, despite its widespread (and, in Europe, mandatory) adoption, no systematic or experimentally rigorous evaluations of its effects on driving performance or behaviour under realistic operational conditions have been reported. Unlike many European countries, where extant legacy ATP systems are being replaced by ETCS, the introduction of ATP in Australia (specifically New South Wales) is novel and represents a significant departure from current practices in which drivers rely entirely on lineside signals and speed signs for their driving authorisations. While ETCS is expected to provide significant safety benefits, it also introduces uncertainties regarding workload, attentional demands and driving behaviour that have the potential to offset the increase in signalling-related safety. To date, only anecdotal (and often contradictory) information has been available about the effects of ETCS on driving practice. Method: To objectively examine these concerns, a cohort of 32 experienced intercity train drivers undertook a 2 x 2 factorial repeated measures design in a high fidelity, full-cab train simulator. Drivers operated with and without ATP on replications of two routes for which they were qualified and which represented differing levels of infrastructure density. Driving operations incorporated a range of normal (non-degraded) scenarios including speed reductions and stopping at platforms and signals. Data were collected on primary performance variables (speed, acceleration, control operations, response times), subjective workload (NASA TLX) and eye tracking metrics. Additionally, a newly developed brake integration metric was used to quantify drivers’ braking strategies. Results: The addition of ATP enabled Drivers to continue to manage their trains effectively with only minor changes to their primary task performance. Effects on average speeds and section times were relatively small and unlikely to be operationally significant. The subjective evidence also suggested that drivers effectively managed the transition to driving under ATP supervision without compromising their perceptions of their own capability to safely and effectively manage their train. However, convergent evidence from several objective measures suggested that drivers managed this transition and maintained acceptable primary task performance at the cost of increases in workload associated with the attentional demands of ETCS. A redistribution of visual attention associated with the ETCS in-cab Driver-Machine Interface (DMI) resulted in a large reduction in time spent attending to the track ahead. This has important safety implications with respect to trackside hazards that are not controlled by ETCS. Quantitative evidence related to braking strategies was also provided for the first time to support the long-held proposition that train drivers engage in adaptive strategies in an attempt to minimise additional workloads at critical times. Discussion: The evidence suggests that while ETCS is a manageable addition to current practice, it significantly increases the perceptual, cognitive and attentional demands of driving when used as an overlay to an extant lineside signalling system. This evidence strongly implies that considerable caution be taken regarding the choice of features to be incorporated into the drivers’ interface as well as the system implementation features that affect the consistency and complexity of its associated procedures and operations. Some specific implications for implementation are discussed.

Practitioner summary: The introduction of ETCS as an overlay automatic train protection system onto an existing complex rail network should provide significant safety benefits. However, it also introduces associated cognitive demands which strongly implies that sound Human Factors design principles be incorporated into its configuration and implementation in order to minimise the potential for excessive workload, attentional demands and distraction to offset the signalling-related safety advantages.

Keywords: Rail transport, simulation, automatic train protection, attention, workload, eye-tracking.

1. Introduction

The European Train Control System (ETCS) was originally developed in the European context to be suitable for high speed, national and cross-border rail networks. However, little rigorous evidence is available in relation to its application as an overlay automatic train protection (ATP) system on complex existing networks similar to
those in New South Wales (NSW; Australia). Because of the differences between the European context and the NSW networks and operations, the proposed application of ETCS here brings with it numerous uncertainties associated with human factors issues and their potential safety, acceptability and performance consequences.

It is critical to the success of a train operator’s safety and business objectives that, when introducing new in-cab technology, the ability of Drivers to continue to optimally manage their trains is not compromised. The importance of sound design supporting the human element in complex systems is being increasingly recognised in contemporary human factors practice in all domains (Reason, 2008). Despite claims about the evidence base and usability of ETCS [Kecklund, L. et al (2001), ERRI (1996)], no systematic evaluation results have been forthcoming from either overseas rail authorities or ETCS suppliers. Anecdotes from suppliers and overseas sources suggest diverse, even contradictory, consequences for driving style, with some suggesting more conservative approaches to driving will occur, some suggesting less conservatism, and some claiming no change will occur.

By contrast, there is an abundance of information from other transport domains which identifies the potential for human factors and safety problems with automated and semi-automated systems and their in-cab interfaces. A recent serious incident on the ETCS-fitted Cambrian line in the United Kingdom has been attributed to, among other things, the Driver’s workload associated with attending to the in-cab ETCS screen at the same time as remaining vigilant to lineside indications (RAIB, 2011). While in-vehicle displays can provide drivers with additional information that may benefit performance, it has been shown that, if not well designed and implemented, the displays could also be a source of distraction and elevated error potential (Young et al, 2003). Problems have arisen when automated or semi-automated systems with in-cab interfaces confused drivers or distracted them from their primary driving task, or when drivers failed to either detect or understand changes in the state of the automated system (Norman, 1990; Young & Stanton, 2007).

Numerous studies have also clearly demonstrated that excessive demands of in-vehicle information displays are leading contributors to errors which have resulted in serious incidents (Lee et al, 2009; Stevens & Minton, 2001). There is also evidence (Lansdown et al, 2004) that additional workload causes drivers to generate compensatory responses in vehicle control, particularly braking, which are directly related to the level of workload and which can result in detrimental effects on performance.

Although the above problems have arisen within other transport domains, the dearth of research specifically examining their potential in rail presents a considerable risk for the rollout of any advanced rail system that features similar in-vehicle characteristics and demands. The nature and extent of the driving constraints imposed by ATP may also directly or indirectly result in changes to driving strategies, since this computational “supervision” appears not to be entirely compatible with current driving skills and practices, train performance characteristics or Drivers’ extant route knowledge. Since considerable resources are invested in Driver training and skill maintenance, it is essential to also understand how these may be affected by the introduction of ATP.

The paucity of evidence-based information about the specific effects of ETCS on driving behaviour, workload and train control under operational conditions, together with the specific, possibly unique, requirements of the Sydney/NSW Trains network, highlights the importance of conducting empirical evaluations to better understand these uncertainties. The evidence from other transport modalities where these issues have been more comprehensively examined underscores the necessity of ensuring sound human factors principles are incorporated in the design, configuration and implementation of such systems in order to avoid adverse consequences. The objective of this research is to evaluate the effects of ATP under normal (i.e., non-degraded) driving conditions and to consider their implications for overall system performance, safety and system acceptability.

2. Methodology

2.1 Test environment

A full-cab, fixed-base train simulator replicated the in-cab displays and controls of an Intercity V-set train cab (with and without an ATP DMI screen), trackside features and realistic signalling aspect sequences. The ATP system was built to the ERA ETCS Baseline 3 System and DMI specifications. The simulator data sampling was event-based unless event separations exceeded 0.1 seconds (a minimum sample rate of 10 Hz). The track vision was based on actual track data and video for the Blue Mountains (BM) and North Shore (NS) routes.

2.2 Study design

The study used a 2 x 2 (supervision condition x route) repeated measures factorial design. Each Driver completed four test combinations comprising two supervision conditions (Non-ATP and ATP) and two routes (Blue Mountains and North Shore). The Blue Mountains (BM) route represented an outer area with low signalling and infrastructure density, and the North Shore (NS) route represented an inner area with high signalling and infrastructure density. Drivers first drove the two routes under NATP conditions (representing
current driving conditions with which they are already familiar), followed by the ATP conditions. The order of presentation of route conditions was counterbalanced. The test run portions of the two routes were selected to be approximately equal in timetabled duration. The differences in route infrastructure density are set out in Table 1.

Table 1. Descriptions of the two simulator test routes.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Blue Mountains</th>
<th>North Shore</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total distance</td>
<td>17.965 km</td>
<td>10.325 km</td>
</tr>
<tr>
<td>Timetabled duration (approx.)</td>
<td>22 minutes</td>
<td>17 minutes</td>
</tr>
<tr>
<td>No. of stations</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>No. of signals</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>No. of speed signs</td>
<td>13</td>
<td>6</td>
</tr>
</tbody>
</table>

2.3 Subjects

Thirty-two NSW train Drivers (30 male; 2 female) with normal or corrected-to-normal vision participated in the study. All were qualified to drive V-set trains and had current knowledge of both routes. The proportions of males and females in the test sample (94% male; 6% female) was consistent with the actual ratio in RailCorp’s driving population at the time (95% male; 5% female).

The Test Drivers were aged from 25 to 68 years, ranged in train driving experience from 3 – 33 years, and had no previous experience with ATP. Screening using a fatigue test battery (Williamson et al, 2001) confirmed that the test Drivers did not differ significantly in levels of fatigue between NATP and ATP sessions.

2.4 Scenarios

The test scenarios were selected to provide “snapshots” of the effects of ATP on some of the major non-degraded driving demands. These are set out in Table 2 below.

Table 2. Test scenarios in Blue Mountains (BM) and North Shore (NS).

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
<th>BM (approx. dist.)</th>
<th>NS (approx. dist.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All stops</td>
<td>Overall run starting and finishing at specified stations within route.</td>
<td>13,130 m</td>
<td>8,150 m</td>
</tr>
<tr>
<td>2. Clear signals</td>
<td>Traversing a section of track under full clear signals.</td>
<td>1,546 m</td>
<td>895 m</td>
</tr>
<tr>
<td>3. Speed sign</td>
<td>Approaching and passing a speed reduction sign.</td>
<td>600 m</td>
<td>600 m</td>
</tr>
<tr>
<td>4. Platform stop</td>
<td>Approaching and stopping at a platform.</td>
<td>1,110 m</td>
<td>1,040 m</td>
</tr>
<tr>
<td>5. Stop at signal</td>
<td>Approaching and stopping at a Stop Signal.</td>
<td>1,075 m</td>
<td>640 m</td>
</tr>
<tr>
<td>6. Clearing signal</td>
<td>Approaching a Stop Signal that clears on the approach.</td>
<td>645 m</td>
<td>345 m</td>
</tr>
<tr>
<td>7. RIFOD</td>
<td>RIFOD (&quot;Return In Face Of Driver&quot;) where a proceed signal suddenly and unexpectedly returns to Stop as train approaches.</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Note: The RIFOD was included to examine any effect of ATP on Drivers’ response times to a sudden, unexpected, trackside safety-critical event.

2.5 Procedures

The test runs were preceded by Driver briefings and a simulator familiarisation phase. It was emphasised to all Drivers that it was the ATP system, and not their driving ability, that was being evaluated and they were encouraged to drive realistically. Eye tracking data collection was conducted separately after the main operational phase to ensure that the use of eye tracking equipment by Drivers not familiar with it did not confound any ATP effects during normal driving. In the eye tracking phase, all operational scenarios were replicated and were also preceded by familiarisation runs.

Before proceeding to data collection, all Drivers were assessed by several criteria, including expert assessment, objective data and the Drivers’ own judgement, for their capability to operate the simulator in a manner consistent with normal driving practice and to understand and use ATP while driving. All Drivers were evaluated as adequately capable with the simulator and the ATP system to qualify for inclusion in data collection.
and proceeded to the operational phase. One Driver later withdrew from the eye tracking phase due to discomfort with the eye tracking equipment.

The Drivers were required to drive each route as an all-stops, time-tabled run in accordance with current network rules and procedures and using their own preferred driving style. As will be the case with actual ATP implementation when used as an overlay to the existing signalling system, Drivers were instructed to treat the lineside signals, along with current network rules, as their authority.

2.6 Human Factors effects measures

2.6.1 Driving strategy

Speed and duration
It has been proposed that, under ATP, the speed Drivers travel through sections may change, either as a direct result of imposed supervision or, indirectly, as a strategic response adopted by the Driver to the presence of ATP. Changes to average speed or duration through sections has implications for network level performance indices such as on-time running.

Braking strategy
Changes to current driving strategies may also manifest as changes in braking strategy (e.g. earlier or later braking) on approach to speed reductions (e.g., signs, platforms, signals). Subtle disruptions in skill-based control has been shown to indicate potentially more profound usage of the Drivers’ limited resources (Holding, 1989; Lansdown et al, 2004). Previous research (McIntosh & Jauregui, 2010) used train braking data over discrete journey sections to develop a new driving metric, the brake integral value, that showed potential to reveal changes in individualised braking strategies.

This measure is calculated by mathematically integrating brake setting values over the normalised time of a braking segment to quantify an individual’s braking profile in terms of both the magnitude and timing of braking effort applied across the duration of a braking segment. The brake integral value represents the proportion (%) of total braking completed at any point in time over the braking segment. Unless otherwise specified in the text, brake integral values are for the halfway (50%) point, by time, of the segment. No such approach to objectively quantifying the effects of ATP on critical aspects of train driving behaviour or strategies has been previously undertaken and it will, therefore, be instructive to obtain evidence for what, until now, has only been unsupported (and often contradictory) speculation.

2.6.2 Workload

While objectively successful task completion may suggest a manageable workload, the associated perceptual, cognitive or physical demands may contribute to fatigue, task overload, or frustration. Subjective workload was therefore evaluated using the multi-dimensional NASA Task Load Index (TLX; Hart & Staveland, 1988; Nygren, 1991). These scales were completed by the Drivers on completion of each of their four test runs.

2.6.3 Attentional demands

Visual attention
The direction of one’s gaze (as indicated by eye movements) in demanding operational environments reliably indicates a person’s attentional focus (Konstantopoulos et al, 2010; Luke et al, 2006) and vision is the predominant source of safety- and performance-critical information for a driver. Importantly, many hazards which can be seen “out the window” are not controlled by the signalling system nor, therefore, by ATP. Failure to detect these hazards due to visual attention being distracted from the primary task of monitoring the track ahead represents a significant risk.

The changes in distribution of Drivers’ visual attention with ATP can be revealed by quantifying the proportion of attention directed towards out-of-cab compared to in-cab visual sources (i.e., “head up” versus “head down” driving). To our knowledge, this is the first time eye tracking has been used to evaluate the effects of ETCS on visual attention and behaviour.

Response time
Attentional re-distribution to in-cab information can result in an increase in response times to unexpected trackside events. Slower response times to safety-critical stimuli such as a RIFOD (or, equivalently, an on-track obstruction, track workers, or similar) represent a significant increase in risk.
3. Results

3.1 Driving strategies

There were no major differences overall between NATP and ATP on either the BM or the NS routes. All Drivers completed the ATP runs at approximately the same speeds (and thus durations) as for current practice. Few Drivers overspeeded on the NS route due to the relatively high station density. However, on the BM runs, the majority of Drivers overspeeded, most on multiple occasions. Both the maximum speed over the posted speed limit and the time spent over the limit differed significantly between the ATP and NATP conditions. On both routes, when overspeeds occurred, ATP clearly exerted a constraint on both their magnitude and duration.

There was a significant interaction between supervision condition and route for train speed when passing a speed reduction sign. Without ATP, the BM featured overspeeding at the sign and the NS featured underspeeding. The effect of ATP was to produce an increase in speed on the NS and a decrease on the BM. In short, the effect of ATP was to produce a convergence toward the posted speeds.

3.2 Braking metrics

3.2.1 Speed sign

On both BM and NS, the 50% brake integral values were significantly lower for ATP than NATP, i.e., significantly less of the total braking was performed early in the approach with ATP than with NATP (see Figure 1). It appears, therefore, that regardless of route, ATP produced a shift toward delaying the use of the brakes during the approach to the speed sign.

![Speed sign: 50% Brake integral](chart1)

Figure 1. Brake integral for speed sign approach.

3.2.2 Platform stop

On the BM, there was a negligible difference in the 50% braking integral between NATP and ATP (see Figure 2). However, on the NS, there was a shift toward earlier braking (i.e., an increase in the integral value) with ATP compared to NATP.

At 75% of the braking time, the BM brake integral indicates a significant shift toward later braking under ATP conditions, i.e., less braking occurring in the 3rd quartile of the braking segment, whereas the NS continued to show a shift toward earlier braking with ATP compared to NATP.

![Platform stop: 50% Brake integral](chart2)  
![Platform stop: 75% Brake integral](chart3)

Figure 2. 50% and 75% brake integral for a platform stop. Note that the BM value has decreased at 75% of the segment time, whereas the NS integral continues to rise.
Since there is no logical relationship between this effect and any direct ATP demand (as ATP does not supervise station stops), this inconsistency, though small in magnitude, may indicate a disruption to the Drivers’ established driving patterns from driving with ATP, though this effect appears to be influenced by the route.

3.2.3 Signal at stop

The brake integral values were significantly higher for ATP than for NATP, regardless of route, i.e., ATP resulted in a shift toward earlier braking on the approach to a stop signal.

![Stop at signal: 50% Brake integral](image)

Figure 3. Brake integral on approach to signal at stop.

3.2.4 Approach a clearing signal

There was a significant difference in the brake integral values between the BM and the NS under NATP (see Figure 4) with the NS showing predominantly early braking and the BM predominantly later braking. There was no significant difference between the routes under ATP, however, changes in the brake integral between NATP and ATP were in different directions for the two routes. On the BM, the ATP brake integral was higher than the NATP value, while on the NS, the ATP value was lower than the NATP value. They tended to converge with braking shifts in opposite directions.

The tendency toward convergence to a non-significant difference under ATP suggests that ATP may be entraining braking behaviour in those situations where, without it, there is greater opportunity for behavioural variation. We can speculate that this may be more likely with a clearing signal where, under NATP, the driving constraint disappears when the signal clears. By contrast, under ATP, the constraint (i.e., release speed) persists to a certain level, even though the signal has cleared. As was noted previously for speed when passing speed signs, these findings may partly reflect differences in the layout and demands associated with the different routes.

![Clearing signal: 50% Brake integral](image)

Figure 4. Brake integral on approach to clearing signal.

3.3 Workload

On both routes, there was a clear, albeit non-significant, trend towards an effect of ATP on Drivers’ perceived workload. This trend did not reach statistical significance due to the variability in Drivers’ ratings.
Table 3. NASA TLX results.

<table>
<thead>
<tr>
<th>NASA TLX Scales</th>
<th>NATP Average</th>
<th>ATP Average</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mental</td>
<td>52</td>
<td>58</td>
<td>+6</td>
</tr>
<tr>
<td>2. Physical</td>
<td>37</td>
<td>41</td>
<td>+4</td>
</tr>
<tr>
<td>3. Time pressure</td>
<td>28</td>
<td>33</td>
<td>+5</td>
</tr>
<tr>
<td>4. Effort</td>
<td>44</td>
<td>50</td>
<td>+6</td>
</tr>
<tr>
<td>5. Dissatisfaction</td>
<td>36</td>
<td>34</td>
<td>-2</td>
</tr>
<tr>
<td>6. Frustration</td>
<td>29</td>
<td>37</td>
<td>+8</td>
</tr>
<tr>
<td>Total (Av.total)</td>
<td>226 (38)</td>
<td>252 (42)</td>
<td>+26 (+4)</td>
</tr>
</tbody>
</table>

3.4 Response time

There was a small but statistically significant slowing of response time by 0.6 seconds under ATP compared to NATP. However, the distributions of the response time data are also of interest. Figure 5 shows that the response times for NATP, with the exception of a very small number of outliers, are tightly clustered about the average of approximately 1.4 seconds.

By contrast, the ATP distribution is considerably more variable (i.e., “spread out”) towards longer response times. The coefficient of variation of these two samples shows that the variability in response times increases from approximately 0.29 (29%) under NATP to approximately 0.49 (49%) with ATP. This represents an almost 70% increase in the variability of the response times.

![Response time distributions](image)

3.5 Visual attention

The proportion of time that Drivers spend attending to the out-of-cab scene (i.e., “head up” driving) under NATP and ATP reduced from over 90% under NATP to approximately 70% with ATP (see Figure 6).

![Visual attention distribution](image)

---

1 The coefficient of variation (CV), also known as “relative variability”, is the ratio of a dataset’s standard deviation divided by its mean. When expressed as a percentage, it provides an intuitive understanding of how variable, or “scattered”, a data set is. A lower percentage indicates a lower variability in the data set; a higher percentage indicates the data set is more varied.
4. Discussion

The evidence from these analyses indicates that the addition of ATP under normal, non-degraded conditions enabled a cohort of relatively experienced Drivers to continue to perform with only minor changes to their primary task performance. Changes to speeds (and thus times) through sections, where they occurred at all, were relatively small and unlikely to have any significant operational consequences.

Similarly, the test Drivers appeared, both from their individual workload ratings and the objective train performance evaluations to have managed the transition to ATP driving quite effectively. This suggests that the fully ERA-compliant Baseline 3 ETCS is a manageable addition to current driving practice. However, it is also clear that primary task performance was maintained by the Drivers at the cost of a significant increase in the attentional demands of ATP.

It is well known that skilled operators in any domain, including drivers in transport, are able to accommodate increasing task demands with no observable reduction in primary task performance, but at the cost of increasing allocation of their limited information processing resources to attain or maintain that level of performance (Hamilton & Clarke, 2005; Wickens, 1989). This study’s analyses of visual attention, response times and subjective evaluations provide convergent evidence to suggest that ATP introduces significant additional attentional demands to the driving task sufficient to show early symptoms of potential performance and safety decrement.

In addition to the above, this study has, for the first time, provided quantitative evidence to support the proposition (anecdotally long held by Drivers themselves) that Drivers engage in adaptive strategies in an attempt to minimise additional workloads. The scenario-dependent changes in braking profiles suggest a redistribution of attentional demands associated with the perceptual-motor task of braking, such that braking was made earlier or later across the braking segment, perhaps to allow for more pressing attentional demands elsewhere in the segment.

This is consistent with minimising the potential for ATP distraction at times when driving skills require maximum attention to be directed elsewhere (e.g., attending to critical tasks such as slowing, stopping, etc.) and provides evidence that the Drivers are making adaptive changes to accommodate the intrusion of ATP into their primary driving task.

The eye tracking data have also, for the first time with ATP (ETCS), provided further evidence of a very large re-allocation of attentional resources to the in-cab DMI, reflecting both a considerable workload increase and a large reduction in the proportion of time spent attending to the track ahead.

The latter effect is an independent risk that presents a significantly increased potential for missing trackside hazards that are not monitored, indicated or controlled by the signalling system itself (e.g., cautionary signs, defects, workers, obstructions, etc.).

In conclusion, while this research provides no evidence against the introduction ATP as an overlay to an existing network, it does suggest considerable caution in the manner in which it is implemented. In particular, its introduction should be subject to appropriate attention being paid to Human Factors design principles which aim to minimise additional workload, distractions and error potential, and maximise safety, performance and system acceptability.

Acknowledgements
This study was a major undertaking and would not have been possible without the invaluable contributions of Tony Hanrahan (Simulation & System Support Manager, TfNSW), Peter Hill (Simcon Services), Sydac Pty Ltd, Peter Brawn (Eyetracker Pty Ltd), Simon Morris (Driver Specialist, NSW Trains), Mick Ablott (ATP Training Manager, TfNSW), Kiri-Maree Taggart (HF Specialsit, TfNSW), Richard White (Operations Analyst, TfNSW), Ken Stephenson (ATP Operations Manager), the RTBU representatives (Andy Holt, Keith McMahon & John Aquilina), and the 32 RailCorp (now Sydney/NSW) train drivers.

References


Stevens, A. & Minton, R. (2001) In-vehicle distraction and fatal accidents in England and Wales. Accident Analysis & Prevention, 33, 539-545.


