To Beep or Not to Beep: Developing a Non Fail-Safe Warning System in a Fail-Safe Train Protection Environment

Richard van der Weide, Kirsten Schreibers and Colete Weeda

INTERGO human factors & ergonomics
Pausdam 2
3512 HN Utrecht, The Netherlands

ABSTRACT

In The Netherlands the most deployed train protection system (ATB-EG) leaves a 'gap' below 40 km/h, leading to about 200 signals passed at danger per year. A new warning system (ORBIT) was to fill this gap to a large extent by warning the train driver when approaching a red signal at (too) high speed. ORBIT is not fail-safe, because it is based on GSM-R/UMTS communication and GPS. The design concept and different design solutions of this imperfect automated system were tested with 17 train drivers in a train simulator, and during four weeks in actual service on one route. Results indicate that ORBIT is effective in identifying hazardous situations. Train drivers accepted the warning system. They preferred a warning that is infrequent, diminishing uncertainty, but leaving just enough time to react adequately. A spoken warning was preferred to a tone warning. Results also indicated that minor adjustments in the algorithm were needed to avoid too many unnecessary warnings.

Keywords: Warning system, Train protection, Human System Integration, Alarm management, Imperfect automation

INTRODUCTION

Train protection systems in The Netherlands

The most deployed automatic train protection system (ATB-EG) in The Netherlands leaves a 'gap' below 40 km/h, leading to about 200 signals passed at danger (SPAD) per year. Although the Dutch railway system is relatively safe, some recent incidents and accidents reinforced the need for extra measures to reduce the number as well as the effect of SPADs. A significant example was a train to train collision in Amsterdam in 2012 involving 1 fatality and 180 people injured. One of the extra measures is a new warning system (ORBIT) to fill the 'ATB-EG gap' to a large extent by warning the driver when approaching a red signal at (too) high speed. In short, ORBIT continuously checks train position, speed, and brake characteristics against the position of the next red signal, and generates a warning when a critical speed/distance curve ("alarm curve") is passed. A schematic visualization of the system can
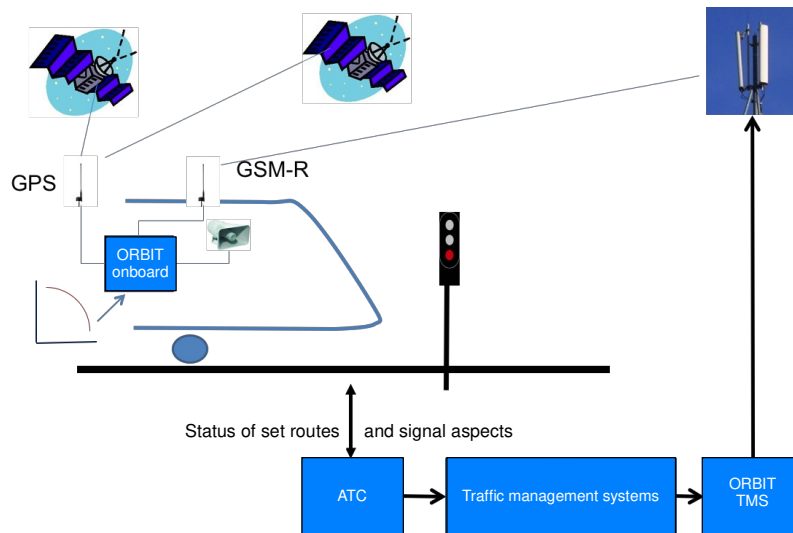
be seen in figure 1.

Figure 1. Schematic visualization of ORBIT.

After passing a yellow signal aspect train speed must be reduced to 40 km/h and the driver must be prepared for a red signal aspect (stop) at the next signal. ATB-EG checks whether the required speed of maximum 40 km/h has been reached or if the brake has been activated. If one of these conditions has not been met the train will automatically be stopped by ATB-EG. However, the system allows passing a signal at danger with a train speed of less than 40 km/h or even with a speed over 40 km/h when a minimal brake has been activated. More advanced train protection systems, with full brake curve supervision like ERTMS/ETCS, will be implemented in the next decades.

ORBIT in addition to train protection systems

To overcome the period until nationwide implementation of ERTMS a relatively simple warning system has been developed: ORBIT. ORBIT is based on knowledge from train traffic management systems about the end of set routes (red signals) combined with GPS coordinates of this signal, and actual position, speed and acceleration of the train based on GPS. This information is evaluated against the brake characteristics of the specific rolling stock. If a red signal is approached with a too high speed ORBIT presents a warning to the driver.

ORBIT is not fail-safe as it based on GSM-R/UMTS communication and GPS, both of which can fail or be unavailable. As a consequence of the operational concept design ORBIT can only guard signals controlled by the train traffic management system, and not automatic signals. In practice this means that about 6000 out of total 11700 signals are covered, and it must be stressed that controlled signals are located at most hazardous infrastructural situations, like points or switches. So ORBIT is planned to be in effect at the most risky locations. In fact, it was analyzed that in theory ORBIT would have been helpful in 84% of the SPADs during 2004-2008, as well as in two major accidents in 2009 and 2012. From a cost/feasibility perspective ORBIT is advantageous compared to systems that need to be integrated into rail infrastructure. Basically, ORBIT uses information already present in traffic management systems, combined with - new to most of the Dutch fleet but commercially available - GPS information and a relatively simple algorithm. The ORBIT onboard unit has no controls, other than a secured power switch that may only be operated with permission. Sounds run via dedicated speakers and entry of train number happens automatically via existing GSM-R (this functionality was not yet implemented, and was done by a remote test facilitator during the test).

Human factors issues

The human factors specialist joined the team of (IT) systems designers to cover relevant aspects. In terms of system performance ORBIT must enhance the driver's perception, assessment and decision to brake in time when

approaching a red signal. In figure 2 (Wickens and McCarley, 2008) this is visualized: ORBIT adds the Imperfect Automated Diagnosis path to the situation where at speeds lower than 40 km/h performance only relies on human perception, assessment and decision. More specific, ORBIT should help prevent a Miss by Human Decision. From this signal detection theory the following issues can be identified:

- Hits by ORBIT must be conveyed to the driver in a perceivable and understandable way (alert salience) to result in the correct action by the driver;

- False Alarms may be detrimental to driver's trust in ORBIT, and thus potentially result in ignoring the warning. The rate and nature of false alarms are discussed below in terms of reliability;

- Misses by ORBIT influence efficacy of ORBIT, but may not have a large effect on trust because SPADs (Misses by human decision) are quite rare.
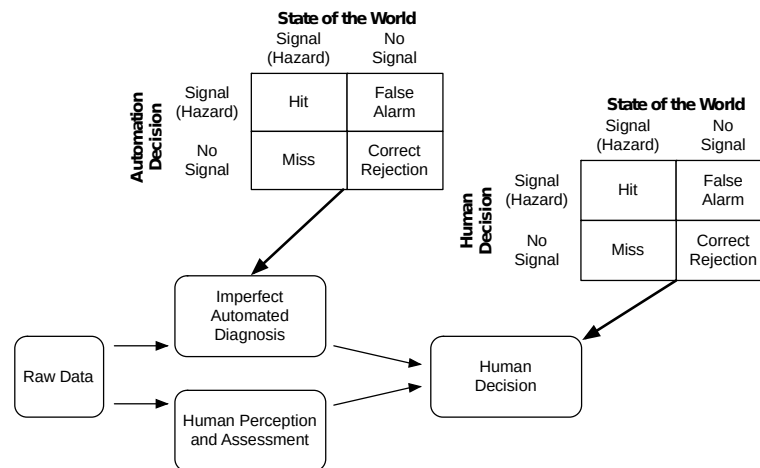

Figure 2. Signal detection theory applied to parallel human and automation alerting system (Wickens and McCarley, 2008)

For an optimal alert salience in a train cab auditory warnings seem to be preferred as the primary task is highly visual; visual attention must not be diverged or overloaded (Woodson, 1981; Sanders and McCormick, 1993; Peryer et al., 2010; ISO 7731, 2003). The choice for an auditory warning for ORBIT was confirmed in preliminary tests with drivers, although some of them would prefer a combination of auditory and visual warnings. For practical reasons this combination was not further tested. Obviously, auditory warnings should be distinctive from existing warnings, discernable above environmental sounds, and fit with the 'message' (Lees and Lee, 2006).

Introducing a non fail-safe system in a fail-safe train protection environment raised questions about required reliability and the amount of warnings a train driver should encounter. Required reliability of ORBIT in terms of true and false alarms and misses could not easily be determined from existing literature. First, terminology is quite diffuse: alarm and alert are not well defined. EEMUA for example (2007) states:

- alarm: audible or visible means of indicating to the operator about an equipment or process malfunction or abnormal condition;

- alert: lower priority notification than an alarm, that has no serious consequence if ignored or missed.

However, in many studies (e.g. Wickens et al., 2009) alarm and alert are used as synonyms which makes it difficult to select required reliability specification just on the priority of the warning. Second, according to EEMUA definitions ORBIT would qualify as an alarm by elimination: the consequence of ignoring or missing the warning

can be very serious. Still, ORBIT would be of lower priority than warnings from the ATB train protection system. It has been suggested that an automated attention guidance system (which ORBIT in essence is), will generally assist human performance so long as the reliability is above about 0.80 (Wickens and McCarley, 2008). A "naturalistic" study of Wickens et al. (2009) however showed that air traffic controllers accepted a 47% rate of false alerts of an automated conflict alerts system. The authors concluded that this relatively high rate was accepted because false alerts were considered "forgivable" as a result of a conservative detection threshold, resulting in more false alerts and less misses. Also, the alert did not distract controller's attention from their primary task as they are both addressing the same high priority task, and the alert was not using sound which would be more intrusive and possibly annoying when false.

During ORBIT development three types of false alerts were identified: 1) inadvertent/overdue alerts, when in fact the red signal has changed to yellow or green, but that change has not reached the train because of long propagation times in TMS or GSM-R/UMTS systems, 2) unnecessary alerts, when the train driver is in full control of the braking movement but an alert is produced anyway, 3) random alerts, when a warning is sounding while no red signal is being approached, possibly because of incorrect GPS-positioning of train or signal. The first two types may be acceptable to drivers, as long as they understand what triggered ORBIT in the actual situation. However, too many of these alerts may cause nuisance. Random alerts are not comprehensible and may cause shock reactions, and undermine trust in the system (Lees and Lee, 2007). Eventually, high rates of false alerts may cause ignoring ORBIT warnings at all. Safety is not at stake directly, because ORBIT does not impede perception of outside signals, but a 'cry wolfe' effect (slow or no reaction to warnings) may be expected (Wickens and McCarley, 2008). Efficacy of ORBIT will become low.

Very high reliability of ORBIT may cause overreliance issues: the train driver may be tempted to wait for the ORBIT warning before starting to brake. Anecdotally, some drivers rely on the ATB-system, and with more recent train protection systems like ETCS, drivers are taught to follow indicated speed-brake profiles. This behavior is unwanted and hazardous with ORBIT, because ORBIT is not fail-safe and only in effect with controlled signals. In this perspective, a certain rate and type of false alerts may not be detrimental to calibrate trust (Parasuraman et al., 2000). Also, this calls for an accurate introduction of ORBIT to drivers to make them understand the nature of choices made in development of this imperfect diagnostic automation (Wickens and MCarley, 2008).

Design of ORBIT

In accordance with the concept of likelihood alarms (Wickens and MCarley, 2008) it was decided to incorporate uncertainty of the warning in the nature of the warning. This was done by introducing a pre-alarm warning ("attention" warning) meant to focus attention of the driver to the next supposedly red lineside signal. The attention warning is triggered a fixed number of seconds before the alarm curve would be crossed; in fact an attention curve is defined in addition to the alarm curve (see figure 3). This will cause the warning to be triggered more often than the average of SPADs which is about once per 10-15 year per driver, and allowing the driver some extra time to act appropriately. The algorithm for both curves takes into account whether the driver is already braking, by delaying the warning with the time it takes for the brakes to couple.
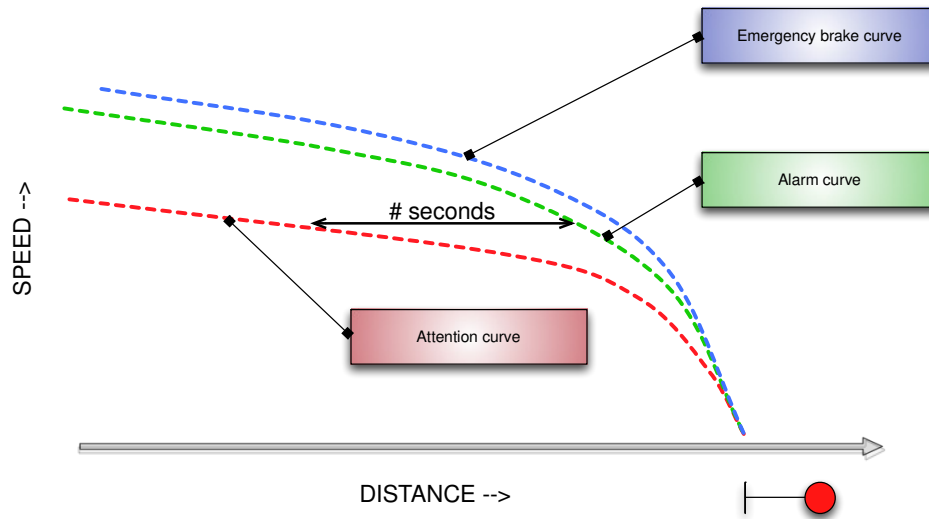
Figure 3. Schematic display of ORBIT-warning triggering curves (x-axis: distance to red signal; y-axis: train speed)

In addition to these curves an algorithm based on positive acceleration and speed within a distance of 20 m of the red signal is used to warn at unauthorized departure.

In this research we focus on alert salience by determining specifications for the (auditory) ORBIT warning, and on the balance between reliability and number of (true and false) alerts taking into account risks of compliance and reliance. The human factors part of all research done in the development of ORBIT consisted of the following parts, following an initiating HAZID session with stakeholders and a literature review:

1. Simulation I: explorative of nature to gain driver acceptance and fine-tuning of parameters;

2. Simulation II: validating driver acceptance and system efficacy;

3. Field test: validating system efficacy and driver acceptance during 20 days in regular service on The Hague – Venlo route.

SIMULATION I

Method

Three experienced passenger train drivers from one company-were involved in five half-day sessions using the ProRail MATRICS simulator extended with an ORBIT simulation module. This simulator consists of two laptop computers, one of which controlled by the test leader, the other connected with a RailDriver™ console for actuation of brake, traction, doors, and deadman placed on a driver's console. The simulated outside view was projected with a beamer (see figure 4). One route (Breukelen-Geldermalsen) of about 45 minutes was used for evaluation of several ORBIT settings. The test route considered a stop train service in a narrow time table, in order to meet as many red signals as possible, including crossing a complex yard (often prone to SPADs). Main goal of simulation I was to narrow down the number of experimental conditions for Simulation II, and to improve the scenarios. The following conditions were tested:

- Auditory warning: tone (attention: single tone 700 Hz, 0.2 s, 1/1 s; alarm: three-tone: 700/930/700 Hz, 0.7 s, 1/0.85 s) or a woman's voice (attention: "Attention, SPAD", alarm: "Brake now!: SPAD!", translated from Dutch). Instruction was given to check the signal aspect when an attention warning was

given, and then act/brake if required, and to brake immediately when an alarm warning was given and check the signal aspect subsequently. Volume of all ORBIT warnings was about 10 dB(A) higher than simulated ambient noise of 60 dB(A).

- Distance between alarm curve and attention curve: adjustable between 3 to 9 seconds (see figure 3).

Drivers were allowed to be present during other driver's tests to get maximum experience with all test conditions. Train drivers are used to presence of colleagues in the cab for e.g. educational purposes. In order to distract drivers, as often happens during SPADs, a secondary visual and cognitive task was added. Within the primary field of view every 15 seconds a new sum was to be solved (+/- 7 or 13 to/from a two digit number between 30 and 70). Evaluation was done by questionnaires and debriefing semi-structured interviews.
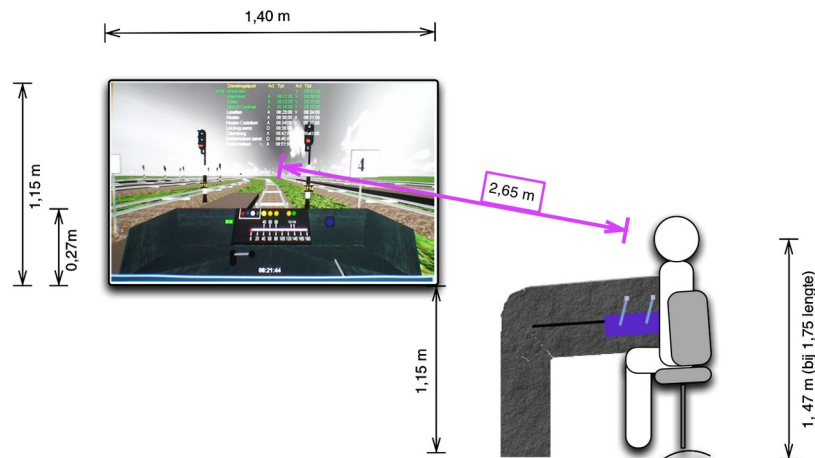

Figure 4. Simulator setup

Results

All three drivers assessed the simulator as being sufficient real to life for this purpose. They were all positive about the concept of ORBIT (8-9 on 10 point scale).

The drivers had a slight preference for the spoken warnings compared to the tone warnings, mainly because the tone had to compete with about ten other tones in the train cab, and the voice directly gives meaning to origin and required action. The fact that a spoken warning may also be audible and comprehensible to passengers in some cabs was identified as a disadvantage.

Contrary to the instruction it was observed that all drivers actuated the brake after an attention warning instead of checking the signal aspect first. They declared that this was their natural reaction to auditory warnings in general. Furthermore, the attention warning was too long in their experience: they suggested that "Attention: SPAD" could easily be replaced by "SPAD", as the auditory warning is an attention in itself.

Because of the urgent character of the warnings, the drivers prefer the ORBIT warnings to be sparse, leaving just enough time to react appropriately. In the simulation this meant that the shorter conditions between alarm and attention curve (3-5 sec) were preferred to longer distances (6-9 sec). It was concluded that the shorter distances were suited for further testing and validation. As preference for voice warnings did encounter some practical dilemmas it was decided to use both in the validating simulation.

Some adjustments were done in the simulator due to technical instability of the software, like a shortened test track. The scenario was adapted to better fit actual driving along this track.

Human Aspects of Transportation III (2022)

SIMULATION II

Method

Due to practical constraints – a limited amount of train drivers with limited time available - a proper counterbalanced design was not feasible. Fourteen experienced passenger train drivers (average 52 years old, 21 years of train driving experience, all male) from four different companies were involved in half-day tests in groups of three (one group of two). All drivers were allowed a 15 minute test drive on a part of the experimental route to get accustomed to the MATRICS simulator. After that each train driver drove Breukelen – Houten Castellum (30 minutes). They could witness the colleague drivers from their test group drive during their practice and test runs to get maximal experience during the half-day. Drivers were not allowed to talk to each other until final debriefing interview. Each driver drove the same experimental scenario once, but with different settings. These settings were:

- distance between alarm and attention curve 3 seconds + spoken warnings

- distance between alarm and attention curve 4 seconds + spoken warnings

- distance between alarm and attention curve 5 seconds + spoken warnings

For practical reasons (time constraints) tone was coupled with the longer distance between alarm and attention curves, because the longer distance potentially triggers more ORBIT warnings and voice was argued to be too intrusive for frequent alerting. These settings were not communicated to the drivers in advance. The spoken and tone warnings were demonstrated during briefing before the test runs. The last signal on the route was set as a technical fault causing the signal to suddenly show the red aspect during approach. An ORBIT warning was thus inevitable. This was to make sure that every driver was confronted with an ORBIT warning at least once. The secondary task from simulation I was used to distract vision and load cognition.

ORBIT loggings were analyzed for number and nature of ORBIT warnings, and reaction times (time between start of ORBIT warning and brake application) with a sample frequency is 1/300 ms. Unfortunately, it appeared technically impossible to simulate system propagation delays. So in fact, simulation tests were done with an 'ideal' ORBIT system in terms of GPS accuracy, almost zero propagation delay, etc.

Results

Due to the small number of test persons and the fact that ORBIT only triggers a warning in exceptional cases it is hard to make quantitative statistical analyses. In table 1 the number of ORBIT warning per experimental condition is shown. Figure 5 shows the distribution of reaction times to ORBIT warnings – defined as the time difference between warning occurrence and brake actuation. As loggings were taken at 300 ms intervals the distribution is shown in categories of 300 ms: median 1.2 s and mode 0.9 s. Adding an average 1.5 s of brake coupling time (for passenger trains with electromagnetic brakes) this would mean that average three and maximum four seconds would be present between ORBIT warning and start of train deceleration.

Table 1: Mean (min-max) number of attention and alarm warnings per experimental condition (to obtain 5 observations per condition one driver performed 2 test drives; the inevitable ORBIT warning at the last - red - signal was not counted)

3 seconds (n=5)		4 seconds (n=5)		5 seconds (n=5)	
attention	alarm	attention	alarm	attention	alarm
1 (0-3)	0 (0-1)	4 (3-6)	0 (0-1)	3 (0-7)	0 (0-0)

Scores for the secondary task did not differ. From table 1 it might be concluded that the three seconds condition causes less warnings. However, from observations and from literature we argue that individual driving styles are of major influence. As all drivers were able to witness the other drivers in their group (with other conditions) responses to questionnaires seem to be more valuable. From these questionnaires seven drivers preferred the three seconds condition, four drivers the four seconds condition, and one driver the five seconds condition. Two drivers did not notice difference. Written clarifications were quite consistent with simulation I: ORBIT should be not disturbing during 'controlled' braking and an ultimate warning leaving just enough time to act. Supporters of the four seconds condition thought the three seconds condition too 'tight', and preferred a second of extra time to act.
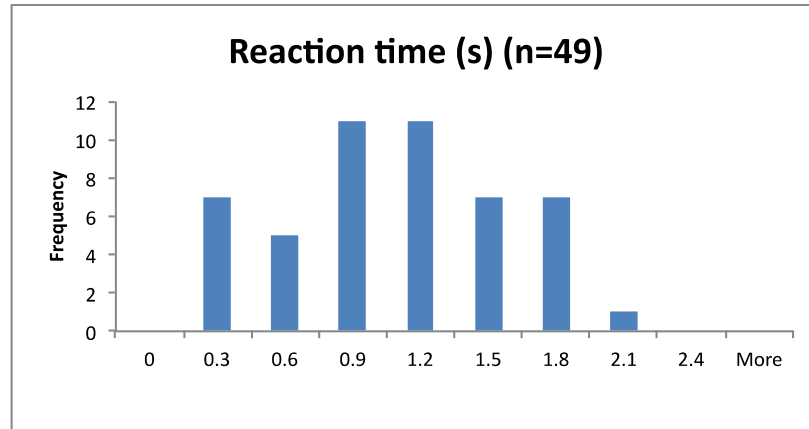

Figure 5. Frequency distribution of reaction times to ORBIT warnings

All 14 drivers preferred the voice warning above the tone warning, and 13 were in favor of nationwide implementation of ORBIT. The one driver who voted against this thought ORBIT superfluous: a well-trained driver should not need this assistance/warning device, and the non fail-safe nature of the device was not appreciated. Average overall appreciation of ORBIT was 8.2 on a 10-point scale (differing from 7 (3x), 8 (6x), 9 (4x) to 10 (1x)).

The vast majority thought volume, distinctiveness, and appropriateness of both attention and alarm (spoken and tone) were good. So design of the warnings cannot explain preference for a timing condition. It was concluded that the three or four seconds distance between alarm and attention curve in combination with the voice warning was best to be implemented in the field test. Taking into account results from reaction time (see above) the four seconds distance was chosen for the field test.

FIELD TEST

Method

Ten intercity trains (type ICR) were equipped with ORBIT onboard units in one cabin (BDS). During 20 days in regular service on The Hague – Venlo route ORBIT loggings were analyzed. Loggings contained information about ORBIT unit number, train number, next red signal guarded by ORBIT, distance to that signal, speed, acceleration, type of ORBIT warning (attention/alarm), and type of ORBIT algorithm that triggered the warning (braking/non-braking/departure). In real-time ORBIT was activated by a remote test facilitator when an instructed driver was present and the BDS cabin was on the front side of the train. Drivers from major stations along the route were instructed about presence and working principles of ORBIT. The instruction was done by information flyers, information on their handheld device, and personal instruction by their manager (who were instructed by our team).

Every ORBIT warning was followed-up by a structured subjective evaluation interview per telephone. Initially,

this was to be performed by the driver's manager. However, it soon turned out that this was too much effort for these managers and a member of the human factors team took over this task.

Human factors specialists analyzed all objective loggings information and interview results, and classified ORBIT warnings to inadvertent/ overdue, unncessary or random.

Results

During the test period 218 valid train journeys on (parts of) this route were recorded, and 5800 red signal approaches guarded by ORBIT. Analysis of loggings showed 78 ORBIT warnings were triggered (see Figure 6). About 70% were only attention warnings and 30% attention followed by alarm warnings. Further analysis of loggings showed that 29 out of 79 warnings appeared (technical) incorrect. Eight of those (10%) seemed to be associated with drift in GPS positioning due to dense and high buildings along the track or a large station roof (mainly Rotterdam). In 21 other incorrect warnings (27%) the warning was inadvertent/overdue: 12 took place due to an error in Traffic Management Systems (not safety related) and the rest was caused by excessive signal propagation times (often more than 60 s, where 2 s is a nominal value) most probably caused by GPS fix or UMTS connection problems in and after leaving a tunnel. Although hard to establish, no automation misses could be identified.

Out of 49 technical correct warnings 42 (54% of total) were evaluated by both train drivers and human factors experts as unnecessary: the speed profile showed that the train was in control. Often, when nearing a standstill, drivers release the brake which may cause a temporary positive acceleration. If this happens within 20 m of the red signal the 'departure algorithm' causes an ORBIT warning. A big part of these ORBIT warnings occurred at a red signal with apparently incorrect GPS coordinates. As the train stop position is at the signal the ORBIT onboard unit thought the train was stopping 5m closer to the signal than it actually was, thus generating a warning.
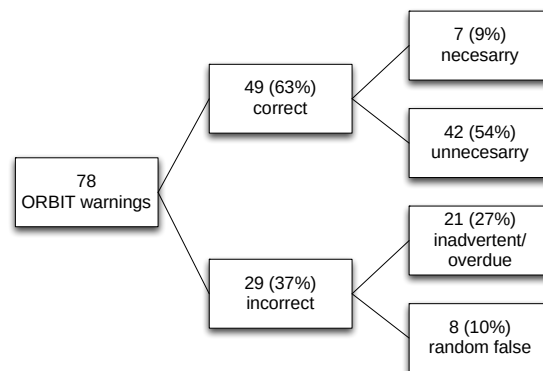

Figure 6. Classification of ORBIT warnings during the field test.

Seven ORBIT warnings (9% of total) finally were considered necessary by the human factors experts, based on visual comparison of the speed profiles to that of other trains that did not generate a warning. The braking was at least uncomfortable, although some drivers declared that they had braked in a 'normal' way and the ORBIT warning was in their eyes unnecessary. In one case, where an attention warning was followed by an alarm warning, the train driver admitted ORBIT was useful urging him to brake in time.

The subjective evaluations with 37 drivers – equaling 57 ORBIT warnings – resulted in 28 completed interview forms. A few drivers experienced ORBIT more than once. From nine drivers only summarized interview results were delivered by their manager, which were not useful for further analysis. Contact with the train drivers regarding the remaining 21 ORBIT warnings, failed.

Some train drivers could not recall an ORBIT warning in the trains. Most probably the speakers of two ORBIT units have been defect during (part of) the field test. In other cases about half of the drivers thought the sound
Human Aspects of Transportation III (2022)

volume was too low. A sample verification measurement in two trains confirmed this feedback: the attention warning had about the same volume as the ambient driving noise at 40 km/h (~55-60 dB(A)), which was 10-15 db(A) less than specified. These results also indicated the need for an integrated test of speaker functionality in the train's onboard unit.

Nevertheless incorrect or unnecessary warnings or low volumes, 80% of the interviewees support nationwide implementation of ORBIT. Average appreciation figure is seven on a ten point scale, ranging from four to nine. Three drivers scored lower than six, two of them arguing that a well-trained driver should not need a warning device, and one driver was really shocked by a technical incorrect warning. All train drivers judged incorrect warnings as not acceptable. The unnecessary signals were however acceptable to most drivers as they understood that ORBIT could generate a warning in that specific situation. Timeliness and duration of the warnings were considered adequate on average

DISCUSSION

The concept of ORBIT appeared to be a valuable measure to reduce the number of SPADs by capturing and guiding the train driver's attention, and improving his situation awareness. Also, as expectations about set routes may drive an incorrect situation awareness, ORBIT may help to correct the driver's mental picture of the situation. This is important because expectations are a major contributing factor to SPADs in The Netherlands (ILT, 2013).

However, during the field test some technical issues arose that generated incorrect ORBIT warnings. These warnings – where drivers cannot relate the warning to a potentially hazardous situation – are not acceptable. After the field test technical measures were taken to overcome errors originating from the TMS. Also, by including a time stamp in the original signal, warnings that are overdue should be avoided. This would introduce a low percentage of automation misses. The same holds for locations with unreliable GPS and GSM-R/UMTS signals: by leaving signals at those locations outside ORBIT coverage and implementing other safety systems at those locations, risk of SPADs can be mitigated.

The relatively high percentage of unnecessary warnings compared to necessary can be overcome by changing some parameters for determining whether a train is braking or accelerating. Over 85% of the unnecessary warnings would then disappear, while no of the necessary warning would have been missed. Still, in accordance with other sources (e.g. Wickens and McCarley, 2008) it was confirmed that "unnecessary" warnings can be acceptable to a certain extent: unnecessary warnings can be interpreted as an understandable or forgivable form of false alerts. They confirm presence of the warning system and can serve to reinforce the driver's assessment of the outside environment.

Contrary to the ATC Conflict Alert system (Wickens et al., 2009) ORBIT uses auditory warnings. Wickens et al. suggest that sound is intrusive and thus potentially annoying when false. Our results suggest that sound is in favor of and preferred by train drivers because the auditory channel is relatively free compared to the primary visual channel. A visual warning in the train driver's console would redirect the driver's outside visual focus, which we consider undesirable. An alternative would be to project the visual warning as a heads up display. This solution was not considered from a cost/feasibility perspective. It might have been an adequate alternative although a potential 'competition' between the visual warning and the lineside signal must be taken into account. Furthermore it must be noted that the auditory pre-alarm attention warning in ORBIT is distinctive but 'gentle'.

The preference of drivers for a warning that is infrequent (not many warnings which are relatively reliable, but leave little time to react appropriately) in combination with a preference for a spoken warning is in accordance with theory. It certainly prevents issues that were identified with extended AWS in the UK where same warning for different states, ambiguous warnings, and a variety of causes were applied (McLeod et al., 2005). The spoken warning directly reveals to the driver which system triggered the warning, what the identified hazard is ("SPAD"), and how to act (alarm: "SPAD! Brake now!).

It was calculated that – when the proposed measures on technology and algorithms are taken and effective – a Human Aspects of Transportation III (2022)

train driver would be confronted with an ORBIT warning once every two weeks on average. For 'conservative' drivers this frequency may well be (much) lower.

Still two risks of ignoring or overreliance on the ORBIT warning were identified. By implementing the mentioned measures the number of incorrect and unnecessary signals are estimated to drop to 0% and 10% respectively. This is expected to mitigate the risk of ignoring the warning. Obviously, proof of efficacy of the measures has still to be delivered. Further risks of overreliance will be mitigated by a thorough introduction of ORBIT to train drivers, emphasizing the not fail-safe nature of ORBIT. Besides it is believed that unnecessary warnings – in the expected low rate – contribute to the understanding of the not fail-safe nature. Also, as automatic signals (about 50% of all signals) are not guarded by ORBIT, it would require a deliberate decision by the train driver to allow oneself to rely on ORBIT in guarded areas. This seems not very likely to happen. However, the risk cannot be excluded. By monitoring the number of ORBIT warnings and the average distance (in seconds) to the attention and alarm curves it is expected to be able to identify drifts in driving behavior at an early stage in future.

Reactions of train drivers to the attention warning were not compliant to instruction: they braked immediately instead of checking the signal aspect first. Although drivers indicated that the nature of the attention warning ("SPAD", spoken in a neutral tone of voice) fitted the original instruction, they naturally actuated the brake handle. In respect of their preference for a relatively late warning, it was decided to change the instruction to allow for natural braking behavior: brake and check signal aspect. For calculation of efficacy of the ORBIT system this is advantageous, because estimated driver reaction times could be lowered. Theoretically, with these measures and recalculations into effect 57% of all SPADs in The Netherlands can be prevented and in 63% the danger point will not be reached. These figures need to be evaluated in practice, but a reduction of several tens of percent is realistic.

In the next phase the effectiveness of ORBIT with freight trains will be studied. These trains are less predictable in their braking capabilities and behavior, because of varying composition, load, and brake settings. The feasibility of a universal attention and alarm curve for freight trains will be studied. Also, special procedures like passing a signal at danger with permission of the train signaler – which will cause ORBIT to trigger warnings – will be subject of study.

REFERENCES

EEMUA (2007), *191 Alarm Systems - A Guide to Design, Management and Procurement*. London: **T**he Engineering Equipment and Materials Users' Association.

ILT (2013), *"SPAD analysis 2012" In Dutch: "STS-passages 2012"*. Inspectie Leefomgeving en Transport. Ministerie van Infrastructuur en Milieu. Den Haag

ISO 7731 (2003), *"Ergonomics. Danger signals for public and work areas. Auditory danger signals"*. Geneva: International Organization for Standardization.

Lees M.N. and Lee J.D. (2007), *"The influence of distraction and driving context on driver response to imperfect collision warning systems"*. Ergonomics, Volume 50, No. 8, pp. 1264-1286.

McLeod R.W., Walker G.H., Mills A. (2005), *"Assessing the human factors risks in extending the use of AWS", i*n: Wilson J., Norris B., Clarke T., Mills A. (Eds.), Rail human factors: supporting the integrated railway. Aldershot, Ashgate.

Peryer G., Noyes J.M., Pleydell-Pearce C.W., Lieven N.J.A. (2010), *"Auditory alerts: an intelligent alert presentation system for civil aircraft"*. The International Journal of Aviation Psychology, Volume 20, No. 2, pp.183-196.

Sanders M.S. and McCormick E.J. (1993), *"Human Factors In Engineering and Design"*. New York: McGraw-Hill.

Schreibers, K.B.J., Van der Weide, R., Weeda, C.E. (2014), *"Human factors specialists to the rescue for (other) systems designers - Case study in railway"*. In: O. Broberg, et al. (Eds.) Proceedings of the XI conference on Human Factors in Organizational Design and Management. Copenhagen (*in press).*

Wickens C.D. and McCarley J.S. (2008), *"Applied attention theory"*. Boca Raton: CRC Press.

Wickens, C.D., Rice, S., Keller, D., Hutchins, S., Hughes, J, Clayton, K. (2009), *"False alerts in air traffic control conflict alerting system: is there a "cry wolf" effect?"* Human Factors, Volume 51, No. 4, pp.446-462.

Woodson W.E. (1981), *"Human factors design handbook"*. New York, McGraw Hill.