**The ERTMS railway signalling system; deals on wheels?**
**An inquiry into the safety architecture of high speed train safety**

John Stoop (1, 2) & Sidney Dekker (2)
(1) Delft University of Technology, the Netherlands
J.A.A.M.Stoop@tudelft.nl
(2) Lund University, School of Aviation, Sweden
Sidney.Dekker@tfhs.lu.se

**ABSTRACT**
ERTMS is the acronym for European Rail Traffic Management System. ERTMS is the future standard for a European train signalling system, enabling interoperable use of the European network without adapting rolling stock to different national signalling systems. It facilitates crossing national borders with a design speed of 300 km/hour eliminating time-consuming change of locomotives at borders. Simultaneously, ERTMS aims at a further automation of railway signalling systems by cabin signalling instead of track signalling.
ERTMS causes the loss of visual detection as a safety separation principle. It applies internal instead of external information and automatic trajectory clearance by automatic train detection and integrity control and a moving block protection, depending on train characteristics. Because automatic acceleration and braking curves are foreseen, the driver only has a monitoring task and is no longer in control. Ultimately, train control automation aims at driver free operations by Automatic Train Operations. Such a fundamental change in safety assurance concepts requires a most reliable engineering design, implementation and operational strategy. In this contribution, the potential of resilience engineering in designing an innovative alternative is explored.

## 1 INTRODUCTION

ERTMS is a part of the renovation and upgrading of national railway systems, facilitating interoperability on the EU rail network, a 25 kV power supply for heavy cargo trains and 300 km/hr high speed trains, a dedicated, ballast-free track, new rolling stock and fully software controlled train surveillance. ERTMS is a trend shift from technical compatibility across nations towards standardisation and harmonisation on the main EU network corridors. The Dutch HSL is part of Paris-Brussels-Koln-Amsterdam-London corridor.
For the Dutch ERTMS development several political choices have been made:

- innovation in Public-Private Partnerships in contracting; mixing public and private interests. On arguments of cost reduction and capacity increase during operations, a full separation between infra provider, service provider and operator is accomplished.
- With respect to technology; the development and implementation phase are conducted concurrent instead of sequential. A simultaneous development of standards and software components is taking place, assuming an Off the Shelf availability of components from various industrial consortia. ERTMS is considered a pragmatic merging of autonomous train control systems (ETCS) and communication systems (GSM-R): ERTMS= ETCS + GSM-R.

The Dutch High Speed Line is the first of international High Speed train corridors to deploy ERTMS and is deployed on the corridor Amsterdam-Antwerp in order to reducing travel times to 120 minutes between Amsterdam and Paris. The level migration from ATP level 1 to full automation level 3 is foreseen in three steps.

| Function allocation | Signalling | Train detection |
|---|---|---|
| Level 1 | track | track |
| Level 2 | train | track |
| Level 3 | train | train |

There have been several disruptions in implementation resulting in software upgrades. ERTMS version 2.2.2 proved to be cross-supplier incompatible but was contractually based deployed, while version 2.3.0 would be the new operational standard. There was no anticipation on necessary, continuous hot upgrades in practice. The level migrations (from level 1 to level 2) were implemented in the Netherlands without fallback options of proven technology such as with the French TVM 430 for the TGV. As a result of time delays and cost increases, the necessity for upgrades and migration and expansion of the testing period was repeatedly discussed in Parliament. These discussions lead to an inquiry into the ERTMS deployment strategy.
In addition, ERTMS is considered the future national standard for the Dutch railway systems, including light-rail developments, merging metro-tram and heavy rail rolling stock on a single interregional network. This harmonisation of train control systems also ultimately serves political goals: a doubling of capacity for half the

costs, while safety performance levels maintain unchanged. To this purpose, several feasibility studies were commissioned searching for new train operating concepts, combining infrastructure, rolling stock, signalling and control, while increasing capacity, reducing costs, lifting technological innovation beyond the present levels of incremental reconfiguration.

## 2 HISTORICAL DEVELOPMENT IN RAILWAY CONTROL

In developing railway signalling 3 historical phases can be discriminated:
1. At the dawn of the railway industry in the early 19th century, the system was modelled after the Napoleontic military organisation. Such an organisational structure was not only necessary due to the military strategic role of the railways in times of war, but served the purpose of compensating the deficiencies in safety technology at the time. A strict compliance with time tables and a scrupulous operation of signals and switches was the dominant safety concept to separate trains in time. Safety technology on the railways developed gradually over time, based on investigations into exploding steam boilers, derailment of trains, deficient pneumatic braking systems, signalling failures and railway crossings with bridges and roads. Until then, railway operations remained very labour intensive. Before technical fallback options were introduced, safety was dependent on self-disciplining and an almost flawless human compliance with regulations. A strict command and control structure by disciplining a large number of railway employees provided the necessary safety on the railways.
Within the Dutch railway company a refined system of financial fines, sanctions and dissatisfactions existed, dealing with trespassers of regulations and operational standards. Even in the early 1950's 5200 administrative verdicts per year were issued, which were eventually reduced to a more acceptable number of about 350 per year after the introduction of the principle of hearing the employees before fining them.

2. Enhanced automation. With the aftermath of a major railway disaster at Harmelen in 1962, the Dutch railway systems entered a second phase of automatic train control by introducing the ATP (Automatic Train Protection) system. The gradual implementation of this ATP system has covered a period of about 40 years since.
This system enabled the railways to establish a more accurately position finding by indication occupancy of a static block by a train on a display at the train control centre. Conflict resolution was enhanced by bringing the trains to an automatic and failsafe standstill if the separation of trains was violated by passing a signal at danger. Reliable communication with train drivers was facilitated by the introduction of several portable and mobile telephone systems. Because the organisational concept of hierarchical decision making remained, a strict separation was maintained between train control and train capacity management. Such a separation was necessary to avoid conflicting interest in decision making within the organisation by negotiating safety versus capacity. Conflict anticipation was resolved by planning and capacity allocation as a responsibility of the capacity management organisation, while conflict resolution was the responsibility of the train control centre. By the introduction of this automation, responsibilities of the train driver did not change, aiming at full and strict compliance with regulations, signalling and driving instructions from the traffic control centre in order to reduce human error by relying on constraining driver performance within strict limits.

3. ERTMS. ERTMS can be considered as the next phase in controlling train driver behaviour by reducing his role to monitoring the system, anticipating intervention in case of disruptions and deviations, evolving into a final driverless train system.

## 3 CONCEPTUAL LIMITATIONS

Full automated control systems however have their conceptual limitations. For reasons of economy of scale and cost reduction, many local train control centres are closed down, replaced by a few large regional or even national centres. A remote situation awareness of such a centralised system under pressure of the need for more detailed and actual traffic information and communication increases the workload of the controllers in case of traffic flow disruptions and incident handling. Additionally, performance based punctuality demands and financial incentives in maximising track capacity initiates conflicts of interest among the business interests of various privatised stakeholders (WRR 2008). This creates a trade-off between short term economical aspects against competing long term public values, creating a 'multiple principal agent problem' (Steenhuisen and Van Eeten 2008). This puts train controllers in a coping situation in which the specialised functionality of their organisation makes their tasks manageable and clearly demarcates their responsibilities, eliminating competing values from their scope. Meanwhile, however, frontline operators face these conflicting values as they emerge in practice. Such conflicts will be solved by their professional expertise and experience, but also may erode standards and norms, adapting operational practice to new demands and values. Eventually, such operational practices may erode the resilience of the organisation, introduces new implicit operating performance standards and causes a drift into failure.

Automation finally, has its limitations by design. With the increase in intensity, the system is loaded to its design limits. The fault tolerance in hierarchical systems increases quadratically with intensity. About its saturation point, the traffic flow becomes instable. At fault, operator induced oscillation becomes possible; fault handling may cause abrupt and progressive collapse of the overall system. To avoid initiating disturbances, an even stricter task performance of the train driver is required. Increasing the punctuality of the time table under high traffic intensity conditions demands an increasing control effort by the traffic controller and train drivers, aggravating the tactical and operational cognitive workload of the traffic control centres. Since the train driver has to drive strictly on clearance based performance and has no mental picture of the surrounding traffic, he cannot and is not allowed to contribute to conflict resolution. This task is solely allocated to the traffic controller, who is forced to communicate simultaneously with several train drivers in picturing a mental model of the actual system state in order to provide the necessary adaptations to the clearances for their task performance. In case of multiple faults and secondary faults eventually a gridlock situation occurs due to which all traffic operations must be terminated and restarted. This is done by a failsafe system breakdown and gradual and safety critical restart of the traffic flow according to the original timetables. Such system instability jeopardizes the capacity demands and public confidence in railways as a high quality pubic transport system.

The underlying organizational mechanisms which threat public values such as safety versus private business values such as capacity and economy can be identified as coping behaviour which has evolved in order to deal with conflicting values (Steenhuisen & Van Eeten 2008). A simultaneous removing of safety margins and introducing conflicting goals during operations is a process which may has unforeseen consequences for the operators in such increasingly dynamic and coupled systems and organisations. Does this however mean that they are unpredictable because there were not designed into the technical system and reveal themselves over time in practice as emergent properties in hybrid systems? If so, will it suffice to discipline organisations with advanced contracts and incentives, piling up requirements without clarifying inevitable trade-offs?

Or can we design resilience into the system to cope with the change in nature, and if so, how should we do that?

Two principal strategies are applied:

- recognition of value conflicts and subsequently, a structuring of the process of communication, coordination and cooperation among all stakeholders in their decision making processes, coping between quantifiable private performance indicators and qualitative public values;
- elimination of the human involvement in disguised bad performance due to ambiguous and hybrid decision making values by developing an innovative train control system, based on modern technology and a new generation of signalling systems.

Both approaches should incorporate additional demands accommodating a doubling in capacity for half of the costs, while maintaining safety at the achieved performance levels.

## 4 STRUCTURING DECISION MAKING PROCESSES

During the High Speed Line project development in the Netherlands, several unforeseen project cost increases and planning delays emerged in deploying the train control software. These disruptions caused questions in Parliament to the Minister of Transport, requesting clarification into the reasons for the software upgrade, necessary migration time and the reasonability and fairness of the testing period. During the inquiry, several value judgements became visible dealing with the project organisation and technological scope.

The main conclusions of the investigations into the ERTMS software upgrade were:

- the institutional environment has complicated the development and implementation of the project. The divisions that have been created during the project between design and construct of the hardware components and the contractual arrangements between stakeholders created a necessity for a complex interface management. This interfacing has not been accomplished
- the two main lines in contracting out the project have only indicated the necessity to create oversight by the end of the project. This division has not led to a role for an architect or systems integrator, responsible for the integral coherence during the implementation of the overall system. The pivotal role of ERTMS as a prerequisite for combining operations of track, rolling stock, train control and signalling system became emergent at the end of the project in the full scale testing phase of the integral system
- the technological development of ERTMS was underestimated. There has been a continuous tension between incremental progress and implementation in an existing railway network on one hand and the ambitions of innovative ERTMS and public-private partnership arrangements on the other hand.

In particular the consequences of several technological design decisions could have been foreseen if such decisions were submitted to a pro-active safety assessment procedure, comparable to other domains such as aviation, digital network providers or telecommunications. Several Points of No Return in the design process have been passed without oversight of their consequences:

- a choice for a new signalling system which was not yet operational at the time, was not compensated for by a qualified fallback option such as the TVM 430 signalling system for the French TGV

- the choice for an innovative ERTMS system in the Netherlands resulting in a system leap in signalling technology was not in harmony with the more incremental process and evolutionary development of the Belgian signalling system on the same High Speed Corridor Amsterdam-Antwerp
- the choice for connecting the Dutch and Belgian system manufactured by two different signalling system consortia at the country border forced the project management to develop a gateway at the border causing high costs and considerable delays in delivering the integrated system for testing and operations. The choice for one system between Rotterdam and Antwerp across the country borders could have been part of the Dutch-Belgian treaty on developing the corridor, in particular because the Dutch paid a substantial part of the Belgian track
- a contractually based testing and deployment of ERTMS version 2.2.2 took place while version 2.3.0 would become the new standard, causing unnecessary complications, costs and delays
- the migration and deployment of ERTMS towards increasingly higher levels transferred the costs and risks of railway signalling from the infra provider towards operator without clarifying the necessity for such migration and its consequences
- the development of ERTMS was considered a conventional technical engineering effort, enabled by a decomposition of the system components in autonomous position finding and communication subsystems. Development and manufacturing of these components was subcontracted across competitive consortia, each dealing with national industrial interests and corridor specific requirements. Each consortium was assumed to be able to deliver these components 'off the shelf' as proven technology. No attention was paid to the software architecture on the traffic management level, while system integration in the testing phase was under pressure of earlier delays and strict time planning
- no precautionary measures were taken to assure a smooth and efficient frequent upgrade of the signalling software during its operational phase, whereas it remained unclear how this upgrade will be settled during regular operations beyond the level of a procedural approach.

## 5 TOWARDS FULL AUTOMATION

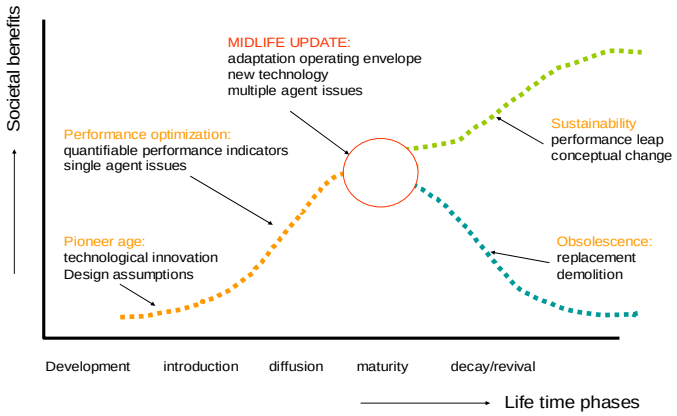
The discussion on full automation originates from the beginning of the process industry. In these days, safety experts in the rapidly developing process industry wondered whether the sector was suitable for a safety approach which was also applied in more conventional sectors.
The process industry applies a design concept in which humans are fallible factors and eliminated by design from the system by automation. Their remaining role is restricted to complying with rules which have been imposed by management. There is no room for the operator in taking critical decisions. Learning in practice is replaced by modelling and by a centralised assessment of all interests by a single party; the corporate management. This design doctrine has become the role model for modern safety management.
Transport systems however, apply two completely different design principles. First, at the control level the system is designed as a support for the operator; it is a human centred design with delegated responsibilities. Second, there is a strict separation between the planning and control level with respect to capacity management and traffic control. It is a distributed responsibility.

First, the delegated responsibility. To prevent accidents and incidents between vehicles, they are separated in time, in distance and by visual detection. This creates a triple redundancy. Time tables, signalling systems and in-vehicle equipment should assure this separation and should support the observations and decisions of the drivers. These three principles are under pressure. High speeds make a direct outside observation impossible. To maximize the availability of capacity and interconnectivity of the network, a maximum traffic density is desirable. ICT applications offer huge opportunities for a rapid reconfiguration in capacity management and traffic process control. Dynamic control opens up the opportunity for maximizing punctuality and minimizing tracking times. Consequently, separation in distance is all that is left. This put high demands on technology and requires good faith of the operator in the supporting technology in case of 'beyond design' situations.

Second, in addition to this delegated responsibility there is another safety principle at a higher systems level, a distributed responsibility. We speak separately of traffic management in addition to traffic control. This separation is introduced in order to prevent a conflict of interest in a situation where one individual or authority should be responsible for balancing safety versus economy. By the increase of ICT opportunities for dynamic adaptation, this principle also has come under pressure. Full automation eliminates the operator and traffic controller, replacing them by computers, in which a black box defines what experience and expertise should be canned into computer algorithms, complying with predefined rules and procedures. Such a view captures any technological development at a rule based level of decision making. This should leave the knowledge based level of decision making solely to the responsibility of managers and governance. Full automation however, leaves non-routine situations which cannot be anticipated for in the design of the operations. They will emerge as unforeseen properties when the system has to perform under pressure.

However, this reductionistic view on full automation does not only remove all redundancy from the system, but also denies the operator a possibility to learn from experiences. Traditionally, there is an expert role for the captain of a vessel and pilot of an aircraft to deal with unforeseen situations. Their collective knowledge represents a capital for the sector which far exceeds the invested capital of each of the companies. By this feedback from practical experience, transport systems could develop into Non Plus Ultra systems: systems which could not be outrivaled because practical experiences were rapidly incorporated in adapted operations. The erosion of both delegated and distributed responsibilities leads to so-called sacrificial decision making. According to the process industry design doctrine, risk decision making is reduced to a single actor issue; one party makes the critical decisions for other parties too. If such safety critical decisions are not explicitly countered in the conceptual design phase or assessed at an institutional level, catastrophic consequences may occur in practice. Restricting strategic decision making to an Environmental Impact Assessment or a Cost-Benefit Analysis is not enough. Doing so, safety is sacrificed against environment and economy. There is a clear need for a proactive Safety Impact Assessment before such concepts are applied in practice. In long living systems within a global network context, such as railways, shipping and aviation, midlife updates create a dilemma from a systems perspective: technological innovation takes place in a saturated and matured system, but bears the burden of teething troubles, performance optimization takes place from an extended single agent perspective due to privatisation and public-private partnerships network configuration, focusing on short term efficiency considerations, while emergent properties are dominant in the overall systems performance due to the saturation phase, integrating new social, external requirements into the public assessment.

Fig 1.Technological development: the life cycle S-curve



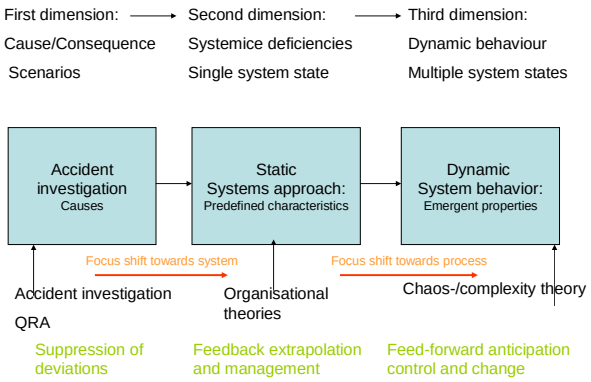

## 6 RESILIENCE ENGINEERING

Rather than speaking in terms of events with an undesired or unforeseen consequence due to which a substandard safety performance of a system emerges in practice, caused by a failure of a component or process, safety can be considered a normal consequence of performance variability. Safety should be achieved by controlling this variability rather than constraining it (Hollnagel 2008). Hollnagel defines resilience as: the ability to effectively adjust its functioning *prior to* or *following* changes and disturbances so a system can continue its functioning after disruption or mishap, while in the presence of continuous stresses. Therefore systems should be able to cope with responses to the critical, potential and factual situations. In order to grasp this coping, a transparency in various stable ad unstable system states should be available supported by the ability to predict, plan and produce.

But what if such transparency is not possible? What if we cannot cope with system complexity and complex causality or lack self-organising and learning behaviour? If we are not capable to come to an agreement on the causes of disasters, should we restrict ourselves to dealing with the consequences or the resilience capability after a disaster? If we cannot analyse the complex reality and cannot achieve consensus, are we deemed to restrict ourselves to a battlefield of subjective opinions, submitted to political will and governance resolve (Rosenthal 1999)? Or do we restrict ourselves to a lower systems level of a single agent and the organisational level, recovering from an undefined threat to its functioning accepting sacrificial losses? After all, resilience has its roots in loss control in warfare, where getting inside the enemy's decision cycle the offensive potential is preserved. In this context, the essence of resilience is defined as a company's ability to make sense of its environment, to generate strategic options and realign its resources faster than its rivals (Hamel & Valikangas 2003).

Secondly, what if a problem solving ability is lacking? Do we apply a pessimistic vision in which engineering design is reduced to a process mixing reuse of known solutions with some new technology and processes, with a sauce of varying thickness of creativity poured over them. Such a concept denies the potential of innovation and technology as a flywheel for progress despite political and institutional controversies (Freer 1949). In addition, the necessity to maintain oversight at the socio-technical systems level is lost. By Hollnagel's definition, resilience engineering has the potential of offering opportunities in solving complex problems by taking into account the dynamics, multidisciplinarity and complexity of systems. A transition into a third systems dimension becomes feasible in applying chaos and complexity theory to the concept of systems control and a re-introduction of the conceptual design phase in system change (Bertuglia 2005, Hendriksen 2008).
This third dimension identifies dynamic systems behaviour beyond the level of linear behaviour which can be explained by static system characteristics such as tight coupling or high reliability. Additional dynamic properties are identified such as deterministic chaos, emergent behaviour, self-organisation, self-conformity resonance and bifurcations. From a safety perspective, the most interesting parameter is the existence of multiple system states -which can be stable or unstable- which contain a characteristic safety performance level by their specific state. This eliminates the debate on acceptable and quantifiable system safety levels, replacing it by an insight into the various system states and their inherent dynamic properties, facilitating an overall assessment.

Fig 2. A third system dimension




## 7 TOWARDS A THIRD PHASE IN TRAIN CONTROL

This systems engineering potential has been demonstrate in a feasibility study into the deployment of a new railway concept beyond the boundaries of present railway configurations, aiming at doubling the number of trains for half the costs per passenger kilometre, maintaining the present safety performance levels.
Regarding the train control functionality, a new Free Ride concept was developed in analogy with the Free Flight concept in aviation, replacing the hub-spoke system by a direct origin-destination concept and automated data transfer for routine information exchnage. Four innovations were incorporated in the Free Ride concept:

- transfer or responsibilities for operational control and safety from traffic control towards the train driver from the perspective of delegated and distributed responsibility and a transfer from a track-bound control to a vehicle-bound control strategy
- replacing a strict hierarchical planning of capacity by a dynamic and flexible allocation and interactive management of disruptions and faults
- introduction of a self-learning software based on principles of Business Model Driven Engineering, Functional Request Specifications, Use Cases, Operating Envelope and elaboration of a Traffic Management Level in the ERTMS software architecture
- certification and validation at an integral systems level, replacing a repetitive upgrade and migration of component certification at a modular level.

The Free Ride concept eliminates the conventional conflict of interest between safety and control, by applying a performance based control strategy instead of a compliance based approach, restricting incident management and handling to the local level of the network.
In order to solve conflicts of interest between safety and capacity at a higher systems level, a new managerial arrangement is required in order to preserve long-term public values against private short-term interests (WRR 2008). In analogy with a Harbour Master and Airport Master, a Rail Master is allocated the strategic safety responsibility in the decision making on dealing with other system performance requirements such as capacity, economy, environment and energy. Finally, a new international, sectorial entity is required in order to assess safety at an integral systems level with respect to systems integration of track, rolling stock, safety assurance and systems certification.

In order to deal with such a systems innovation, on one hand an interdisciplinary approach is a prerequisite, creating synergy between three rationalities of design, social and technical sciences and practical expertise and experience (Stoop & Dekker 2007). Replacing a technological/substantive approach by a process/negotiative approach in which process drives out content, has created disaster, as demonstrated by the introduction of ERTMS in the Dutch railways. Combining three rationalities of a technological construct, a social construct and a local construct facilitates communication, learning and adaptation across actors at all levels and life-cycle phases of a system. Each rationality contributes to the overall systems design: from a technical perspective a dynamic modelling software development is required, from an local operator view, a new cognitive engineering modelling is a prerequisite for delegated responsibilities, from a social perspective new organisational and institutional entities have to be incorporated in the systems concept.

## 8 CONCLUSIONS

In assessing the safety of the ERTMS system development several conclusions can be drawn:

- actors with different rationalities are located at three different phases in S-curve, creating value and control conflicts. A need for a multi-agent process approach is emergent, but not sufficient;
- technological innovation creates major uncertainties: engineering is not a standard technology application which can be bought Off the Shelf: it also contains software design concepts change, system integration, oversight/consequence analysis and integral system certification;
- shifting from a technological perspective in systems development towards a social engineering is not sufficient; there is a need to integrate the technical, human and organisational/institutional design across the various system life phases, taking into account the various system states that may exist in practice.

Resilience engineering consists of three dimensions: technological engineering design, process design and a systems architecture dimension. Without such an encompassing approach, introducing ERTMS is nothing but deals on wheels.

## REFERENCES


Freer (1994). ICAO at 50 years: Riding the Flywheel of Technology. *ICAO Journal Vol.49 No 7, September 1994, pp.19-32*

Rosenthal (1999). International *Conference on the Future of European Crisis management. The Hague, November 7-9th, 1999.* Challenges of crisis management in Europe.

Hamel and Valikangas (2003). The Quest for Resilience. *Harvard Business Review, reprint R0309C, 13 pag, September 2003, wwwhbr.org*

Bertuglia & Vaio (2005). Non-linearity, Chaos and Complexity. *Second edition, Oxford University Press, Oxford*

Stoop et.al. (2007). HSL safety signalling system ERTMS. An independent investigation into the usefulness of adapting the ERTMS safety signalling system. *Commissioned by the Research and Verification Department of the Dutch Parliament.* Delft University of Technology, 23 May 2007 (In Dutch with English summary)

Stoop and Dekker (2007). Are safety investigates proactive? *33rd ESReDA seminar Future Challenges of Accident Investigation. Ispra, Italy, November 13-14, 2007*

Hollnagel (2008). Remaining Sensitive to the Possibility of Failure. *Ashgate Studies in Resilience Engineering*

Hendriksen (2008). Feasibility of the Complexity theory in learning from Naval Disasters. *Delft University of Technology, 2008*

Steenhuisen and Van Eeten (2008). Invisible Trade-Offs of Public values: Inside Dutch railways. *Public Money & Management, pp 147-152, June 2008*

WRR (2008). Controlling Infrastructures. An investment assignment. *Wetenschappelijke Raad voor het Regeringsbeleid, Den Haag, 2008* (In Dutch)