The railway as a socio-technical system: human factors at the heart of successful rail engineering

J R Wilson1,2, T Farrington-Darby3, G Cox1, R Bye4, and G R J Hockey4
1 School of Mechanical, Materials, and Manufacturing Engineering, University of Nottingham, Nottingham, UK
2 School of Safety Science, University of New South Wales, Sydney, Australia
3 Centre for Rail Human Factors, University of Nottingham, Nottingham, UK
4 Department of Psychology, University of Sheffield, Sheffield, UK

The manuscript was received on 10 April 2006 and was accepted after revision for publication on 1 December 2006.

DOI: 10.1243/09544097JRRT78

Abstract: High-quality engineering and operations management are key to meeting all the requirements of a successful railway—quality of service, reliable and safe performance, and maximum possible use of capacity. However, the railway is a socio-technical system and therefore has human factors at its core, which requires a strong integrated ergonomics contribution. Moreover, this contribution must be at a systems level rather than providing point solutions to particular equipment, interface, workplace, or job problems. This paper draws from the first two human factors projects in the EPSRC Rail Research UK programme, interpreting them for an engineering audience. The paper first emphasizes and gives examples of the need for a systems ergonomics contribution to engineering an improved railway. Then the available literature is summarized in a structured fashion. Finally, a short summary is provided of the research which has started to develop a distributed cognition model of work on the railways, especially across functional groups of signalling, control, and train driving.

Keywords: human factors, systems ergonomics, rail network control, distributed cognition, rail signalling

1 INTRODUCTION AND BACKGROUND

In every country with a rail network of any importance, the relevant operational, regulatory, and government bodies are trying to achieve something similar. This is to move more people and goods, on time and safely, to the satisfaction of their customers. The need is for a more reliable, higher quality, and safer railway within a system where there are considerable restrictions on capacity. In the UK for instance, high speed, cross country, commuter, and freight services must share a limited track capacity with each other and with those carrying out emergency repair, planned maintenance, and major enhancements and renewals. Although engineering and operations improvements are vital, these can only be achieved through understanding and integration of the key rail human factors.

This need to carry out rail engineering with an integrated input from human factors has been understood in terms of fundamental and applied research. For fundamental research, the main European research network, EURNEX, has one of its ‘activity poles’ dedicated to human factors (with a different one dedicated to safety, emphasizing that human factors is about far more than safe systems, important as these are). The UK research network, Rail Research UK (RRUK), subject of this special issue, is predicated to an extent on a strong engineering and economic research foundation, but has had human factors projects as part of its core since its inception.

At the level of practical application within rail engineering, one example is the Ergonomics National Specialist Team of Network Rail (the UK rail infrastructure owner). This team includes ergonomists,
This system is regarded as a socio-technical system and the head of ergonomics reports directly to the Chief Engineer. As a simple example of the perceived importance of human factors for rail and of the value Network Rail place on ergonomics contribution, the team has grown in 5 years to 13 members with another four attached PhD students.

For some readers, their familiarity with ergonomics/human factors may be limited to its (invaluable) contribution to providing dimensional and performance information for design of equipment, interfaces, and workspaces (e.g. train cab design), working environments (e.g. signal boxes), minimization of risk from manual handling (e.g. track working), and design of information display systems (e.g. lineside signage or train movement displays). However, ergonomics increasingly works at a systems level, being central to systems engineering. The drive behind this comes from contributions such as better understanding of organizational failure in accidents [1] and acceptance that cognitive task performance is situated in a setting [2] that influences it strongly, and is spread across people, places, and times [3]. This has led to systems level human factors studies of air traffic controllers [4], emergency crews and controllers [5], planners and schedulers [6], and critical care staff [7]. This human factors approach is also now central to study and improvement in the rail network.

In essence the railway is a large, complex distributed socio-technical system with many difficult (and some easier) engineering problems at its core. This system is regarded as a socio-technical system because it meets the classic criteria for this: it is a purposeful system that is open to influences from, and in turn influences, the environment (technical, social, economic, demographic, political, legal, etc. [8]); the people within it must collaborate to make it work properly; and success in implementation of change and in its operation depends upon as near as possible jointly optimizing its technical, social, and economic factors. It is an excellent example of a modern complex socio-technical system – the health service and emergency services are offers – which has many more layers of complexity than the traditional focus of socio-technical systems in the manufacturing industry. Events, operations, people, and technical systems are widely distributed in time and space: they are often mobile, people must collaborate using refined social as well as technical skills, and the system is distributed across regional, national, and cultural boundaries, leading to additional problems of interoperability [9]. Designing and building for success in such socio-technical systems cannot be solely the province of engineers. An interdisciplinary effort is required in which the discipline of ergonomics (or human factors – the terms, disciplines, and professions are one and the same and are used interchangeably, as in this paper) is central to understanding and improving such systems.

This paper draws from work carried out to date on human factors projects within the EPSRC RRUK. From the first project (B2 – [10]), the paper makes the case for ergonomics/human factors within rail systems engineering, and then points the reader to the large and varied literature available in rail human factors. Knowledge valuable to the design and operation of a human-centred railway is summarized, including work on human factors integration. Then the paper introduces the work carried out in project B3 [11] to model the interacting work activities across three key functional groups – signallers, controllers, and train drivers. The purpose of this paper is to stress, for an engineering audience, what is meant by human factors (at the systems, social, cognitive, and physical levels), and bring to the reader a large quantity of the good rail human factors work that has already taken place. The purpose also is to set the scene for a wider understanding of human factors; that is, the need to move beyond understanding performance, behaviour, and task/system fit for individual workers in rail and to examine, understand, and improve work in distributed teams in future.

2 PEOPLE, ENGINEERING, AND THE RAILWAY

People are central to all rail activities, from planning and (re)building the network to operating the infrastructure to using its services, and so the interactions that are the province of ergonomists – people–people, people–tasks, people–equipment/software, people–environments, people–organizations – are many and varied. The stakeholders include signallers and controllers (electrical, infrastructure fault, and traffic); drivers; station and on-train staff; planners, engineers, and managers; track (maintenance) engineers and workers, lookouts, and site safety controllers; and passengers and the general public (the last both legitimate – e.g. at level crossings, and illegitimate – e.g. trespassers). A simple representation of the interconnections between all these people and functions can be found in reference [12]. The rail system tasks include vehicle control, systems process control, monitoring, planning, and physical work, occurring in settings such as vehicle cabs, control rooms, outdoors, and large buildings and spaces. The artefacts used include VDUs, signals, paper, CCTV, hard wired controls, handtools and large engineering plants, and vehicles. Therefore, the roles and interactions involving people are multiple and so the human factors contribution must be multiple also in order to contribute to better design of interfaces, tasks, systems, equipment, jobs, and environments.
Such multiple human factors contribution, within an integrated systems ergonomics perspective, will support engineering of a better railway. Two examples of this need to account for ergonomics to support engineering at a systems level are given next.

First, considerable efforts are being made to introduce improved maintenance and inspection methods and regimes. The capacity, reliability, and safety of the railway may improve dramatically if better sensor-based inspection technology, greater functionality in road/rail vehicle and engineering trains, and improved design of track workers' tools and equipment are deployed. To be effective, the design of all of these new technical systems must fit the real needs and capabilities of their users and must be human-centred. Possibly the greatest potential for improvement in maintenance (whether small reactive repair jobs or large and long route enhancement projects) will come through improved planning, communications, briefings, and control of engineering work, which requires a considerable human factors contribution to the design of the relevant work systems and work organisation [13].

Differing numbers of specialists from different functions have to come together into teams with varying degrees of integration, and must coordinate and cooperate together while distributed over time and space. The roles include planners and managers, engineering supervisors, persons in charge of possession, controllers of site safety, overhead line engineers, trackside workers, engineering vehicle managers and drivers, and inspectors. The wider collaborating network will also include signalers, controllers, and passenger and freight train drivers. Among the studies undertaken has been a recent one on the work of engineering supervisors [14]. The nature of their role means that their activity is, or ought to be, highly collaborative. They are responsible for a number of work sites along a length of track that can be anything from a few metres to several kilometres. Their collaborative activity stretches over time (translating outline and detailed plans made hours, days, or weeks before being brought into activity on the shift, and handing back a safe track to the railway operations) and over space (with their work gangs spread out along the track, engineering and road/rail vehicle drivers coming into and out of the site, and the signalers and controllers many kilometres away). The teamwork is virtual in the sense that they are not all co-located and work for several different companies. And they and their gangs are mobile. Early observations and interviews revealed the critical aspects of the role that include communications, shared planning, and conducting briefings, all central to collaborative work.

In such settings, a conceptual framework of the issues must be built up in parallel to the collection of empirical evidence, and this underpins understanding and eventual recommendations. Tasks and interactions between all the actors concerned were observed with different degrees of detachability and obtrusiveness according to circumstance and the requirements of rich data collection. A variety of interview techniques were employed, dependent on whether the focus was individual performance or group/collaborative behaviour. The critical part of such studies (apart from gaining access in the first place) is making sense of all the rich data from various sources: combining, sorting, reducing, and representing these in a manner so as to be useful, but at the same time so that there is a traceable evidence trail [15].

Subsequently, the authors have become heavily involved in a major Network Rail initiative to develop new processes, procedures, and rules for all levels of engineering work. The ergonomists are central to this effort. The project managers have recognized that before they start to develop new work systems there must be a benchmark for the current situation. This has involved a long programme of site visits, observations and interviews, and mixed function workshops, to construct a functional analysis across planning, access, engineering train movements, delivery, and handover. As phases from this are validated with further structured workshops, the authors have set up a series of Human-HAZOP (hazard and operability) workshop sessions, in order to identify current potential risks (for safety and for engineering work performance), opportunities for failure, and safety controls. This methodology and the benchmark information will then be in place for comparison with similar (if more predictive) exercises carried out for proposed new methods of working, in order to better predict likely improvements in effective and safe systems of work.

As a second example, across Europe, politicians and some rail industry insiders are waiting for the implementation of the European Rail Traffic Management System (ERTMS) to allow substantial improvement in the control of the rail network (although some rail industry insiders are somewhat less sanguine however!). In whatever version it may be implemented (ranging from full levels of automation through to various forms of intelligent decision support system and new communications networks), ERTMS should allow trains to run with closer separations but more safely, although acknowledging concern over its cost effectiveness given the enormous investments required. All this investment will be futile unless the new systems are implemented with a full understanding of how skilled experienced signalers and controllers operate the system now and of how
to rationally implement human-centred automation. Human factors have historically identified problems that occur because of a belief that automation will solve operational problems, or that better technical systems per se will improve performance. Bainbridge’s well-known ‘ironies of automation’ [16] have been found in many industries. These include: (a) automation, often regarded as being highly reliable and having cost advantages over human-centred systems, rarely provides all the functionality promised, does fail, and recovers less well from this than do people, while often costing far more than was promised in development and running costs; (b) the very people that designers believe are unreliable and therefore should be designed out, in fact keep many operational systems up and running, despite the failures of technical systems; (c) the very skills that people can bring to a hybrid (automation plus people) system – of intervention, intuitive, and deductive reasoning, problem solving, etc. – are those that are hard for them to develop and maintain if the design philosophy is for them to be ‘monitors’ only. Automation and ‘technical fixes’ rarely deliver all they promise for the price they charge. There are always gaps in the capabilities of computers and machines; when they fail it can be catastrophic (for performance continuity if not always safety), and heedless implementation can leave poor quality jobs for the people in the system, which render them incapable of contributing their expertise to optimize system performance or to cope when things go wrong (the ‘out of the loop’ syndrome).

The authors would argue for people involvement in operating complex systems, especially where there is a cadre of highly experienced and skilled and motivated staff. Why under-utilize such a valuable resource? The role they would play is that of supervisory controllers. However, what happens if experienced and skilled employees leave the industry, and the workforce profile is of many new recruits in signalling and control (as is found also in industries such as steel, aerospace manufacturing, and medical technology as well as rail)? If technical system change ‘divorces’ the operator from the process they are operating, then they may never pick up deeper levels of expertise. If entrants into the rail industry arrive with lower level technical skills and worse social skills, then will they pick up the technical and social network knowledge needed to intervene and optimize? If employees are likely to leave jobs after only a few months or years, how can mutually supportive work groups operate? The understanding that will be needed requires new studies of levels of signaller supervisory control and ways to avoid the ironies of automation and will build upon thorough studies of signaller and controller competences and workload [17–19].

3 HUMAN FACTORS STUDIES OF RAIL SYSTEMS

The extent of rail human factors research and, as a consequence, application has fluctuated over the years. In part, the level of effort has paralleled society and government (dis)interest in the railways, and consequently levels of investment. There has possibly also been a perception that the railways, where apparently nothing much changes quickly and where most problems are resolved easily, is not a dynamic and exciting environment for research. Now pressures from the technical, organizational, safety, financial, and political climate [20] have created a clear need and welcome for high-quality human factors to support analyses, developments, and change implementation. Across the world in fact, rail human factors research and application has undergone a renaissance, funded by both research grant awarding bodies and by industry itself.

The authors have recently produced reviews [20, 21] and collations [22, 23] of rail human factors literature. See also the annual Rail Safety and Standards Board CD-ROM of human factors research. In the rest of this section, the authors summarize some of this literature, relating it to different types of rail engineering endeavour. These relate variously to signalling (primarily electrical, mechanical, and control engineers), telecommunications (electrical and control engineers), electrical and plant (electrical and mechanical engineers), buildings and structures (civil engineers), track (civil engineers), cab and rolling stock (design, mechanical, and electrical engineers), IT systems (software and control engineers), and delivery processes (systems and manufacturing engineers).

3.1 Audit of work systems ergonomics and attitudes of staff

Before redesign of any facility or implementation of new systems, there is a need to assess, across a wide range of work environments and systems, current ergonomics factors. This might be through expert assessment or audit of workforce knowledge and opinions. An example of the former is Network Rail’s Baseline Survey [24] and of the latter is the Railway Ergonomics Questionnaire [25].

3.2 Understanding train driver behaviour and consequences for performance

A major concern of rail engineering is the design of train cabs, cab systems, and lineside information systems. Once general ergonomics information and feedback from all staff is available a deeper understanding of how and why people in key rail functions
behave, the consequences for performance, and the implications of systems change for such behaviours is required. At a general level, current research addresses the fundamental elements of the train drivers’ role and performance, including their route knowledge and the underlying psychological components of train driving [26–28]. Much research in this area has investigated the potential causes of human error [29] and the extent to which the in-cab environment supports the driver’s ability to maintain situational awareness [30]. A part of reducing potential for driver error and of increasing their effective (on-time) performance lies in the design of their jobs and job aids [31] as well as understanding and optimizing – neither too high nor too low – their workload.

In particular, there has been a broad and relatively well-studied research field of train driver vigilance and perception, their recognition of and acting upon signs and signals. This includes also investigations into signals passed at danger (SPADs) and the appropriate design of signage and signalling systems (see later). Of all rail human factor topics, these probably have been the most studied over many years, going back to the 1960s and 1970s [32, 33] and even before. Recently in the UK, motivated by the Ladbroke Grove rail crash and by reports of incidents not leading to injury, there have been various studies of SPADs [34, 35], predictive tools [36], and development of tools to identify the risk of SPADs at different signals [37, 38]. Because of the desire of some to find a technological fix to such incidents, there has been related research into the use of vigilance devices and reminder appliances [39, 40]. Modern observation techniques such as the measurement of eye movements and of direction of gaze allow interpretation of drivers’ behaviour and of the possible reasons for it [41, 42].

One use of eye tracking is as another way to investigate the onset, manifestation, and consequences of fatigue (e.g. dwell or fixation times will become longer as people get fatigued). Studies here have examined the prevalence of sleep apnoea [43]. Research related to fatigue has also examined the effects of work rosters [44], used observation and self-report to study the effects of long (>6 h) journey times [45], run simulator studies [46], and developed checklist tools such as the Fatigue Index [47] as well as prototypical preferred roster patterns [48]. Related to impairment through fatigue is the incidence and effects of drugs and alcohol use on performance [49].

### 3.3 Understanding behaviour and performance in rail traffic control

From the point of view of electrical and control engineering for new rail traffic control centre systems, human factors research was probably scarcer from the 1960s to 1980s than for train driving, research, but included studies of the cognitive processes [50] and collaborative processes [51] involved. In recent years, a number of work systems design and human performance concepts have transferred from other industries and systems, aviation in particular, into the railways. This research has dealt with the mental workload of signallers [17, 18, 52]; teamworking and situation awareness [53, 54]; reasoning [55]; expertise and competences [19, 56]; and information interfaces [57, 58]. At one level above signalling and control, human factors is also now becoming a part of the study and understanding of the work of planners, for instance in timetabling, organization of possessions (of the track, for maintenance), and for emergency handling [59–61]. (Note that those involved in controlling the movement of trains are called variously in different countries controllers, signallers, planners, and dispatchers.)

### 3.4 Design of human machine interfaces

Systems engineering meets electrical and control engineering in the development of human–machine interfaces, generally involving interaction with computer systems. There have also been systematic attempts to understand and model the driver’s activities in detecting, recognizing and acting on signals and signs, and consequently to provide a rational basis for positioning of lineside information [62, 63]. There have been many recent efforts to improve design of train cabs and of the information interfaces within them [64]. There are also contributions to do with significant changes to the information interfaces in rail traffic control and signalling [57–59] and in use of modern and personalized IT to support maintenance work [65, 66]. The advent of ERTMS will mean that a major effort is required to provide appropriate information interfaces [67], especially in light of the changed nature of communications required between signaller/controller, driver, and maintainer.

### 3.5 Automation

As indicated earlier in the paper, a number of new technical systems are at the proposal, feasibility testing or pilot implementation phases: for example all the systems to do with ERTMS – and this will mean radical changes in the work and role of rail staff. Human factors contributions to the debate and to systems design and implementation are appearing [68], and a particular key contribution will be in migration and parallel running with old and new systems [69, 70]. More generally, there is considerable debate about how to integrate useful components of
automation (for instance, in decision support) with human skills to provide more reliable and effective total railway systems [71].

3.6 Track engineering

A relatively neglected area of rail human factors until recently has been engineering on the track—inspection, maintenance, and renewals. The contracting out (and sub-contracting and sub-sub-contracting) of this in the UK, and the accidents at Hatfield and Potters Bar, brought new focus onto the area, and there has been a renewal of research with a variety of approaches and methods being used. Examples are assessment of maintenance communication errors [72], track workers’ safety culture [15], attitudes to work (and their relation to accident rates [73], and the performance of various key functions such as that of engineering supervisors [13].

3.7 Human reliability, reporting systems, procedures, and violations

Because of the recent great concern among public, media, and government over rail accidents and safety, this application domain has become the focus for advances in reporting systems [74–77]. Recent years have seen the transfer from industries such as nuclear and aerospace of human reliability frameworks and of human reliability assessment (HRA) methods to identify potential for error especially related to train driving [78, 79]. Indeed, rail has been the focus for improvements in HRA techniques, which should then transfer back into other domains [80].

Every rail operating country has a ‘Rule book’, a set of operating procedures that in many cases has grown in an unstructured fashion, with rules being generated and added to deal with concerns as they arise, but with few attempts to rationalize these or to assess for consequent inconsistencies or redundancies. The impact of procedures—for good or ill—is recognized by the human factors community, and research and application work to improve these is taking place in several countries [81]. Linked with this are efforts to improve understanding of violations (deliberate breaking of the rules, often to improve performance or even just to get the job done at all), which may or may not be ‘caused by’ inappropriate rules [82], and also of safety culture, generally both within organizations and also across national or company boundaries [9].

3.8 Design for neighbours—passengers and public

Moving away from rail staff and looking at other stakeholders, a reasonably long-standing and continuing theme of rail human factors research has been in ride quality and passenger comfort [83, 84]. Support for passengers also comes through the interfaces to the information and ticketing systems with which they are provided [85, 86] and the design for their movement around stations and in boarding and alighting trains [87, 88]. In less positive circumstances, there is also a need to understand how people behave in emergencies and how to best help them evacuate carriages when necessary or to support their rescuers [89–91] and to design carriages to reduce the chance of injury [92]. Even more removed from the proper functioning of the railway, it is a sad fact that for some in society their contact with the railway is through trespass, vandalism, crime [93] and suicide—which added together comprise by far the greatest cause of deaths on the railway [94, 95]. This is highly related to a current area of great concern and human factors attention, where the public and the railways interface at level crossings [96, 97].

4 HUMAN FACTORS INTEGRATION

Much of the preceding section concerns generation of human factors knowledge—to do with research. In the past few years also there have been considerable strides made in application of that knowledge within structured design processes. Guidelines and standards are being produced that are increasingly appropriate to rail application in order to guide those planning and engineering the networks of tomorrow [98].

Increasingly, in some industries, and especially in the military, such standards and guidelines are applied within the framework of a Human Factors Integration Plan (HFIP), which itself will usually be specified within a HFI Standard (HFIS). One part of the early RRUK project work was to assess the impact of HFIS/P within the UK railways. This required confidential interviews with key people in rail engineering and rail human factors, because the written detail on such plans and standards is often hidden in commercial project documentation (although see references [99, 100] for some published information).

The rationale for HFIs is that for human factors to be adequately addressed, it is essential that it is managed as integrated within the whole project rather than as something bolted on as an afterthought. HFIPs are literally the formal definition of how human factors will be integrated into a system life cycle and a description of the related assurance procedures. They are aimed at project managers, design and construction engineers as much as at human factor experts. They may be developed and applied just to particular projects or system designs, but more usually will be generic documents produced by an organization or group of
organizations, which can be operationalized for any particular project or system.

An HFIP will often define who is involved and with what degree of responsibility, who the stakeholders are and how they are to be engaged in the process and their views accounted for, and how human factors expertise will be coordinated and maintained. Appropriate standards, guidance, methods, tools, and analyses will be defined, and operational concepts and requirements for human performance will be described and explained (see reference [101]).

The interviews that the authors carried out with HF specialists and systems engineers found evidence for increasing use of HFIP within rail. HF specialists at Network Rail and Rail Safety and Standards Board have observed HF input into business processes through their organizations, with HF awareness at board level. HFIPs based on company standards and guidance documents are in many cases used to bring about HF assessments, problem identification, design solutions, and evaluations at project and programme level. They are being utilized to ensure that HF remains on the agenda as project plans emerge and when changes and critical decisions are made.

HFI can be problematic, and the challenges faced using HFI and HFIPs in rail have been mirrored in all industries where there are different disciplines working together to ensure a design solution is delivered on time and in a cost-effective way. The worst case scenario is when HFIPs are drawn up but not used—thus meeting the requirements of contracts or standards but without actually impacting upon systems ergonomics. The challenge is to ensure that high quality and appropriate HFIPs are drawn up and agreed by staff with the right level of HF competence and that this process includes defining appropriate timing and levels of HF intervention. Suitable means by which to measure HFI success is also a current focus.

One of the difficulties of HFIPs and HFISs is, to paraphrase the words of a number of design engineers and operational staff, that it may be a very worthy document but the document itself, and particularly putting it into practice, can be akin to ‘pouring concrete into the veins of creative design’. In words from elsewhere, from the head of an ergonomics group at a UK rail organization, the last thing that the human factors team, the engineers they work with or the project managers want to do is have a 600-page search for one or two key design aspects. As a consequence, there has been a strong preference in rail for a more ‘light touch’ approach to HFI, recognizing that a systematic process will reap benefits, but only if it does not become a burden [102].

5 STUDY AND FIRST MODEL OF THE WORK OF CONTROLLERS, SIGNALLERS, AND TRAIN DRIVERS

As proposed earlier, and without wishing to overstate the case, there has been something of a shift in the approach of human factors in domains such as rail over the past few years. The change in perspective has been coherent with systems engineering, with human factors taking a holistic, socio-technical view: socio-technical for the reasons given earlier, and holistic in that all aspects of human performance and interaction with systems—physical, cognitive, social—are accounted for in an integrated fashion and across the system life cycle. An important corollary of such a view is that the systems are examined and analysed in the light of their operations being distributed—temporarily, spatially, and functionally.

Activity in a system will take place continuously or discretely over a period of time and the artefacts or people of interest through that system over time can be ‘followed’ (an equivalent here is following the ‘patient journey’ through a health system). Action in such a system also takes place over a wide geographical area (an equivalent here is studying the work of a forest fire fighting crew as they carry out their work in a command centre, on the ground in the forest, and from the air). Activity is also functionally distributed (for instance a design and development team containing concept designers, engineers, architects, stylists, production engineers, systems engineers, marketing specialists, and customer representatives). One good example of this trend to take a systems viewpoint when studying and making design improvements to rail systems can be found in the area of design for passengers at stations. Early work to examine how to support passengers through architectural design or information display carried out classical studies of foreground and background design of signs, placement in key positions, etc. More recently [88], such studies have tracked passenger journeys, from leaving home or at least arrival at the station through all aspects of access, information finding and train boarding, and then subsequently alighting at the other end. Within this approach, the perspective has been of universal access or design for all.

In order to understand the railway as a distributed socio-technical system, it would help to have systems level models (including those of distributed cognition) of the whole rail system and human–human and human–artefact interactions and representations [12, 103, 104]. The early projects of RRUK made a first attempt to model the interactions of rail functions and operations on the UK network from a systems ergonomics point of view.
The methods employed within the studies of the rail network—concentrating in turn on driving, signalling, and control and then on the interactions between them—have built upon those developed and used by the authors in other studies and contexts, for instance in planning and scheduling, and especially in earlier work on train driving and rail network maintenance. The process used here by and large consisted of divergent (spreading the net widely) and convergent (honing in on key issues) information-gathering phases, using qualitative approaches and methods from ethnography, sociology, and field study of human factors. The relationship with ethnography is in recognition of the impact of work culture, the fact that the study observer is a participant in the workplace too, and the need for traceability of the interpretations made. Critically, the authors have been informed and supported, especially in validation, by the involvement of subject matter experts (SMEs) throughout [105]. A more detailed explanation of study methodology with respect to the study of controllers appears elsewhere [19].

The multi-method approach adopted accessed data and findings at different levels of detail and proximity to the individual operators, referred to as levels of analysis. The approach has many parallels with the well-known cognitive work analysis (CWA) [106], a framework for the analysis, design, and evaluation of human machine systems; this is primarily a formative (in design) analysis approach rather than providing a purely descriptive or an unrealistic normative analysis. The five CWA phases of analysis are work domain analysis, control task analysis, strategies analysis, social organization and cooperation analysis, and worker competency analysis. A full CWA is known to be time consuming to perform. The domain analysis is usually the most substantial (at least in published studies) and involves a detailed breakdown by levels of abstraction of what goes on in the domain and is decomposed into different parts of the system. Because the authors believe that CWA is more applicable to control room type work that takes place in one setting, is usually most relevant to studies of the human–computer interface (although see reference 107), is difficult to apply to the evolutionary design of existing systems [108], is often based upon simulator studies, and does not sufficiently recognize that the social aspects of work are central rather than one element, the authors decided that it was not appropriate to conduct a full CWA in the RRUK model of rail operations. However, core theoretical constructs from CWA and from naturalistic decision making [109] underpinned the research programme and many of the techniques were adapted to develop the model framework.

The model framework itself was flexible in the early stages of the work, being modified to account for practical or theoretical issues encountered during the fieldwork and also to account for different perspectives and needs when investigating, in overlapping studies, control, signalling, and driving. In the early version used to guide data collection and analysis, the model framework had stages of domain (or system), function, activity, task, task element, and process analysis, and in this form it helped to guide the questioning process of the study (Table 1). Subsequently, the model was refined to be a first framework of distributed cognition in controlling the network, and the stages (or levels) modified somewhat in order to be of explanatory value in understanding the complex interrelationships and expertise needed to perform successfully (see next section).

By moving through different phases of analysis and with each one building on the other, a picture of the system under investigation is formed. The intention was to build a picture of rail operations that informs design decisions in a way that acknowledges the adaptiveness of the human in dealing with complex and unanticipated situations as well as supporting them better through their development within a rail operations role.

The first study stage was to set up site access and the planning of initial site visits to meet managers. In addition, these first contacts enabled collection of some data for later analysis at the function level. Subsequently, managers (driver, control centre, and signaller) were interviewed at a number of sites to provide an overview of the roles of the different functions within the industry, how work is organized, what the main roles are, and the background of the relevant staff (function level findings). (The research reported here differs somewhat from the general case of CWA and ethnographically informed field study, both because the authors already had a very good generic domain understanding through their previous rail research and also because the domain organization, Network Rail, were also partners in the research.)

During site visits, the researchers drew control facility layout diagrams, marking the position of the different types of operators and the equipment on their desks (function level findings). In addition, there were initial observations made at the worksite; first discussions held with job holders, managers, and SMEs; and first analyses of available documentation (work/activity level findings). For these interviews and all subsequent observations and interviews, it was emphasized to participants that the projects:

(a) were carried out by an independent research group;
(b) had a clear system for holding original data confidential, and a firewall between the research
group and the client (the employer of some of the
participants);
(c) would have arrangements for feedback and
dissemination of results agreed in advance;
(d) did not identify individuals associated with what
was said or observed for inclusion in the reporting
of findings.

A second wave of longer site visits or train cab rides
took place at a number of sites/routes (for instance,
for control alone, five control centres, 24 shifts, and
216 hours observation). This phase was focused on
detailed data collection from individuals, captured in
the form of field notes written during observation and
‘on the job’ interviewing.

Observation and interview were also used to inves-
tigate the culture and ways of working, considering
the interactions of the individuals within their func-
tion and with those outside (e.g. drivers with sig-
nallers). The interview process became more efficient
and more structured as the visits proceeded. Observa-
tions were collected when sitting with individual sig-
nallers, controllers, or duty managers, or when taking
cab rides, but the work of others was also noted in the
observation records. The notes distinguished the con-
temporaneous observations from the comments or
interpretative notes that the researcher made regard-
ing an observation at any time subsequently. Data
collection was in the form of recorded data points
within field notes. Data points consisted of:

(a) directly observed actions or events such as com-
munications and use of information systems or
manuals, and responses to incoming information
or communication;
(b) responses to questions from the investigator,
where possible a word-for-word recording verified
with the participant at an appropriate time;
(c) diagrams and visual representations produced by
the investigator or the participants; many dia-
grams of infrastructure layouts with the relevant
tags were produced or used during interviews with
controllers, for instance;
(d) copies of notes made by participants during their
work.

After collection, data were treated through a system
of emergent coding to identify features of controllers’
work in a more generalizable form, given the context-
based method of data collection. To begin with, the
codes were at a very high level based on Fig. 1. A sec-
ond wave of coding occurred where more detailed
information was required. For example, the ‘work
type’ code was decomposed to give subcodes, of
‘cognitive work’, as situation awareness, distributed
cognition, problem solving, decision making or mon-
itoring, or again the ‘activity’ code was opened up
and decomposed into the types of events dealt with,
for example, train failures, infrastructure failures,
fatalities and injuries, requests for possession of the
line by others, responding to a query.

<table>
<thead>
<tr>
<th>Level of decomposition</th>
<th>Definition</th>
<th>Typical questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>The system being considered in terms of core business, key players and physical nature</td>
<td>What is the key business for UK rail? What is the basic structure of the industry, and who are the governing bodies and component organisations? Who are the key players? What are the components and features of the rail system? What are the roles of each function in rail operations? How are functions organised to satisfy top level goals and contribute to rail service delivery? What are the features of each function's work systems, e.g. staff, environments, work settings, technologies, etc?</td>
</tr>
<tr>
<td>Function</td>
<td>The role of the different operational functions in the domain</td>
<td>What are the main activities for a given functional role? Which activities will provide focus for task level analysis? What interactions/collaborations/conflicts occur? What are the main tasks that are performed? What interactions, collaborations, and conflicts occur? What are the key steps (3-6) that make up each task? What is the operator doing in each step and why? What information and knowledge is required? How do operators know what information they need, where it is located and how it can be accessed? Where are the lines of responsibility for each of the tasks? Can fundamental HF research be used to describe the essence of what is occurring in the task elements? Are there any common or generic processes occurring? Which aspects of the process are most challenging?</td>
</tr>
<tr>
<td>Work activity</td>
<td>The high level categories of activities performed within a functional role</td>
<td></td>
</tr>
<tr>
<td>Task</td>
<td>Tasks that are performed to satisfy work activity</td>
<td></td>
</tr>
<tr>
<td>Task element</td>
<td>Task elements</td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td>What are the processes that underlie the tasks?</td>
<td></td>
</tr>
</tbody>
</table>
The next level of investigation was to consider, at a process level of analysis, how work activities and tasks are actually performed. Rather than explore every task and task step through hierarchical task analysis or an exhaustive cognitive task analysis (CTA), the investigators made use of the task diagram interview described as part of the applied CTA (ACTA) method. This technique offers simplicity and speed in eliciting the processes underlying the key activities previously identified and, importantly, identified the challenging aspects of the work. The task diagram involves decomposing the activity into three to six task steps or subtasks, in order to maintain a level of superficiality at an early stage. SMEs who were motivated to participate were selected for interview; who were articulate (this is essentially a verbal exercise with limited time); who were experienced in a variety of positions or roles, providing a spread of routes types of train and/or sites; and were able to provide a rounded picture or the job. These requirements, especially to be experienced and articulate, are likely to make these non-representative samples, but the needs of the study were to gain maximum insight and broad and deep knowledge rather than an ‘average’ knowledge.

The interviews lasted between 1.5 and 2.5 h. The general opening question in each case was ‘Tell me about your job – What is it that [your function] do?’ Then example activities, planned and unplanned, were introduced into the interview. The identified tasks were decomposed into three to six steps, and the task diagram that emerged was one that was found to be repeated for all types of events and incidents. To test this assumption, the participants were asked to continue the interview by listing all the types of incidents that made up their work and to verify that these were covered and could be explained by the task diagram. The SMEs were also questioned about any other activities that did not fall into either the task diagram or the coding category description of ‘managing incidents or events’. The task diagram that emerged was verified further by several additional experienced SMEs.

6 DISCUSSION – TOWARDS A DISTRIBUTED COGNITION MODEL OF RAIL OPERATIONS

This paper has made the case for the central role that human factors and systems ergonomics play in understanding and helping to develop an improved rail network and rail service. The relevant literature has been summarized at some length in order that the reader can make a first search for background knowledge relevant to their own interests and needs. The substantial study of the interrelated functions of signalling, control, and driving that is reported has produced many findings and interpretations that are currently being used in defining fundamental operational principles to inform any future radical redesigns (of control technologies for example), and to exemplify the strategies of ‘experts’ in order to feed these into training and competency programmes. The method itself, and the underpinning model, is also being employed to develop a tool useful to frame future related research.

Cognitive processes, such as decision making, inference, reasoning, and learning, are central to the safe and efficient operation of rail systems. Traditionally, cognitive ergonomics studied and developed designs on the basis of an understanding of these processes for individuals interacting with artefacts such as control panels or VDU displays, and did so often through laboratory or simulation experiments. More recently, in recognizing the prevalence of work in complex distributed systems as described early in this paper, cognitive ergonomics frequently studies this work from the perspective of distributed cognition (thinking spread among different people and computers) and ‘cognition in the wild’ (literally thinking as it occurs in real and messy settings). In the case of rail, as one example, signallers communicate with drivers, signallers in other sections and colleagues in control in their own and other companies, and make use of VDU-based control systems, printed simplifiers, paper manuals, and so on. This makes any reasonably valid laboratory study of such work difficult if not impossible. It also means that in field studies, it must be understood how the temporal, social, and material distribution of cognition supports and constrains operator performance.

The model framework used to guide data collection and analysis has been modified and is now the basis for a web-based distributed cognition tool. This will neatly represent our key findings as regards the knowledge, communications, and expertise that
the functional groups studied bring to their roles in making the rail network actually work. This web tool is being examined in cooperation with people from the industry in order that it be practically applicable and useful. The first representation of the model for the tool is shown in Fig. 1; it now has stages for domain (or system), function, activity, scenario, and strategy analysis. The levels can be explored vertically as well as horizontally, each level being associated with specific questions that were or could be asked.

Level one analysis is of the rail domain. For the purposes of the model this provides an overview of the industry and the key stakeholders. The importance of this level of analysis is that at the top level, purposes and values for the system are made explicit. This will enable operators to make sense of new situations and to act to meet the target states required. This level of analysis can also reveal where different stakeholders may contribute with different secondary objectives to meet an industry-wide common objective. These differences can be traced down through the subsequent levels to explain the different needs to potentially collaborating operators in different companies.

At the next level key functions such as signalling, driving, and controlling are described in terms of their contributions to the domain level but also their collaboration with each other and the nature of the roles within each individual function. This forms an operations layer which, by eventually including elements such as maintenance and planning, will represent what it takes to run the railway on a day-to-day basis. This is an important level that links the vast, complex, and open socio-technical rail system to the functional groups that make-up rail operations. This level provides the physical, technical, political, historical, cultural, and social context of more detailed work analysis at the lower levels.

The authors have developed a work activity level of representation. The nature of the work within a function and the categories of activities are described here. This level is the bridge between the role of the function within the domain and the tasks that operators in a function perform to achieve the work. Task level analysis in the model focuses on the types of tasks that form the work of operators and, in the version shown here, is subsumed within activity. Instead of detailed step-by-step prescriptions, the tasks are viewed in terms of task types and key steps, activities, or processes that are considered challenging by operators.

Scenarios at the next level are developed together with the SMEs and other relevant job holders. These provide a platform to better understand, in rich detail, the nature of people's work, the settings and the constraints, and the skills and knowledge that signallers, controllers, and drivers bring to bear. In particular, these scenarios are used to probe challenging aspects of tasks, which then provide data on strategies, knowledge, skill, expertise, situation variation, and task idiosyncrasies. For an understanding of distributed cognition – between signallers, controllers, and drivers, and their computer and other interfaces and artefacts – the authors are particularly interested in the strategies that they have developed and use.

Although the need is often stated for distributed cognition models to explain activities in complex socio-technical systems, few exist and they are not easy to define and represent. Our approach here has been to describe, explain, and understand the work of key rail functions separately but with a common methodology, and subsequently to bring the individual models together to jointly describe the collaborative thinking and decision-making as well as communication links involved. The research to do this has built upon a rich seam of rail human factors research of the past few years. The models and knowledge gained will be vital, often central, to engineering of the successful railway of the future.

ACKNOWLEDGEMENTS

This work was made possible only due to the efforts of our colleagues and SMEs within a number of rail companies, and the authors thank all of them for their time, expertise and, above all, patience. The authors acknowledge the contribution to the RRUK projects as a whole from colleagues Beverley Norris and Sarah Jackson. They are grateful to the two anonymous referees for their insightful recommendations on how to improve the paper, although any problems remaining in doing so are, of course, our responsibility. The work was informed by our membership of EURNEX (The European Rail Network of Excellence) and was funded by the EPSRC RRUK, with matched funding from Network Rail. However, the views expressed are the authors’ own and do not necessarily represent the views of any of these organizations.

REFERENCES

6 MacCarthy, B. L. and Wilson, J. R. (Eds) Human factors in scheduling and planning, 2001 (Taylor and Francis, London).


32 Branton, R. Investigations into the skills of train driving. Ergonomics, 1979, 22, 155–164.


36 Wright, L., Dabekaussen, M., and van der Schaaf, T. Predicting the causes of SPAD incidents. In People and rail systems: human factors at the heart of the railway (Eds J. Wilson, A. Mills, T. Clarke, and


61 Van Wesel, W. Task oriented support for train shunting planning. In People and rail systems: human factors at the heart of the railway (Eds J. Wilson, A. Mills, T. Clarke,


105 Cox, G., Farrington-Darby, T., and Bye, R. From the horse’s mouth: the contribution of subject matter experts (SMEs) to rail research. In *People and rail systems: human factors at the heart of the railway* (Eds J. Wilson, A. Mills, T. Clarke, and B. Norris), 2007 (Ashgate Publishing, Abingdon, UK), in press.

106 Vicente, K. J. *Cognitive work analysis*. Toward safe, productive, and healthy computer-based work, 1999 (Lawrence Erlbaum Associates, Mahwah, New Jersey).


