Early Route Drivability Assessment In Support Of Railway Investment

W Ian Hamilton, Technical Director, Human Engineering Limited[1]
Emma Lowe, Human Factors Advisor, Network Rail[2]
Charlotte Hill, Senior Human Factors Consultant, Human Engineering Limited[1]

1. Human Engineering Limited, Shore House, 68 Westbury Hill, Bristol, BS9 3AA
2. Network Rail, 40 Melton Street, London, NW1 2EE

KEYWORDS: Signalling scheme, workload prediction, GRIP process, drivability.

Abstract

Considerable investment is being made in new rail schemes in the UK. A major area of investment is the design of new and revised infrastructure on the railway network to improve capacity, safety and performance. In making these investments and changes it is important that the anticipated benefits are likely to be achieved in practice. One of the key factors that can influence the success of route change is the drivability of the new scheme design. At present, drivability can only be assessed relatively late in a design lifecycle when it is difficult and most expensive to make any modification.

Drivability is influenced by the design of a signalling scheme, owing to the fact that the infrastructure features and characteristics of a route determine the driver's task. For example, a driver will read a speed sign, and respond by accelerating after the train length has passed the sign. A driver may then pass a radio change marker board, read it, and respond by checking that the radio has changed to the correct channel. Consequently drivability is closely related to the human factors concept of workload.

Network Rail's Ergonomics Group commissioned Human Engineering Limited to develop a tool to help assess the drivability of schemes at an early point in their development. The emerging Route Drivability Tool (RDT), and how it can be integrated within the GRIP process, is described in this paper.

Railway Investment Process

Network Rail is responsible for the operation, maintenance, renewal and enhancement of the national rail network in order to meet its statutory and regulatory requirements and the aspirations of its stakeholders. To fulfil this goal, Network Rail must make intelligent investment and engineering decisions that deliver schemes that can be operated safely, reliably and efficiently.

The Guide to Railway Investment Projects (GRIP) (Reference 1) describes how Network Rail manages and controls projects, throughout their lifecycle to ensure success and compatibility with existing rail operations. All qualifying renewals, enhancements and capital expenditure projects must comply with the GRIP process.

The GRIP process is structured with an overarching policy manual that is applicable to all projects. Similarly there is an underpinning project management manual that is applied to all projects. The process is also supported by:

- An Enhancements and Developments Manual,
- A Renewals Development Manual,
- The Delivery Manual,
- The Third Party Manual, and
- The Investment Regulations, with guidance on obtaining funding authority.

All projects must be able to demonstrate their compliance with these processes. This is achieved through a seven stage project lifecycle. Each stage is required to deliver an agreed set of products. The lifecycle is illustrated in Figure 1.

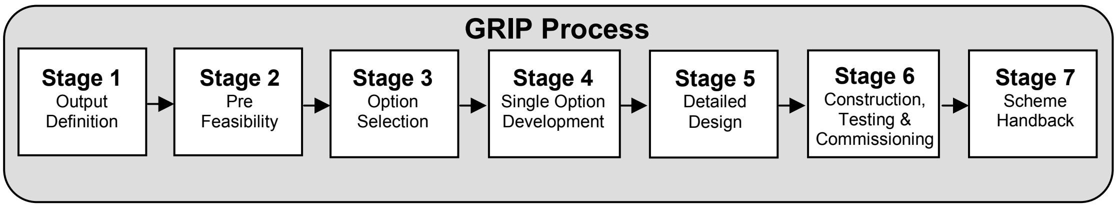


Figure 1: Lifecycle

Each of the stages represents a discrete business or technical milestone and GRIP defines the products that should be produced at that stage. For each product GRIP defines its purpose, input dependencies, format and content, and there is a responsibilities matrix that provides specific guidance on the preparation of products and cross-references to applicable standards and templates. Each stage is also subject to a gate review to confirm that all management, business and technical milestones are achieved.

Human Factors Implications For Investment

New railway infrastructure projects must be able to demonstrate that anticipated operational revenues will justify the capital investment for their development. In the case of projects such as Thameslink 2000 and Cross Rail, operational revenues will depend on running frequent, high occupancy services, stopping at numerous stations, with short dwell times. These requirements introduce route design and signalling scheme complications that place greater workload demand on drivers and signallers. Consequently, human factors can be a key determinate in the ultimate success of the scheme.

Track upgrades and new signalling scheme projects are under pressure to deliver shorter, more reliable journey times, more frequent services, higher track capacity, greater levels of safety and reliability, and increased value for money. This can result in scheme designs that feature higher operating speeds combined with more frequent speed changes. Stock specific signage is also a feature of the proliferation of lineside signs. All of which demands vigilance and discrimination by the driver, placing a greater reliance on route knowledge. There can also be increased complexity from more variable routing. This means more lineside signage and signalling, plus a host of other junctions and associated indicators. New cab displays and communications systems also contribute to the information processing demands on drivers. And reduced station dwell times and journey turn-around times can add time pressure to normal duties.

Understanding these demands and the drivers' capacity limitations for vigilance, attention, information processing and memory are key inputs into the design activity. There is considerable evidence that high workload can reduce performance effectiveness and increase the propensity for error, prevent effective recovery from mistakes, and impose occupational stress that can affect health.

Network Rail acknowledges the importance of human factors for investment projects and has produced a specific standard (Reference 2, RT/E/P/24020) and guidance (Reference 3, RT/E/G/00027) on incorporating ergonomics (human factors) within engineering design projects. These deal with the requirements to consider the human interface, and the phases of human factors work are fully aligned with the GRIP process – see Figure 2. In particular, the guidance document provides details of the expected ergonomics technical milestones and outputs that are relevant to the stages of the GRIP lifecycle.

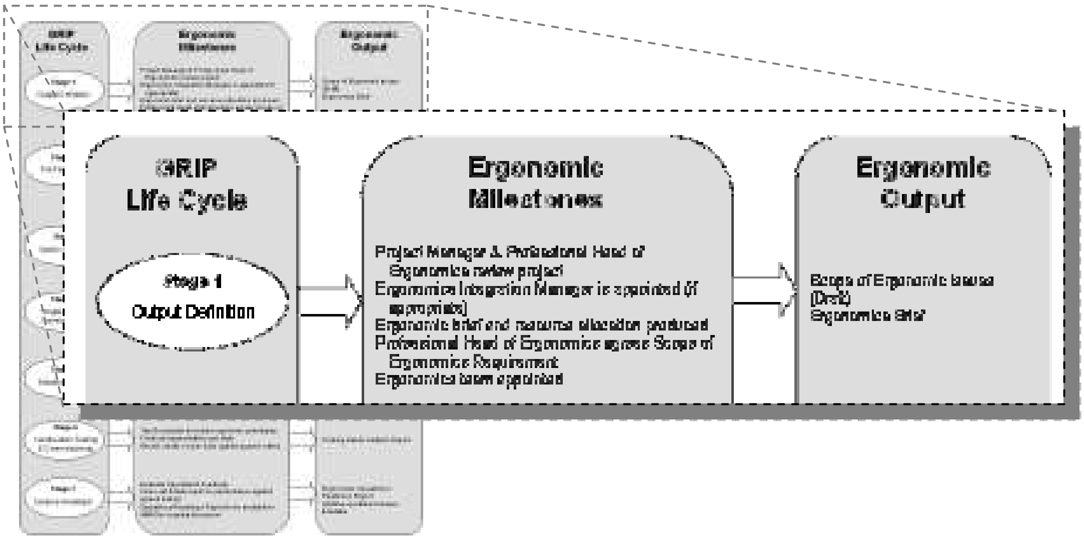


Figure 2: Engineering lifecycle with ergonomics milestones and output

One such output (defined in Note 11 of Reference 3) is the workload assessment report. Workload assessments are undertaken to establish the feasibility of the demands imposed on the operator by the proposed design of the user interface. The route drivability technique (RDT) provides a means of assessing the workload imposed on drivers by the design of the infrastructure and its proposed operating conditions. The next section introduces the RDT and explains how the concept of drivability has been operationalised.

Route Drivability = Workload

Drivability is influenced by the design of a signalling scheme, because the driver's task is determined by the infrastructure features and characteristics of a route. For example, a driver will read a speed sign, and respond by accelerating after the train length has passed the sign. A driver may then pass a radio change marker board, read it, and respond by checking that the radio has changed to the correct channel.

As such, the type of infrastructure features and characteristics of a route will govern the tasks that a driver performs. The location of these features and the linespeed of the route determine when these tasks occur and the time available to perform them.

Previous studies (References 4 and 5) showed how increasing the number of infrastructure features on a route increases the number of tasks to be performed by the driver, and increasing the linespeed reduces the amount of time available to perform these tasks. This, in turn, increases driver workload, which may reduce drivability, all other things being equal.

Thus, decisions taken about the design of a route will directly affect the workload of the driver, and analysis of workload as a dependent variable of route design can be diagnostic of the drivability of the route when it is built.

At present drivability can only be assessed late in the design lifecycle, and tends to be evaluated as part of other processes such as risk assessment and signal sighting where there is a reliance on structured expert judgement. Consequently, the arguments for design modification on the basis of drivability may be difficult to qualify and are impracticable at mature stages of delivery, and so other engineering and budgetary factors may take priority. The RDT provides a means of making this evaluation early and on the basis of objective, quantified criteria.

Behavioural Building Blocks

The scientific basis to the RDT is described in detail in Reference 6, which describes the development and main features of a general model of human information processing capabilities applied to railway engineering. The model utilises cognitive task analysis and modelling techniques to describe driver performance in relation to infrastructure design features and operational conditions. The model is capable of predicting the performance time, workload and error consequences of different operational conditions.

Features on the route prompt driver activities. A feature will prompt a group of tasks – the "behavioural building block" – a meaningful group of goal-directed tasks. The building block is "anchored" to the feature. Features were identified and selected for modelling. The construction of corresponding behavioural "building blocks" that describe how the driver responds to each route feature, and how they are "anchored" or related to the corresponding features, was undertaken as a scoping study (Reference 7).

The scoping study presented a hierarchical task analysis that described the core driving activities identified in the relevant Railway Group Standard (GO/RT3251, Reference 8). The task analysis covered revenue service operation, including the movement of passengers and freight.

As there are a very large number of different features present on the UK railway, and the driving task covers many different circumstances, it was decided to capture the most important aspects of the task to an appropriate level of detail. Features were prioritised according to the following criteria:

- How much the feature, and its associated task, affects the drivability of the route, as determined by previous work and consultation with Network Rail.
- How much is known from theory or observation about how it affects the driver.
- Whether it can be modelled, either quantitatively or qualitatively.

An example of an anchored building block is presented in Figure 3. The anchor is a critical event that prompts the behaviour, like the sighting of a signal or the sounding of an alarm. The anchor may occur at the start or end, but most often occurs partway through the task. Anticipation and expectation play an important role in expert driver performance, and these parts of the task, called the "early" component, will be performed in advance of the prompt. The parts of the task that occur after the anchor event are called the "late" component. The task is represented on a time axis, the length of the task indicating the performance time.

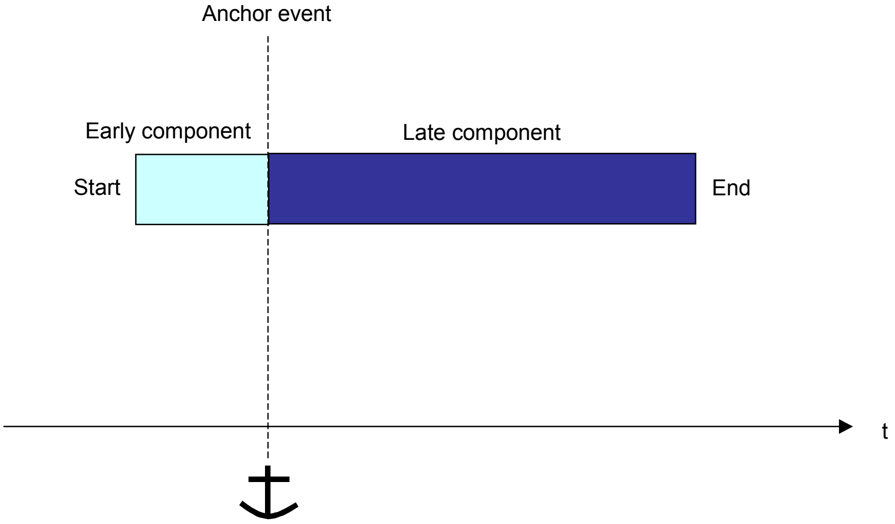


Figure 3 - A task-group anchored to an event

GOMS framework

The drivability methodology is based on the Goals-Operators-Methods-Selection Rules model of skilled human performance (Reference 9, Card, Moran and Newell, 1983). It assumes that drivers perform the driving task in a goal-directed manner, and use rules to select the most appropriate behaviour to complete the task. The hierarchical task analysis of train driving explicitly describes the goals the driver aims to meet, and the methods they have available to them.

Simple rules can be constructed to model driver behaviour in relation to task events (for example speed changes, maintaining progress, response to particular features). These have been shown in the validation study (see below and Reference 5) to be closely related to real driver behaviour, especially when that behaviour is clearly bounded by task goals and environmental constraints (for example, braking for a new speed restriction or stopping at a station to meet the service pattern).

Perceive-Recognise-Decide-Act Cycle

The perceive-recognise-decide-act cycle (PRDA) provides a general strategy for driver responses to task events. The PRDA cycle divides behaviour into three stages of action:

- A perceptual stage - when the feature is detected and recognised
- A processing stage - when the response to the feature is selected
- A response stage - when the response is made. The response may be a physical act, but may also be a cognitive decision that can influence future behaviour

The model assumes that drivers are prospective, and have a plan of how they intend to drive the route. It is based on the experience and knowledge of how they usually drive the route, and is constantly updated to allow modifications in strategy as required. The driving plan allows the driver to predict future events and prepare for them in advance, indeed certain activities require this for efficient performance.

The driver continually checks that information and driving events are consistent with the driving plan. New information or unexpected events are incorporated into the plan through the cognitive re-planning task. To reflect this, three additional stages are also included:

- An anticipation stage – when the driver prepares for the activity
- A search stage – the driver actively searches for the event in order to release the behaviour
- A checking stage – following the event the driver checks that his action is appropriate and that any information collected is consistent with (or added to) his driving plan.

Therefore the general strategy used in the behavioural building blocks can be summarised by the following tasks:

- Anticipate feature
- Search for feature
- Perceive feature
- Recognise feature
- Respond to feature
- Check/update plan

Where evidence existed (from studies or driver interview) that further behaviours would be required for the driver to perform a particular activity these were added, but the guiding principle was that the behavioural description would be the simplest one that could account for realistic performance.

For the purpose of modelling, behavioural building blocks are broken down into action descriptions. These are classified into either visual, auditory, cognitive or psychomotor behaviour. This so called VACP task analysis is a standard method in human factors which is associated with techniques for performance time estimation and workload analysis, (e.g. demand ratings). Further details are provided in Reference 6.

The following infrastructure features were identified as priority features for inclusion in the methodology:

- Signals, banner repeaters and route indicators
- AWS magnets
- General information boards, radio change and whistle boards
- Permissible speed signs and advanced warning boards
- Gradients and neutral sections
- Stations, platforms and stopping cards

As an example Table 1 shows a description of the ‘AWS Horn Response’ behavioural building block.

Task	Action description	VACP	Demand	Performance time range (s)
Anticipate AWS horn	Anticipate AWS horn about to sound	Psychomotor - Discrete Actuation	1	0.076 to 0.420
		Cognitive - Encoding/Decoding, Recall	2	0.167 to 0.527
AWS Horn Anchor	-	-	-	-
Hear AWS horn	Hears and recognises the horn	Auditory - Detect/Register Sound	1	0.075 to 0.370
		Cognitive - Sign/Signal Recognition	1	0.100 to 0.540
Check AWS alert	Compare AWS alert with expectations	Cognitive - Evaluation/Judgement (Single Aspect)	2	0.167 to 0.527
Cancel AWS horn	Return hand to desk	Psychomotor - Discrete Actuation	1	0.076 to 0.420

Table 1 – VACP Description of the "AWS Horn Response" Behavioural Building Block

The development of the behavioural building blocks for all the priority features is described in full in Reference 10.

Summary Results of RDT Validation Study

The measures of drivability are the workload variables of time pressure, task demand level and cognitive conflict between concurrent tasks. The theoretical basis of these measures is described in Reference 6. In a previous study (Reference 5) the RDT's performance time and workload prediction features have been validated for use in drivability assessment. A cognitive task analysis was used to model driver behaviour in response to route features encountered when driving through the core section of the London Thameslink Railway. Performance time and workload predictions were then derived using the various performance time and workload prediction algorithms in the cognitive model. The details of these are also reported in Reference 6.

The modelled scenario was then driven in both a train cab and route simulator (of the same route) with N=17 driver subjects. Observational data were obtained for driver task performance, a secondary task measure of workload (i.e. a self-paced mental arithmetic test), and for subjective ratings of workload (using NASA TLX and Instantaneous Self Assessment). These data were then used to compare the drivers' behaviour with the predicted driving strategy. Task performance times were compared with predicted execution times, and predicted workload values were compared with secondary task and subjective data.

The behavioural rules in the cognitive task analysis had predicted, with a high degree of accuracy, how all 17 drivers responded to events. For example, it was possible to predict with precision, the speed profile driven by drivers over the entire route. This is illustrated in Figure 4, showing the close match between the predicted driving speed profile for the route (the black line) and that driven by the drivers in the simulator (the coloured lines). The thick, pale green line shows the speed profile for a real train on the same route. This shows that the modelled driving behaviour, simulated driving, and real world driving speeds all coincide.

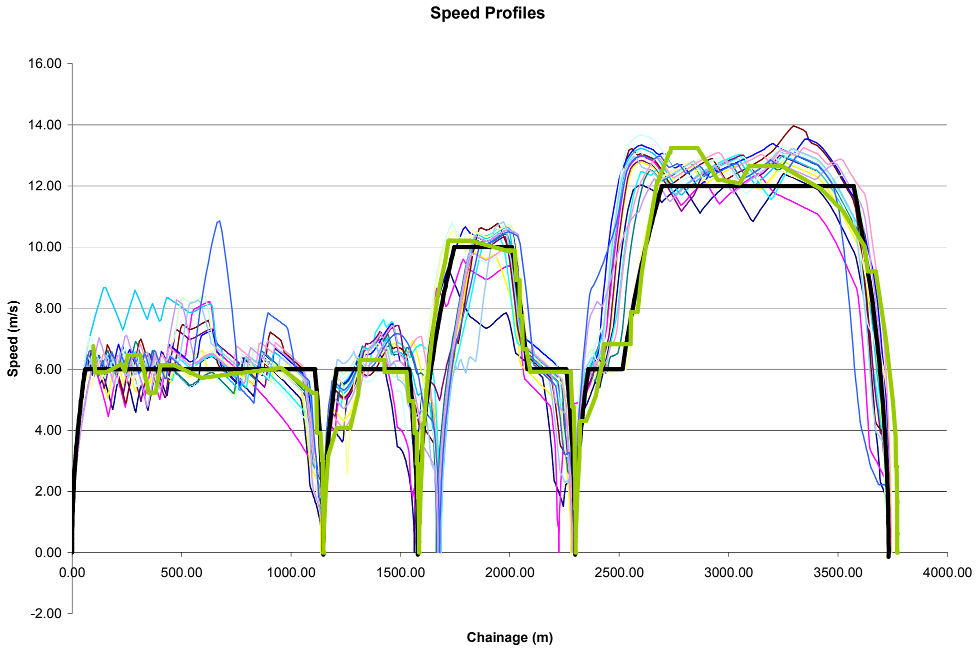
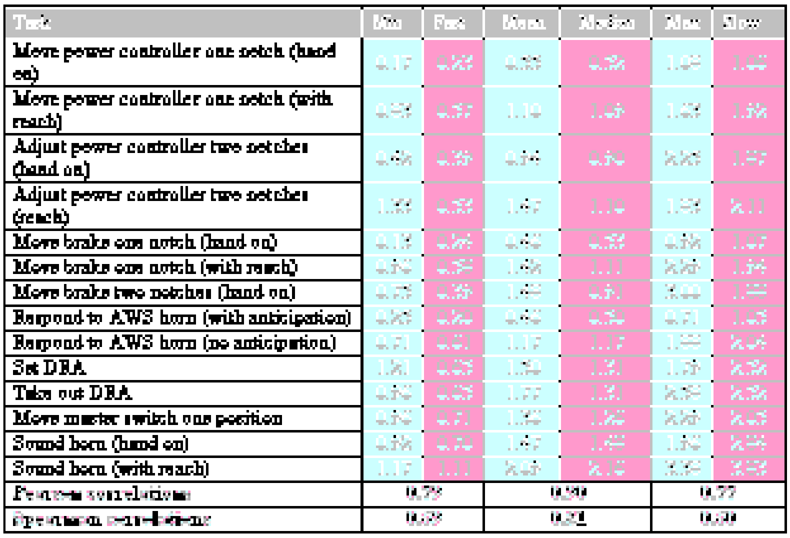


Figure 4: Speed profiles

The kind of behaviour that proved most difficult to predict accurately was closed-loop control behaviour, such as maintaining a constant train speed. Although the goal of maintaining speed was clearly revealed by the cognitive task analysis, there was a high degree of individual difference in the traction power adjustment strategy employed by drivers. However, the strength of the match to the overall speed profile driven on the route as a whole (Figure 4) suggests that this difficulty does not undermine the general utility of the method.



Task	Min	Fast	Mean	Median	Max	Slow
Move power controller one notch (hand on)	0.17	0.23	0.35	0.36	1.06	1.05
Move power controller one notch (with reach)	0.53	0.37	1.10	1.09	1.63	1.96
Adjust power controller two notches (hand on)	0.46	0.39	0.74	0.70	2.23	1.57
Adjust power controller two notches (reach)	1.33	0.37	1.47	1.10	1.53	2.11
Move brake one notch (hand on)	0.13	0.26	0.46	0.52	0.96	1.07
Move brake one notch (with reach)	0.36	0.58	1.42	1.11	2.20	1.94
Move brake two notches (hand on)	0.75	0.39	1.48	0.91	3.00	1.55
Respond to AWS horn (with anticipation)	0.25	0.20	0.46	0.30	0.71	1.05
Respond to AWS horn (no anticipation)	0.71	0.61	1.17	1.17	1.55	2.06
Set DRA	1.21	0.63	1.50	1.31	1.79	2.32
Take out DRA	0.36	0.63	1.77	1.31	2.56	2.32
Move master switch one position	0.36	0.71	1.36	1.26	2.20	2.05
Sound horn (hand on)	0.96	0.70	1.47	1.46	1.36	2.54
Sound horn (with reach)	1.17	1.11	2.09	2.16	3.56	3.53
Pearson correlations	0.79		0.89		0.77	
Spearman correlations	0.75		0.81		0.80	

Table 2: Comparing predicted (Fast, Slow & Median) and observed (Min, Max & Mean) timings.

The behavioural performance time algorithms provided reasonable estimates of actual performance for most behaviours. There was a very close match between predicted typical times (Median) for most overt driving behaviour and actual observed behaviour – see Table 2.

The ranges too were quite accurately predicted, although it is clear that the algorithms tend to over-estimate the slowest (Slow) performance times. This consistent error can become a significant over-estimate when timings are added together for extended sequences of action. For example, station departure preparation actions were predicted as taking typically 7.9s but were observed to have taken only 6s. Further work on this has already provided some evidence that predictive accuracy for slowest times can be improved by using a set of algorithms corrected using a z-score method, (Reference 11).

The predicted workload values for time pressure, demand and conflict were compared with the subjective and secondary task measures from the simulator study. The RDT produced many more workload peaks than were observed in the simulator data.

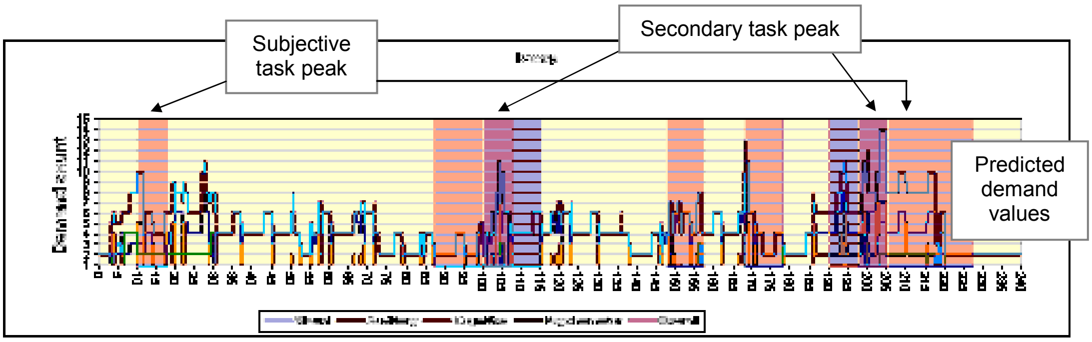


Figure 5: Workload profiles for demand values

However, for every peak in the subjective and secondary task data there was a corresponding peak in the predictions. On Figure 5, the mid grey areas show the peaks for subjective workload ratings, while the dark grey areas show peaks in the secondary task data. Both the predicted measures and the observed measures show peaks at arrival into stations and during speed control under high vigilance demand periods.

It is suggested that the additional peaks in predicted data reveal the greater sensitivity of the predictive methods to short duration fluctuations in workload. This is probably due to the higher sampling rate of the predictive method compared to the intermittent and post-hoc methods of data collection relied upon for the subjective reporting and secondary task methods.

Overall the results of the validation study have provided sufficient evidence to justify the continued development of the RDT.

RDT In Support of GRIP

The RDT now exists as a software supported method for use in predicting the drivability of a route (see Figure 6). The RDT is implemented as a Microsoft Access database application that allows human factors specialists and signal scheme designers to select the features of a planned route and quickly assess the feasibility of the driving task.

The basic RDT procedure involves:

- Entering infrastructure features and their distance along a journey, based on a signalling scheme diagram or other source of information
- Running the model to calculate the speed profile and workload rating
- Interpretation of the workload ratings
- Discussion and iteration of the journey, incorporating human factors advice to optimise the driver's workload.

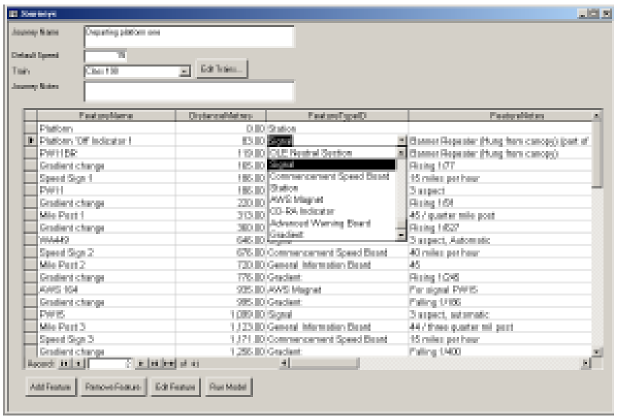


Figure 6: The basic RDT screen

This assessment allows human factors and workload considerations to be made in a timely and efficient manner. The outputs from the technique provide a way of identifying problem areas in drivability terms, and allow for rapid iteration to test the effectiveness of proposed solutions. The RDT has already been used in support of a renewal project (Reference 10) and has demonstrated that it is sensitive to differences in the existing and new route designs under consideration, and identified the increased demands on drivers associated with the revised route.

The predictive analysis of route drivability offers a number of advantages to Network Rail:

- To improve the safety, operational efficiency and reliability of routes
- To incorporate human factors early in the project lifecycle, facilitating human factors integration throughout the project cycle
- An objective, quantifiable and diagnostic methodology to support structured expert judgement (SEJ) in decision making.
- The RDT supports the GRIP process at stages 3 to 5, as illustrated in Figure 7 below.

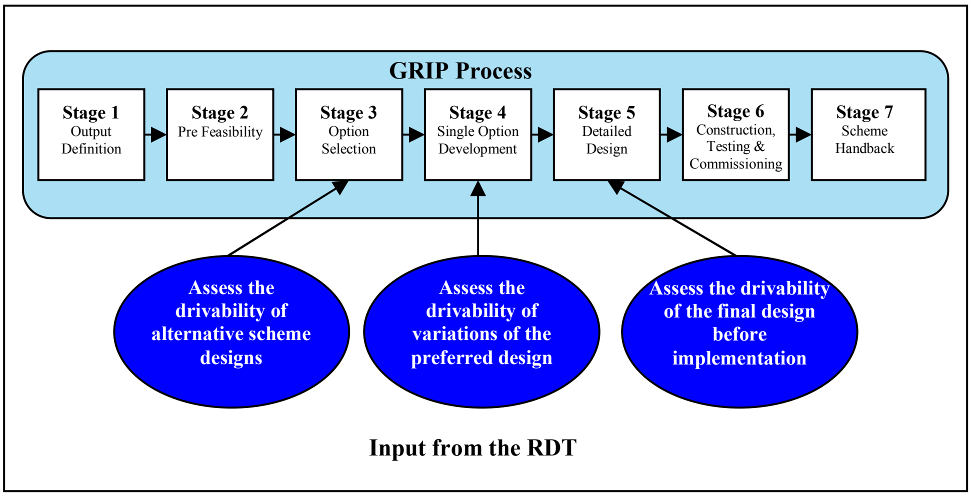


Figure 7: Inputs from the RDT to the GRIP stages.

The RDT methodology is also complementary to the design process. For example, new schemes are increasingly being modelled in CAD packages (like SolvSighting) to estimate sighting distances for particular features and establish that they satisfy the required standards. The RDT uses sighting distances to anchor when the driver begins particular tasks, so the output from CAD methods would improve the accuracy of a RDT model of a route.

Ongoing Work

Work on the development of the RDT is continuing with the list of route features, and their associated building blocks, being extended to include such things as signal aspects and route indicators. The extension also includes further development of the workload outputs. In particular, a cognitive conflict parameter is to be added that will provide a clear indication of where excessive workload occurs. This will also identify precisely the route design features that have caused workload problems so that these can be targeted for corrective action. There is also work to develop an engineering procedure to support the application of RDT within projects and to specify competency requirements for its use. This work will be concluded in 2006.

References

1. **Network Rail**, *The Guide to Railway Investment Projects (GRIP),* 2005.

2. **Network Rail Company Specification**, 2003. *Incorporating Ergonomics Within Engineering Design Projects: Requirements,* RT/E/P/24020, Issue 1, August 2003.

3. **Network Rail Company Specification**, 2003. *Incorporating Ergonomics Within Engineering Design Projects: Guidance Note,* RT/E/G/00027, Issue 1, August 2003.

4. **Li, G., Hamilton, W. I. and Clarke, T**., (2003) *Driver recognition of railway signs at different speeds – A preliminary study.* In: Contemporary Ergonomics 2003, (ed. P.T. McCabe), London: Taylor & Francis, 373-378

5. **Hamilton, Lowe & Blanchard**, (2004). *Validation of ATLAS Workload Prediction Tool for Early Human Factors Analysis*, HEL/NR/03860/RT2, Issue 01, 1st April 2004.

6. **Hamilton, W.I. & Clarke, T.** (2003) *Driver Performance Modelling and its Practical Application to Railway Safety.* In Human Factors in Rail Conference, October 2003, York.

7. **Hamilton, Lowe, Blanchard & Maddock**, (2004), *A Scoping Study for Route Derivability Assessment*, HEL/NR/041022/RT1, Issue 01, 11th August 2004.

8. **Railway Group Standard**, *Train Driving* GO/RT3251, Issue 4, October 2004.

9. **Card S.K., Moran T.P., & Newell A.,** (1983), *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, Inc – New Jersey.

10. **Hamilton, Lowe, Blanchard, Hill & Maddock**, (2004) *Early Route Drivability Assessment*, HEL/NR/041024/RT1, Issue 01, 3rd December 2004.

11. **Hamilton & Newman** (2003), *Cognitive Task Analysis for Signal Sighting Hazards*, HEL/RSSB/03862/RT1, Issue 01, 8th October 2003.