FINAL PUBLISHABLE REPORT

CONTRACT N°: GRD1/2000/25361 S12.319626

ACRONYM: HASTE

TITLE: Human Machine Interaction and the Safety of Traffic in Europe

PROJECT CO-ORDINATOR : Prof. Oliver Carsten, Institute for Transport Studies, University of Leeds

PARTNERS:

University of Leeds (UNIV-LEEDS)	UK
Human Factors Research Institute (TNO)	NL
Swedish National Road and Transport Research Institute (VTI)	S
Delft University of Technology (Delft)	NL
Volvo Technology Corporation, Sweden (Volvo)	S
MIRA Ltd. (MIRA)	UK
Technical Research Centre of Finland (VTT)	FIN
Universidade do Minho (UNM)	P
Transport Canada (TC)	CA

REPORTING PERIOD: FROM 1st January 2002 TO 31st March 2005

PROJECT START DATE: 1st January 2002 DURATION: 39 months

Date of issue of this report: May 2005 (Version 1)
August 2005 (Version 2)
October 2005 (version 3)
November 2005 (Version 3.1)

Project funded by the European Community under the 'Competitive and Sustainable Growth' Programme (1998-2002)

INSTITUTE FOR TRANSPORT STUDIES
DOCUMENT CONTROL INFORMATION

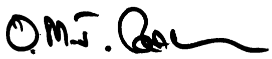
Title	HASTE Final Report
Author(s)	O.M.J. Carsten, N. Merat, W.H. Janssen, E. Johansson, M. Fowkes, K.A. Brookhuis
Editors	Natasha Merat, Oliver Carsten
Reference Number	HASTE Final Report
Version Number	3.1
Date	November 2005
Distribution	Project and European Commission
Availability	Restricted, until approved
File	H:\Safetygroup\HASTE\Final Publishable Report_v3.1.doc
Authorised	O.M.J. Carsten
Signature	

Table of contents

Executive Publishable Summary

The partnership in HASTE consisted of eight European partners and one partner from a country with a cooperation agreement. Expertise in the area of driver behaviour and system evaluation was guaranteed by the inclusion of academic and research institutions, each of which brought a particular speciality to the project. The two industrial partners, one of whom is associated with a vehicle manufacturer, ensured that the project goals were realistic and timely and provided important input with regards to design and manufacturing.

At the start of the HASTE project, the international state of the art in terms of methodologies for assessing the safety implications of HMI was highly problematic. There were various pieces of advice on issues to consider in product development, with the most notable being the original version of the European Statement of Principles, and there were various proposed tests and quasi-standards, which generally lacked proper validation and which also typically were not designed to evaluate usage and performance *while driving.*

Thus the available advice, tools and metrics did not permit, in any straightforward way, judgements to be made about the safe use of a particular IVIS during driving. There were, as a result, no criteria which could be used by a manufacturer, a system supplier, consumer organisations or the public authorities to determine whether a particular system meets a minimum threshold of safety in actual use, or for that matter, to rate or rank different products in terms of their safety while in use. Also little was known about the effect of cognitively demanding *non-visual* IVIS on driving performance.

With the huge growth in the market for satellite navigation systems and other "nomadic" devices, the need for a suitable safety-related test procedure has become even more urgent. Tens of thousands of devices are in daily use without any assurance that they can be operated safely in a vehicle, Furthermore, the purchaser of such a device is not able to obtain an information about the safety merits of one system as opposed to another and is thus left ignorant about this crucial aspect of product quality.

There was thus a clear need for a new test regime for the assessment of IVIS which:

- Was technology-independent
- Had safety-related criteria
- Was cost effective
- Was appropriate for any system design
- Was validated through real-world testing

Therefore, the main *aim* of HASTE was to develop methodologies and guidelines for the assessment of in-vehicle information systems (IVIS). The intention was to devise an assessment regime that was independent of the design of an IVIS, and that was based on an evaluation of driving performance while using the system, as compared with driving performance when not using the system. A major technical and scientific *objective* of

HASTE was the identification and exploration of the relationship between traffic scenario, driver and IVIS. This relationship was investigated by studying behavioural, vehicle, psycho-physiological, and self-report measures.

The technical and scientific objectives of the programme of research were:

- To identify and explore relationships between traffic scenarios in which safety problems with an IVIS are more likely to occur
- To explore the relationships between task load and risk in the context of those scenarios
- To understand the mechanisms through which elevated risk may occur in terms of distraction and reduced Situation Awareness
- To identify the best indicators of risk (accident surrogates)
- To apply the methods devised to the evaluation of real systems
- To recommend a pre-deployment test regime that is both cost effective and possesses the validity to predict performance
- To recommend an approach for the preliminary hazard analysis of an IVIS concept or design

The first phase of the project defined the methods, metrics and scenarios in which IVIS-related safety problems are likely to occur. In particular, Deliverable 1 attempted to refine the knowledge on the impact of IVIS on driving performance. An IVIS may have negative or positive consequences on driving. However, since IVIS-related performance decrements are crucial for the final safety judgement, in HASTE the emphasis was on the *negative* effects of IVIS on driving performance.

The methods, metrics and scenarios identified in Deliverable 1 were applied to the first experimental stage of the project, which attempted to establish the effect of different types of distraction imposed by an IVIS on the driving task, as well as identifying the best risk indicators for assessing IVIS use during driving. Distraction was either from a visual or cognitive ‘surrogate’ IVIS task, designed specifically to control the level of distraction in a systematic manner. The visual task, not unexpectedly, led to poor steering behaviour and degradation of lateral control of the vehicle. By contrast, with the cognitive task, the major negative effect was more on longitudinal control, particularly in car following, rather than on lateral control. Indeed, the cognitive task appeared to ‘improve’ lateral control of the car, although eye-movement patterns showed that this was coupled with a greater concentration of gaze towards the centre of the road, at the expense of the periphery.

The risk indicators identified in the first experimental phase were then used to work towards the development of a **testing regime** for IVIS that was both simple and valid. This simplified test regime was applied in a second set of experiments, for the evaluation of a series of tasks performed on some *real* IVISs while driving. Conclusions from these studies influenced the final phase of the project: the formulation of guidelines for a future test regime for IVIS, a kind of ‘cook book’, which would provide a practical testing and scoring procedure for the safety assessment of IVIS, when used during a drive. Specifically, between four and six behavioural parameters are thought sufficient to evaluate any system that is offered for assessment. The most informative tool for assessment of a system is thought to be a reasonably advanced driving simulator, used in a rural road setting, and requiring an average of 15 participants.

It is hoped that the prototype test regime recommended in HASTE will be developed further, in order to achieve an efficient and fully validated procedure that can be applied to all types of IVIS equipment including nomadic devices and handsfree mobile phones. The intention in HASTE has been to recommend a test regime that is as cost effective as possible, and can be used both as a *pre-deployment* regime and for *final verification* of IVIS tasks. A suitable test regime must be practical, meaningful, and repeatable during product development, in order to decide which in-vehicle tasks a driver might reasonably be allowed to access and perform whilst driving. The HASTE test regime could be presented in a variety of ways, including code-of-practice, ISO standard, Pass/Fail criteria or as a testing procedure within the primary new car assessment programme (PNCAP), i.e. in order to provide information to consumers Finally, it is hoped this regime will be used both by governmental organisations and by OEMs.

An additional item of work carried out by HASTE has focussed on how a preliminary hazard analysis of an IVIS concept or design could be carried out. This would assist designers very early in the design process and would, it is hoped, prevent more hazardous designs being developed beyond the conceptual stage. The work began with a review of the state of the art in the preliminary safety analysis of automotive systems in general to identify those that had the most potential for application to an IVIS HMI. A new evaluation methodology was then defined for the hazard identification and risk analysis of an IVIS concept or design. The effectiveness of this selected methodology was then validated by applying it to the HMI of existing IVIS system designs and the risk assessment results compared with the evaluation results obtained elsewhere in the project.. The results from this comparison process then guided the definition of a final IVIS HMI assessment methodology called the Driver Operability Procedure (DOP). Guidance has been given as to how this procedure could be applied within an industrial design and development process.

Disclaimer

The research reported herein was conducted under the European Commission Competitive and Sustainable Growth Programme. The project was carried out by a consortium comprising: the Institute for Transport Studies, University of Leeds; TNO Human Factors Research Institute; the Swedish National Road and Transport Research Institute (VTI); Delft University of Technology; Volvo Technology Corporation; MIRA Ltd; Technical Research Centre of Finland (VTT); Universidade do Minho; and Transport Canada. The opinions, findings and conclusions expressed in this report are those of the authors alone and do not necessarily reflect those of the EC or of any organisation involved in the project.

1 Objectives of the project and state of the art

1.1 Project objectives

Driver distraction is a major cause of accidents and the introduction of new mobile devices and in-vehicle information systems into road vehicles has the potential to exacerbate the problem of distraction by creating new sources of interference with the driving task. Hence the concern among safety experts and policy-makers over the impact of mobile phone use on traffic safety. Arising from this concern and focussed particularly on concerns about IVIS such as satellite navigation systems, there have been a number of initiatives in Europe, North America and Japan. The rationale for the HASTE project was that there was a need for a fundamental investigation of the link between distraction and driving performance with a view to the creation of a performance-based evaluation regime for assessing the safety of different HMIs.

The technical and scientific objectives of the programme of research were:

- To identify and explore relationships between traffic scenarios in which safety problems with an IVIS are more likely to occur
- To explore the relationships between task load and risk in the context of those scenarios
- To understand the mechanisms through which elevated risk may occur in terms of distraction and reduced Situation Awareness
- To identify the best indicators of risk (accident surrogates)
- To apply the methods devised to evaluating real systems
- To recommend a pre-deployment test regime that is both cost effective and possesses the validity to predict performance
- To recommend an approach for the preliminary hazard analysis of an IVIS concept or design.

1.2 Pre-existing state of the art

At the start of the HASTE project, the international state of the art in terms of methodologies for assessing the safety implications of HMI was highly problematic. There were various pieces of advice on issues to consider in product development, with the most notable being the original version of the European Statement of Principles, and there were various proposed tests and quasi-standards, which generally lacked proper validation and which also typically were not designed to evaluate usage and performance *while driving.*

Thus the available advice, tools and metrics did not permit, in any straightforward way, judgements to be made about the safe use of a particular IVIS during driving. There were, as a result, no criteria which could be used by a manufacturer, a system supplier, consumer organisations or the public authorities to determine whether a particular system meets a minimum threshold of safety in actual use, or for that matter, to rate or rank different products in terms of their safety while in use.

With the huge growth in the market for satellite navigation systems and other "nomadic" devices, the need for a suitable safety-related test procedure has become even more urgent. Tens of thousands of devices are in daily use without any assurance that they can be operated safely in a vehicle, Furthermore, the purchaser of such a device is not able to obtain an information about the safety merits of one system as opposed to another and is thus left ignorant about this crucial aspect of product quality.

There are three types of standard or procedure that can be employed for the evaluation of HMI (Parkes, 1995):

1. ***Product or design standards.*** These take the form of specifying the physical aspects of the system, for example a minimum screen size or a particular layout of the control buttons. These are easy for the designer to follow, but they suffer from the drawback that they are technology-dependent and therefore tend to stifle innovation. They also do not guarantee the usability of the entire system or the safety of driving while using the system.
2. ***Procedural standards.*** These take the form of prescribing a programme of analysis and testing to be used in product development. They generally require an inspection of certification authority to enforce their use; they often require extensive documentation; and they can be laborious to apply. In the case of HMI, they will not guarantee safe performance, but of course they can protect the system manufacturer who can show that rules, regulations and advice were followed in the system design process.
3. ***Performance standards.*** These specify a minimum level of task performance, which must be met while the system is being used. They might specify this level for the primary task of driving, or for the secondary task of interacting with the in-vehicle system. Such standards are technology independent and do not limit innovation. If they incorporate an assessment of performance in the primary task of driving, they can provide an objective assessment of whether a minimum level of safety is met. However, they require research effort for their development and validation, and in actual use they may require testing by a particular test house or with specific equipment.

In the automotive world, performance standards are quite common in the vehicle design area, notably in such areas as braking and crashworthiness. However, there were, at the commencement of the HASTE project, no general performance standards or performance tests that could be applied to assess the safety of interaction with an IVIS while driving. Product standards are surprisingly infrequent in the automotive environment, so that not even pedal placement is specified by regulation.

In the area of HMI, the primary focus in the last fifteen years at a European and also national level has been on the development of ***procedural*** guidelines and pre-standards. A PROMETHEUS MMI Checklist was produced in 1991 and was followed by the DRIVE II HARDIE Design Guidelines (Ross et al., 1995) and Handbook (Ross et al., 1996). Further elaboration and refinement led to the UK Safety Checklist for the Assessment of In-Vehicle Systems (Stevens et al., 1999) with a 12-page form, 5 pages of instructions and 26 pages of supporting information. However, for a particular question such as "Is the IVIS free from reflections and glare under all ambient light conditions?" an extensive set of reviews and tests may be required.

There is little doubt that following the procedures recommended in such checklists can help to produce a better-designed system and can help to identify design errors and problems. But the sheer laboriousness of the procedures recommended is likely to mean that shortcuts will be taken. Perhaps more serious, the procedures are in the main subjective and cannot provide a certainty that a minimum level of performance has been met.

In terms of the laboriousness of such checklists, there has been considerable effort at both national and European levels to reduce them to a set of major principles. The outcomes are the UK Code of Practice (Department of Transport, 1994), the German Code of Practice (Wirtschaftsforum Verkehrstelematik, 1996), the ECMT Statement of Principles of Good Practice (ECMT, 1995) and, more recently, the European Statement of Principles on Human Machine Interface from the HMI Expert Task Force (European Commission DGXIII, 1999). But such codes suffer from the fact that, while they enshrine very worthy principles of good design, they do not provide a regime for *assessing a design*. These deficiencies in the current methods have been pointed out in, amongst others, the ETSC review of Intelligent Transportation Systems and Road Safety (1999).

The principles in such documents as the European Statement of Principles (ESOP) are difficult to dispute and may well be helpful in encouraging good design. But the availability of encouragement and sensible *advice* needs to be distinguished from a *test regime* to assure that a reasonable level of safety has been achieved.

An alternative to using the procedural tools is to use a cut-down performance test. These are often applied "off-line", i.e. away from the driving task in a laboratory environment. Here, the most noteworthy test is the 15-second rule submitted in SAE (Society of Automotive Engineers). The "15-second rule" (Farber, 2000) has been proposed as a minimum level of performance in visual attention to an in-vehicle navigation system. The standard suggests that whilst the vehicle is in motion, the longest total task time that should be allowed is 15 seconds:

> "SAE J2364 seeks to define a performance based compliance procedure aimed at limiting the amount of time drivers are permitted to devote to navigation system tasks in moving vehicles"

A number of issues arise when considering the 15-second rule:

- Are static and dynamic task times correlated (i.e. tasks that take a long time to perform statically will take the same amount of time when performed dynamically)?

- Is dynamic task time correlated with “eyes off road time” (i.e. the longer the task takes to perform, the more time the driver will spend with their eyes on the task instead of the road)?
- Some tasks may pass the 15-second rule when conducted out of the driving environment, but would fail in a dynamic situation.

The idea of a total task time has also been disputed to the fact that it has been found that drivers tend to “chunk” large tasks into smaller sub-tasks of between 1 and 2 seconds glance duration (Zwahlen, 1998; Wierwille, 1988; and Dingus, 1986). By using total time on task as a surrogate measure for safety, it implies that ten glances, each of duration 1.5 seconds are less safe than a single 14 second glance. Finally, the method fails to cover cognitive load as opposed to visual load. There is considerable evidence that cognitive load can lead to distraction and reduced Situation Awareness. Situational awareness can be regarded as consisting of three levels, perception of elements in the current situation, comprehension of the current situation and projection of future status (Endsley, 1995). So it was clear that cognitive distraction needed to be a focus of the HASTE work.

1.3 Approach in the HASTE project

The HASTE project began with the premise that what was required by system producers, automotive OEMs, public authorities and consumers was a set of quantifiable benchmarks of driver performance while using an IVIS. The creation of such benchmarks needed to start with a fundamental investigation of the role of task load in perception and decision-making and the influence of such task load on driver behaviour. Such research was identified as the number one human factors research need in the area of Intelligent Transport Systems by a U.S. IVI (Intelligent Vehicle Initiative) Human Factors workshop in 1997 (Battelle Research Group, 1998).

It was clear that the checklist and statement of principle approach offer guidance in design and can help in diagnosing problems. Such procedures can therefore be seen as complimentary to a performance test regime. Cut-down performance pre-standards such as the 15-second rule may have relevance to some kinds of load, but do not address driving performance, (which is after all the ultimate issue), are not able to measure cognitive load and are irrelevant to many forms of interaction and dialogue (e.g. voice messages and voice commands).

There was thus a clear need for a new test regime for the assessment of IVIS which:

- Was technology-independent
- Had safety-related criteria
- Was cost effective
- Was appropriate for any system design
- Was validated through real-world testing

The development of this test regime needed to begin from first principles, i.e. with an assessment of the influence of both visual and cognitive task load on driving performance and safety. This would show which indicators of driving performance and safety were most diagnostic and allow surplus test conditions and situations to be eliminated. The prototype test regime thereby developed could then be applied to the testing of some real IVIS systems (or tasks on those systems). This second stage would permit further refinement of the test regime. Also to be investigated were:

- Whether results obtained using different groups of subjects drawn from different European countries were consistent
- Whether on-road testing was required in addition to or as an alternative to simulator testing
- Whether simulator quality affected the quality or reliability of results

The HASTE work was thus intended to be both ambitious (particularly in addressing the fundamental aspects of the link between distraction and driving performance) and useful, in that it would produce a practical test regime. That test regime would, it was hoped, become a universal tool for assessing the safety of an IVIS.

1.4 Developments since the start of HASTE

The HASTE project began in January 2002. Since that date there have been a number of developments in parallel with the HASTE project.

In the U.S., the CAMP (Crash Avoidance Metrics Partnership) Driver Workload Metrics Project brings together Ford, General Motors, Nissan and Toyota to develop performance metrics and test procedures for both visual and cognitive aspects of driver workload from telematics systems. The intention is that, in the future, vehicle OEMs will be able to use these workload evaluation procedures to assess what in-vehicle tasks might be accessible to a driver while the vehicle is in motion. The work was supposed to be reported by the end of 2004, but to date no public reports have been issued.

In terms of statement of principles, the European Statement of Principles of 1999 was followed by a more extensive North American Statement of Principles (AAM, 2002) and a Japanese set of guidelines for in-vehicle display systems (JAMA, 2004). The JAMA guidelines include stipulations on display positioning, the amount of detail to be shown in map displays and recommends a static occlusion test for assessing visual demand.[1]

A new draft of the European Statement of Principles has recently been issued (European Commission DG Information Society, 2005). The new version is considerably lengthier than the original version and has considerably more explanation and elaboration. However, this document acknowledges that for

[1]In the occlusion method, users are assessed in their interaction with a system while wearing a set of goggles equipped with shutters. The JAMA Guidelines specify that the shutter pattern shall be 1.5 seconds open and 1.0 second closed. The number of 1.5 second chunks used for viewing a display provides a means of assessing visual demand. In the JAMA procedure, this is done statically, i.e. not while driving. The latest version stipulates that, when using the occlusion method statically, the total shutter open time shall not exceed 7.5 seconds, i.e. five openings.

many principles no pass/fail criteria are provided. A number of overall "design goals" are proposed, of which the first is: "The system supports the driver and does not give rise to potentially hazardous behaviour by the driver or other road users" (page 7). No procedure for assessing compliance with this design goal is provided. It is also stated that: "the principles are not a substitute for regulations and standards and these should always be taken note of and used" (page 4).

A very useful assessment of the utility of the statement of principles as an assessment tool has recently been provided by a project carried out on behalf of Transport Canada (Morton and Angel, 2005). This project asked a group of experts to provide critical reviews of the AAM principles. In addition, the AAM principles were used to guide an expert assessment of four different navigation systems as installed in their vehicles by OEMS. This was intended to indicate whether the application of the principles could produce a reliable evaluation. Compliance, as assessed by the inspectors, is provided in the report, principle by principle. The report concludes:

> In the assessment of the AAM guidelines, principles appeared to be valid but often insufficiently detailed and too vague in the accompanying elaborations. This led to poor reliability of results between inspectors. A focus group of inspectors came to the consensus that the guidelines were valid but elaborations needed further work. On the whole, the AAM guidelines were found to be valid. With revisions, the AAM guidelines may be sufficient to ensure safe operation of telematic systems, but insufficient in current form due to inadequate scientific support, incompleteness, and poor reliability of results.

Other significant conclusions in this report were that the AMA principles lamentably failed to address cognitive distractions and that a number of principles required or would benefit from dynamic (as opposed to purely static) evaluation, i.e. evaluation while a driver was performing the driving task. This latter point is particularly relevant to assessing Principle 2.1, which states that "systems with visual displays should be designed such that the driver can complete the desired task with sequential glances that are brief enough not to adversely affect driving." This principle is equivalent to Design Goal 1 in the new version of the ESOP, and is acknowledged in the Canadian report as the core of the AAM principles.

It should be noted that the AAM document is considerably more elaborate and detailed than the counterpart European document, even in its revised version. It is fair, therefore, to conclude that there is little chance that the ESOP will ensure that only safe systems come to market. The new version of the European Statement of Principles can provide advice, but is not intended to serve as a test procedure. And there is serious reason to doubt that any such statement of principles can ever serve as a robust methodology for a test regime. Indeed, it can be argued that there needs to be a clear distinction between ***advice*** on how to improve design, including what considerations to embody in design, and a ***test regime*** for assessing IVIS. The goal of HASTE has been to produce the prototype of such a test regime, and that goal is as valid now as it was when the project was launched.

2 Scientific and technical description of the results

The workplan for this project was structured into Workpackages. Each of the research Workpackages (WP1, 2, 3 and 4) approached separate parts of the research question and technical objectives described above. The results of each of these Workpackages had implications for the direction of the subsequent ones. The relationship between all Workpackages is outlined in Figure 1.

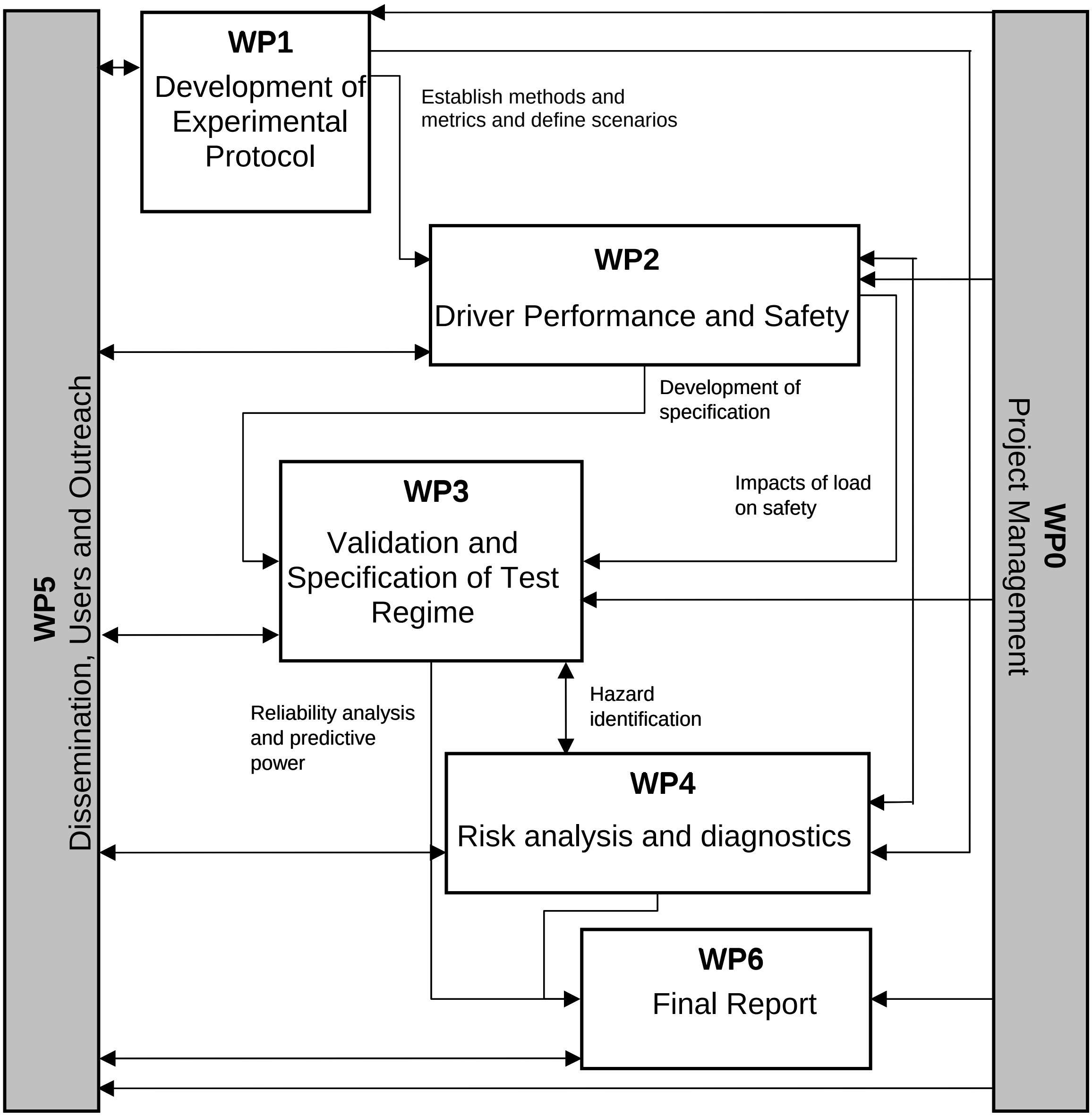


Figure 1 – Interaction between Workpackages

2.5 Workpackage 1 – Development of Experimental Protocol

Since the *aim* of HASTE was to develop methodologies and guidelines for the assessment of in-vehicle information systems (IVIS), the main effort was in studying drivers' behavioural, vehicle, psycho-physiological, and self-report measures while engaged in driving with an IVIS.

In Deliverable 1, a preliminary experimental design was presented that aimed to investigate and improve the understanding of the relationship between IVIS, driver and traffic scenario. Each of the elements has a very large number of factors that are relevant and thus of interest for investigation. However, it was necessary to reduce the considerable number of *potential* factors into a recommended set. Here, theory could provide guidance on prioritising important areas and issues on which to focus and thus how to construct an appropriate experimental design. Methodological matters such as diagnosticity, validity and sensitivity determined how to measure the effects of driving with IVIS on performance in both primary (driving) and secondary (interaction with the IVIS) tasks and therefore determined the selection of particular dependent indicators.

Rather early in the review and planning, it became evident that one major methodological problem was the lack of a reliable means of administering a "dose" of IVIS and of varying the dose administered. Considerable attention was therefore focussed on the development of suitable "surrogate" IVISs, which could provide a clear distinction between type of load being imposed on the driver and also the severity of that load in a controlled and reliable manner. These surrogate IVISs could then serve both to increase understanding of the effects of using an IVIS while driving and as a way of providing benchmarks against which driving with various real IVISs could be compared.

The following sections summarise the recommendations for the experimental work in HASTE. It was anticipated that practical and technical considerations would affect the implementation of the design as conceived here.

2.5.1 Secondary tasks

The literature review examined a number of candidate tasks that could potentially be used as surrogates for an IVIS in the WP2 experiments, which examined driver performance under task load. The result of this review concluded that, ideally, tasks should have a number of features:

- They should have clear modality (visual, auditory)
- Tasks requiring cognitive processing should be distinguishable insofar as possible from those that require only visual attention
- They should be manipulable in terms of task difficulty

The review also suggested that as well as visual and cognitive tasks, suitable manual tasks should be available to represent menu search and use via buttons, keys or touchscreens[2]. A number of suitable tasks which have been designed to measure different levels of perceptual and cognitive ability in

[2] Due to limited resources, manual tasks were not included in the WP2 experiments

human operators were identified. A list of factors which can control task difficulty was also proposed. Finally, the incorporation of crossmodal paradigms, e.g. visual tasks supplemented by auditory signals, was recommended for some of the perceptual tasks, to compare the effect of unimodal and crossmodal secondary task presentation on driving workload[3]. This review also revealed that there was not a readily available set of tasks which could be manipulated to create various levels of difficulty.

Subsequently, a set of pilot studies were carried out to identify whether it was possible to develop tasks that imposed perceptual and cognitive demand at varying levels. Whilst each of the experiments was accomplished in isolation (i.e. as a primary task) for this stage of the project, it was anticipated that all tasks would be performed in combination with driving (i.e. as a secondary task). The kind of load placed by each of the chosen tasks (i.e. simple visual attention, memory for visual/auditory information, visual-manual co-ordination/memory), was one that is usually required by different forms of IVIS, either in isolation or collectively with other loads. Promising candidate tasks were identified and further work was carried out to refine the selection and range of difficulty as well as to confirm that the tasks were suitable for use while driving.

2.5.2 Scenarios

A review of the circumstances in which interacting with an IVIS is likely to lead to safety problems, identified a set of the most critical parameters. They were:

- Driver age - with a focus on older drivers
- Driver vigilance - in the form of boredom resulting in low vigilance
- Road type - as urban, rural and motorway environments
- Junctions - as a parameter of road infrastructure
- Pedestrian facilities - as a parameter of road infrastructure
- Other road users - as creating the possibility of crossing patterns
- Special events - as potential hazards

Based on this review, *combinations* of circumstances were prepared in order to create the test roads and events for the experiments in Workpackage 2.

2.5.3 Driving performance measures

The focus of HASTE was on the influence of IVIS use on safety, i.e. on performance of the primary driving task of driving. “Performance” here is defined to include both the control level and the tactical level of the driving task. The selection of appropriate parameters and tools for measuring driving performance was clearly critical for the project. A preliminary set of mandatory driving performance measures were therefore identified. They were:

- Steering wheel reversal rate
- Lane exceedence
- Lateral position

[3] Crossmodal paradigms were also excluded from WP2, but addressed in WP3

- Standard deviation of lateral position
- Time to line crossing
- Speed
- Standard deviation of speed
- Time headway
- Distance headway
- Time to collision
- Reaction time to unexpected events
- Subjective ratings in the form of the Lund observer protocol, which is derived from the Wiener Fahrprobe observer rating scale (Risser, 1997).

An additional set of optional measures was also selected. These included physiological measures such as heart rate variability and eye movement patterns. These measures were only optional because the required tools for these measures were not available at all sites. Measurement problems were addressed, so that data collected at the various sites could be compared.

2.5.4 Workload measures

The review looked at the whole range of workload measures. These included: studying performance on the primary tasks, the application of secondary tasks to measure primary task load, the use of visual performance workload measures; subjective workload rating scales, and physiological measures. A prioritisation of the various measures was made and the following were proposed as being desirable and practical:

- Primary task performance
- Subjective workload measures: NASA TLX (Hart & Staveland, 1988) and RSME (Zijlstra & Van Doorn, 1985)
- Peripheral Detection Task
- Glance frequency and glance duration
- Secondary task performance

2.5.5 Participants

The review concluded that, if tests were to be performed with “the average driver”, participants had to meet the following criteria:

- Age: 25-50 years
- Gender: Both male and female
- Total driving experience: between 10,000 – 1,000,000 kilometres

The review also stated that some subgroups deserved particular attention. These were: older drivers, aged 60 and over, and novice drivers, aged up to 24 with an annual distance travelled less than 10,000 km and holding a licence less than one year.

2.5.6 Test procedures

The deliverable did not stipulate the final experimental design, selection of indicators or test procedures used in the subsequent experiments in

Workpackage 2. That selection was made subsequently, taking into account practical aspects and the resources available, which of necessity required some hard choices to be made. The report thus provides more general guidance and perhaps is more broadly useful as a result.

2.5.7 Summary of Workpackage 1

In sum, the main challenge of Workpackage 1 was to find a balance between transferability of results and feasibility of design when formulating a preliminary experimental design. Information was acquired on a diversity of environments, drivers, and IVIS. that are together sufficiently broad to allow the formulation of generic guidelines. At the same time, it had to be possible to conduct the experiments in such a way that they yielded adequately reliable data, and could also be carried out within the given amount of time and resources.

2.6 Workpackage 2 – HMI and Safety-Related Driver Performance

The objective of the WP2 experiments was to investigate the impact of IVIS task load on driving performance, attention and workload. To achieve this, two surrogate IVISs (S-IVISs) were created: one imposing (non-visual) cognitive load and the other visual load. Using these S-IVISs, it was possible to vary secondary task load systematically. The interaction between these tasks and driving was assessed separately, using driving tools in the laboratory[4], simulator and field.

This Workpackage also intended to identify the advantages and disadvantages of the different assessment methods (laboratory, simulator, field), and finally to identify the road types and scenarios which were most productive for testing the effects of IVIS on driving. Different groups of drivers were used and scenarios varied in accordance with the protocol and procedure for safety assessment of IVIS, as outlined in HASTE Deliverable 1 (Roskam et al., 2002). Using this deliverable as a guideline, a set of parameters which were considered highly relevant in assessing the safety impact of IVIS were chosen for the WP2 experiments. These were:

- Scenario parameters:
    - urban, rural and motorway environments
    - critical events, or road complexity level
    - junctions as a parameter of road infrastructure
- Individual parameters:
    - average and older drivers
    - nationality

2.6.1 Overview of experimental design

All combinations of road type and S-IVIS were covered at least twice in the simulator and field experiments, resulting in a total of 17 experiments and 527 participants. All simulator experiments included a standard rural road. The idea of this distribution was to have comparable experiments and to test

[4] The term laboratory refers to a low-cost alternative to a full scale simulator. The software used for this system is the same as a full scale simulator, but the driver controls and image generation are less immersive.

reliability across simulators, and validity between simulators, laboratory and field trials. The field trial situation was considered to be the reference situation (see Figure 2), although of course, the comparability to the field experiments depended on the scenarios.

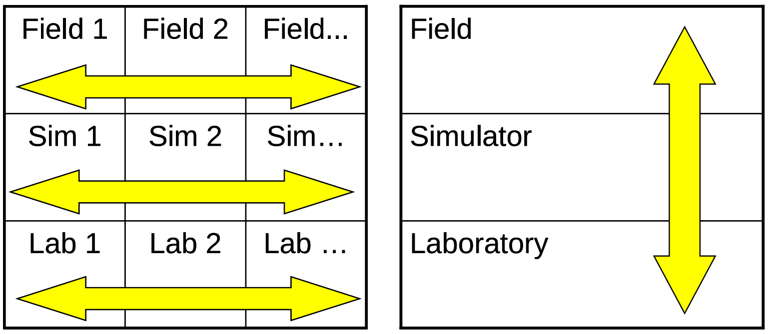


Figure 2 – Strategy for testing reliability and validity

Based on the recommendations form Deliverable 1, a group of "average" drivers (aged between 25 and 50) were recruited by each experimental site. To study the effect of ageing on driving performance during interaction with an IVIS, older participants (aged over 60) were also included in one simulator and one field study.

2.6.2 Factors and Levels

The experimental factors and number of included levels are listed in Table 1. Experiments on different road types and different age groups are considered separate experiments and are thus not included as factors. No statistical comparisons were made between S-IVIS type.

Table 1 – Factors and levels used for WP2 experiments

Factor	Levels
S-IVIS complexity level (SLv) (Within subjects factor)	4 (including baseline)
Road complexity level (RLv) (Within subjects factor)	3 (for Simulator and Laboratory - rural road) 2 (in Simulator- Motorway and Urban road) 1 or 2 (in the field trials)
S-IVIS task (Between or within subjects factor – but not compared)	2 (the cognitive task and the visual task)

A generic experimental design (Table 2) was developed for the simulator and laboratory experiments. The same design was then adopted in the field trials as far as possible. The factors in Table 2 were within-subject in all experiments. This design included three road complexity levels and four S-IVIS complexity levels. Nine activations of the S-IVIS were included in the simulator experimental drives, evenly distributed over S-IVIS and road complexity levels. Since there was a separate baseline drive, there were three observations for each RLv in the baseline drive. The number of S-IVIS activations was either 6 or 9 in the field trials.

Table 2 – Number of drives used for each road/S-IVIS level in WP2

	S-IVIS level			
	Baseline	**S-IVIS Level 1**	**S-IVIS Level 2**	**S-IVIS Level 3**
Road Level 1	3	1	1	1
Road Level 2	3	1	1	1
Event	3	1	1	1

2.6.3 The Surrogate In-Vehicle Information Systems

The rationale for choosing surrogate in-vehicle tasks has already been outlined in section 2.5.1. A detailed description of the two tasks chosen for this purpose is outlined below.

The Visual task

The design of this task was based on visual search experiments frequently used in experimental psychology. According to Treisman's Feature Integration Theory (Treisman, 1988) the speed at which a visual target is identified within a display is affected by its visual similarity to other objects in that display. Visual search experiments have shown that unique features of a target object allow it to 'pop out' of the display, resulting in faster decision times. Difficulty in target identification will therefore increase as the non-target objects become more similar to the target in colour, shape and/or orientation. In addition, increasing the number of objects in a display is shown to increase reaction time to targets, but only when a target object must be recognised by a conjunction of features (e.g. colour and shape).

This information was therefore used to create a visual task that could be presented at different levels of difficulty, based on the number and visual characteristics of 'non-target objects'. Participants were asked to decide if a target upward facing arrow was present within a display of arrows. Response was required "as quickly and accurately as possible", and given by pressing the '*yes*' or '*no*' buttons on a touch screen display, which was placed to the side of the steering wheel (Figure 3). The level of difficulty of this task was manipulated by the number and direction of the distracter arrows (which faced left, right or down), as well as the size of the display. The upward facing arrow was present on 50% of occasions . The displays for this task were presented at a system-paced rate of every five seconds, and each block of the task contained 6 displays, resulting in a 30 second 'burst' of the task. Performance on this task was measured using reaction time, proportion of correct responses and proportion of false responses.

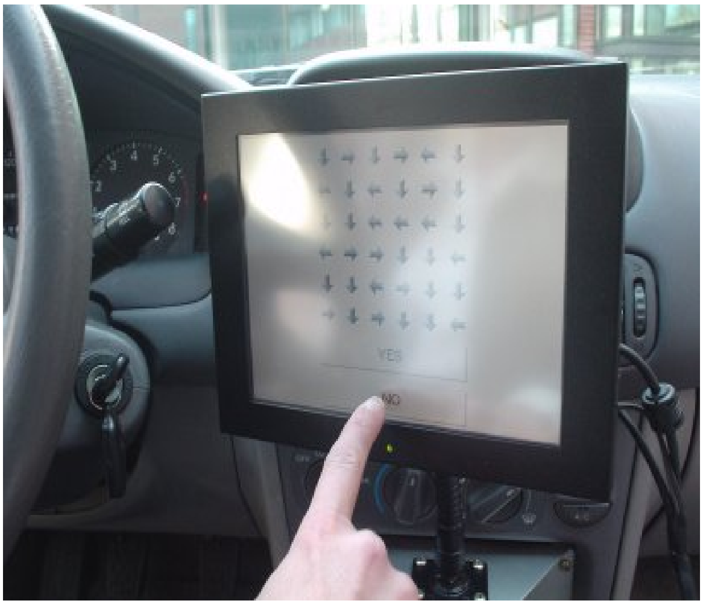


Figure 3 – Visual S-IVIS in a car (Helsinki)

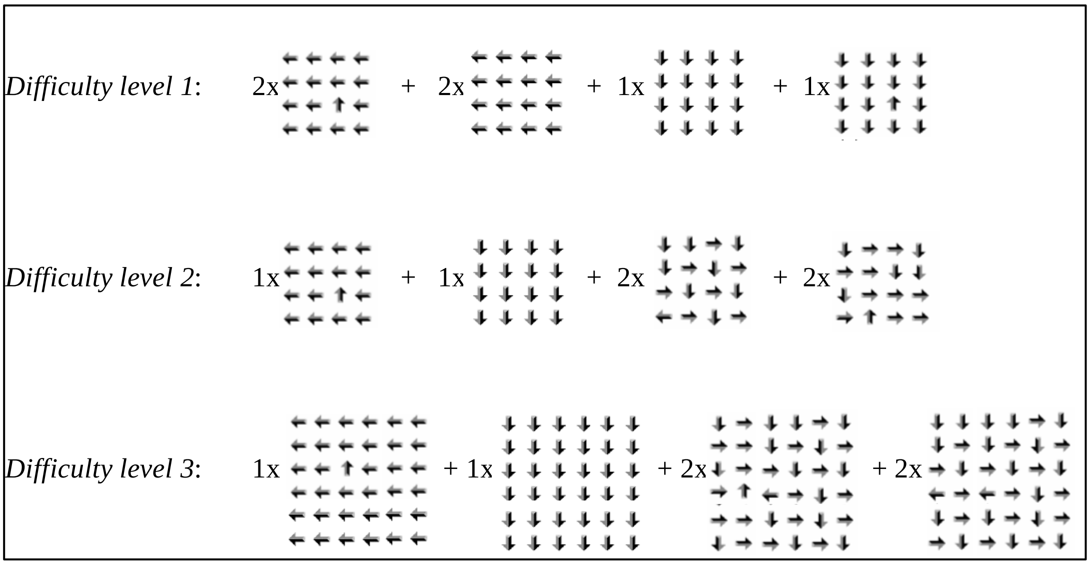


Figure 4 – Example of displays used in the visual S-IVIS

The Cognitive task

This task was adapted from a visual version of the Continuous Memory Task, used by Veltman and Gaillard (1998). The auditory continuous memory task (aCMT) involved the presentation of fifteen complex sounds, at a rate of one every 2s. Participants were asked to keep a separate count of their ‘target sounds’ each of which was identified at the start of an experimental block. Task difficulty was manipulated by increasing the target sounds from two to three to four. Participants’ response was given verbally and recorded by the experimenter. Performance in the cognitive task was measured by proportion of correct and false responses.

2.6.4 S-IVIS measures and analysis

In order to investigate if and how drivers prioritised between the driving task and the S-IVIS, an analysis of drivers' performance on the S-IVIS secondary task was considered important. For example, drivers might have chosen to reduce their effort, or even abandon the secondary task when driving became more demanding. The S-IVIS performance indicators were analysed using Univariate or repeated measures ANOVA. The effect of S-IVIS difficulty (SLv) in the driving condition was analysed; the factors included were SLv and road complexity level (RLv). Also, a comparison between using the S-IVIS as a single task (static test) or dual task (during driving) was made.

The distribution of the Visual and Cognitive S-IVIS tasks across the different road types and assessment methods is outlined in Table 3 .

Table 3 – Distribution of experiments over S-IVIS, assessment methods and road types

	Sim	**Lab**	**Field**
Urban	2 Visual, 2 Cognitive		3 Visual, 2 Cognitive
Rural	5 Visual, 5 Cognitive	1 Visual, 1 Cognitive	3 Visual, 2 Cognitive
Motorway	2 Visual, 2 Cognitive		3 Visual, 2 Cognitive

2.6.5 Assessment methods and roads

Driving simulators of varying complexity were used in the WP2 experiments (see Figure 5). The main advantage of using driving simulators and laboratories for IVIS testing is the excellent prerequisites for experimental control. These tools allow the exposure of all drivers to the same road and traffic conditions. Also, a perfect replication of these conditions between baseline drives and experimental drives is possible.

Figure 5 – Moving base driving simulator (left), static driving simulator (centre) and laboratory set up (right)

The field studies performed with instrumented vehicles were chosen because of their ecological validity, and the fact that they often provide scenarios that are not always possible in the simulator. Not only does this mean real accident risk and high driver motivation, it also includes more unpredictable aspects such as behaviour of other traffic participants, adverse weather conditions, unexpected road constructions, etc.

All road types were incorporated in the WP2 experiments. To ensure adequate cross-site and cross-cultural comparison, all simulator and laboratory studies used the same rural road design (see Table 4). Field studies were conducted at five sites, and incorporated motorway, rural and urban roads, either exclusively or in combination.

Table 4 – Simulator and Laboratory Experiments

Site	Urban	Rural	Motorway
Leeds, UK		✓	
Minho, Portugal		✓	
TNO, Netherlands	✓	✓	
Transport Canada	✓	✓	
Volvo, Sweden		✓	✓
VTI, Sweden		✓	✓

The length of the standard rural road route used in the simulator was 29 km and the signed speed limit was adapted to national standards (90 km/h in Sweden, 96km/h in the UK etc). Each lane was 3.65 metres wide. A lead vehicle was always present, which the driver was instructed not to overtake. This road contained three levels of road complexity, as defined by road curvature and behaviour of the lead vehicle. These were:

RLv1: Straight roads, requiring minimal workload compared to other scenarios.
RLv2: Gentle s-shaped curves, which required some negotiation by the driver.
RLv3: Discrete critical events, which necessitated a major reduction of speed by the driver. As well as requiring a reasonable degree of interaction with the simulator and the lead car, this type of scenario was also thought to impose maximal driving difficulty (see Figure 6).

Figure 6 – Example of a rural road critical event

The motorway road used by Volvo and VTI simulators was 46 km long and had two driving lanes in each direction plus a hard shoulder. The lane width was 3.75 metres, and the speed limit was 110 km/h. (see Figure 7). The road curvature was sampled from a road in Sweden. At predefined locations along the road, there were cars to be overtaken, and cars overtaking the participant. The road included two levels of complexity; normal driving and events. In the normal driving condition, there was random curvature, occasional overtaking cars and cars to be overtaken. The events were caused by other vehicles where on three occasions, other vehicles cut in front of the participant. These vehicles were either merging from slow travelling queues or overtaking the participant. The latter alternative involved the car braking.

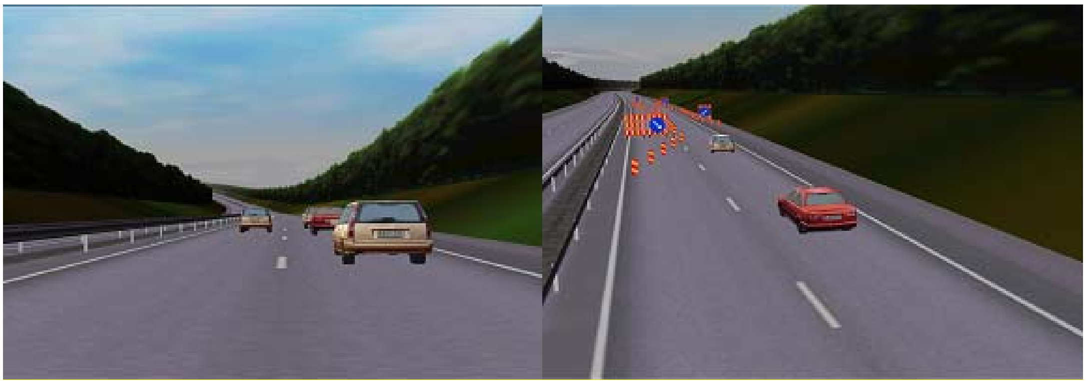


Figure 7 – The motorway environment, showing a roadwork event (right).

Urban roads were included in both simulator experiments and field trials. The simulator urban roads had a speed limit of 50 km/h and included two or three road complexity levels; straight sections, junctions (in one of the two experiments) and critical events (See Figure 8). The field experiment on urban road included 30-60km/h sections and incorporated several scenarios, such as junctions and zebra crossings.

Figure 8 – Urban simulator environment, critical event

2.6.6 Driving performance and workload measures

The effect of an IVIS on driving behaviour is not always associated with an increase in accident risk, or a reduction in safety. For instance, drivers may compensate for the distraction by reducing speed, choosing to drive a less demanding route or by increasing their distance to other road users. To examine the effect of the two surrogate IVIS tasks on driver behaviour, all partners used a set of mandatory measures which were well defined in advance to avoid any cross-site variations. Driving performance was therefore measured using the following category of measures:

- Self-reported ratings
- Lateral control measures such as steering reversal and standard deviation of lateral position
- Longitudinal control measures such as speed and distance to a lead vehicle
- Physiological workload measures such as heart rate variability and gaze behaviour
- Observations of driving performance made by an accompanying expert observer

The *self-reported ratings* were given by drivers following their interaction with each ‘burst’ of IVIS. Drivers were required to rate their self-assessed quality of driving performance on a linear scale between 1 (“I drove very badly”) and 10 (“I drove very well”). The *expert observations* were conducted in the field by an experienced driver, using a slightly modified version of the Wiener Fahrprobe observer rating scale. A more detailed description of the measures is provided in the HASTE Deliverable 2 document (Östlund et al, 2004).

2.6.7 Analysis of driving performance and workload measures

Driving performance, workload related measures and performance on the S-IVIS task were all analysed for this Workpackage. Driving and workload data were analysed to examine the safety indicators' sensitivity to cognitive and visual/motor load. The purpose of the S-IVIS data analysis was to investigate if and how drivers prioritised between the driving task and the S-IVIS task.

To generate comparable results between the experiments, a common analysis method was designed by partners at VTI. However, slight changes were introduced to account for occasional differences in design between measurement tools or road types. For instance, whilst the rural simulator road contained three levels of complexity, only two levels were used in the motorway simulator experiments. For all analyses, Univariate Analysis of Variance (ANOVA) and a 5% level of significance were used.

2.6.8 Cross-test-site meta-analysis

A cross-test-site meta-analysis was conducted to evaluate the sensitivity of particular measures, and establish which measures were the most reliable indicators of driving performance and workload. This meta-analysis was highly facilitated by the strict standardisation of the included measures, experimental design and scenarios. For each separate study, effect sizes were calculated using Cohen's d (Cohen, 1988), but only if the effect of S-IVIS on driving was significant at the 10% level. An effect size is the difference score between an S-IVIS level and baseline divided by their common standard deviation (which makes it effectively a z-score). As a convention, the following are used in the literature as descriptive of effect sizes that may occur: 0.2 = small, 0.5 = moderate, 0.8 = large, 1.0 = very large effect.

2.6.9 Participants

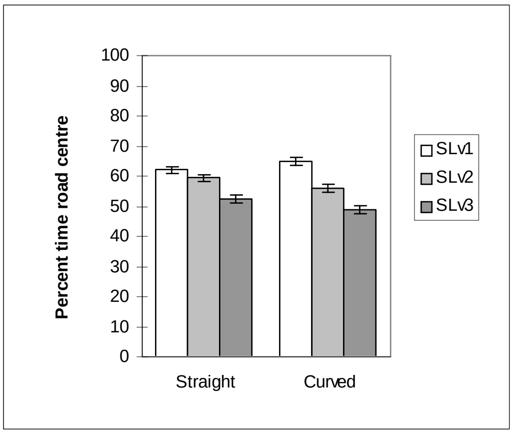
All experiments included "average" drivers, as described in WP1. Older drivers were also recruited for one simulator study (Leeds), which studied the effect of the cognitive task in a rural setting and two field studies (VTT), which included both S-IVIS tasks in a mixture of urban and rural roads. The effect of the visual S-IVIS on elderly drivers could not be studied due to simulator sickness, possibly associated with moving visual attention between the S-IVIS and the simulator road. To study the effect of any cultural differences, Portuguese drivers were compared to British drivers in a laboratory study using both visual S-IVIS (Minho).

2.6.10 Results

i. Effects of S-IVIS on driving performance

The two types of S-IVIS had quite different effects on driving performance. The visual task had pronounced effects in terms of steering and lateral behaviour; with both steering activity and lateral position variation increasing as an effect of the visual task (Figure 9 and Figure 13). The cognitive task caused reduced lateral position variation, leading to an "improved" steering behaviour (Figure 13), though there was also a tendency for drivers to shift away from the road edge with increased task load. This "improvement" in steering behaviour was accompanied by an increase in glances focussed on the road ahead (measured by Percent Road Centre[5]), at the expense of the periphery (Figure 10). Further, the visual task resulted in decreased speed, possibly as a compensatory strategy to reduce the workload, or perhaps as a result of reduced feedback from the road (Figure 11). This effect was not found for the cognitive task. The effect of the two S-IVIS tasks on headway was somewhat contradictory across sites. Whilst some sites reported no effect of this task on driving performance, others reported a much reduced headway compared to baseline driving. In the field experiments, the accompanying observers identified deteriorated speed and yielding behaviour as an effect of both S-IVIS tasks. Drivers' rating on their own driving performance was found to be very sensitive to demand from S-IVIS, showing a reduction as a result of both S-IVISs (Figure 12).

Results suggested that drivers were not always able to manage the trade-off between primary and secondary task, and that there were many indications of driving performance deteriorating most when the secondary task demand was at its highest. Elderly drivers were particularly poor at this task management.



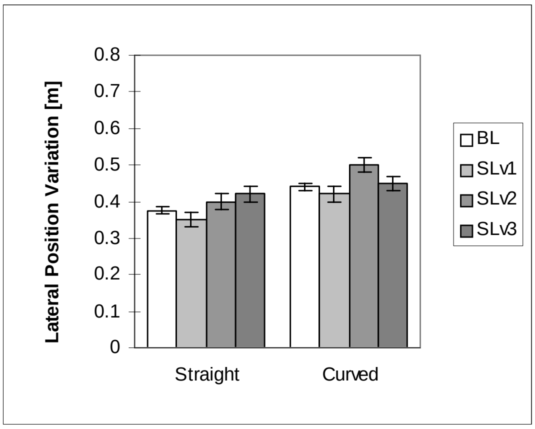


Figure 9 – Effect of the visual S-IVIS on Percent Road Centre (left) and lateral position variation (right) (simulator rural road)

[5] This is computed as the percentage of driver gaze fixations within one minute that fall within a specified area representing the road centre.

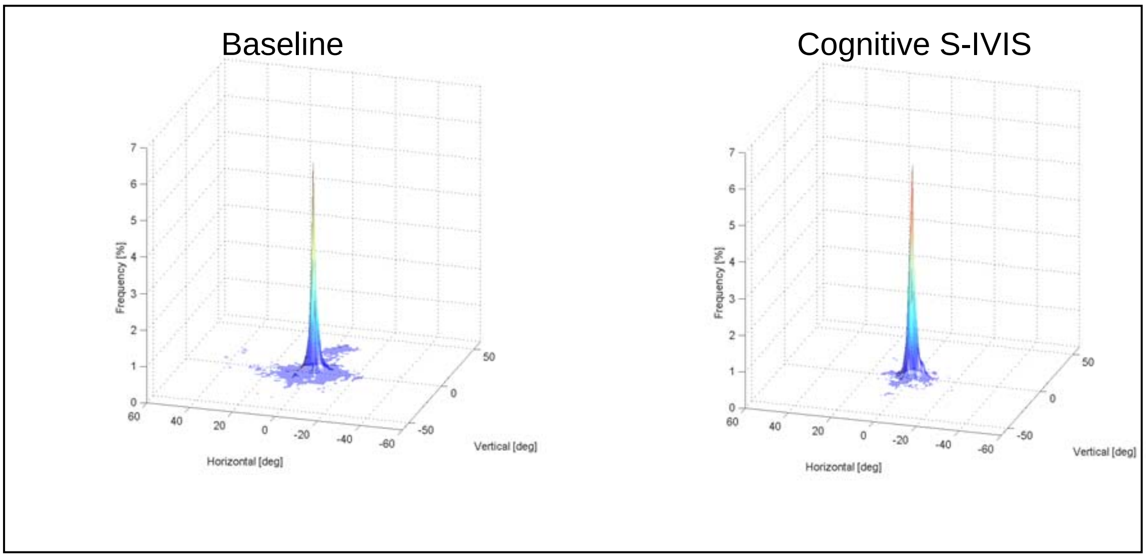


Figure 10 – Effect of the cognitive S-IVIS on gaze concentration towards road centre (motorway field)

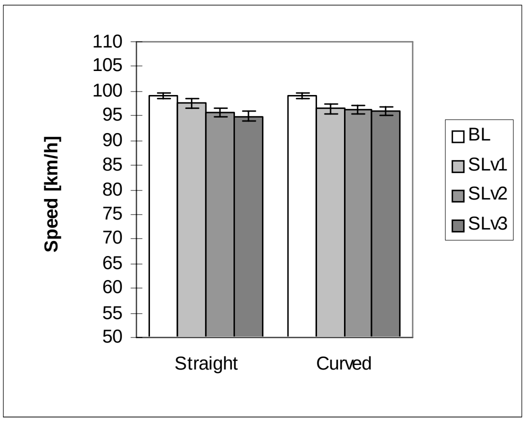


Figure 11 – The effect of the visual S-IVIS on speed (simulator rural road)

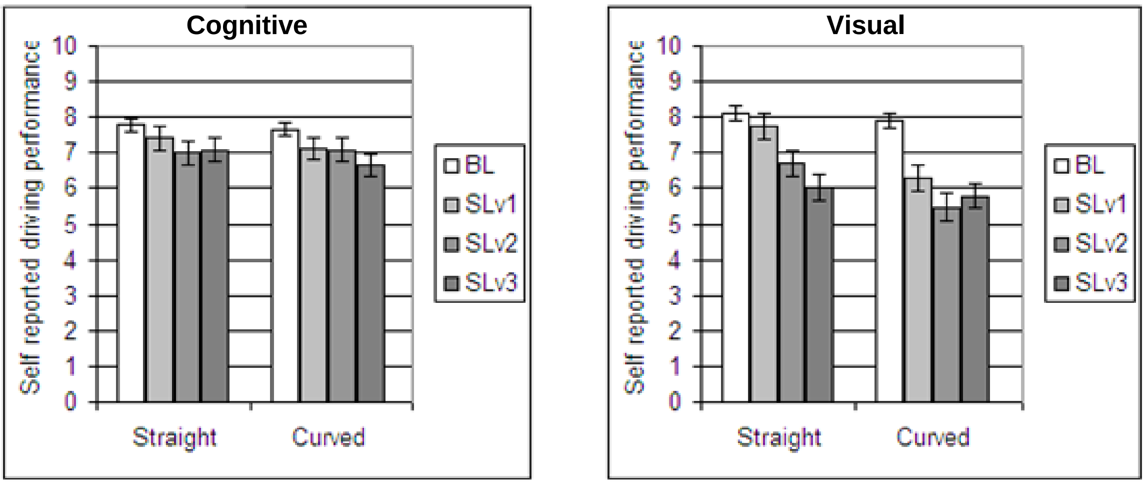


Figure 12 – Effect of the two S-IVIS on self-reported driving performance (simulator rural road)

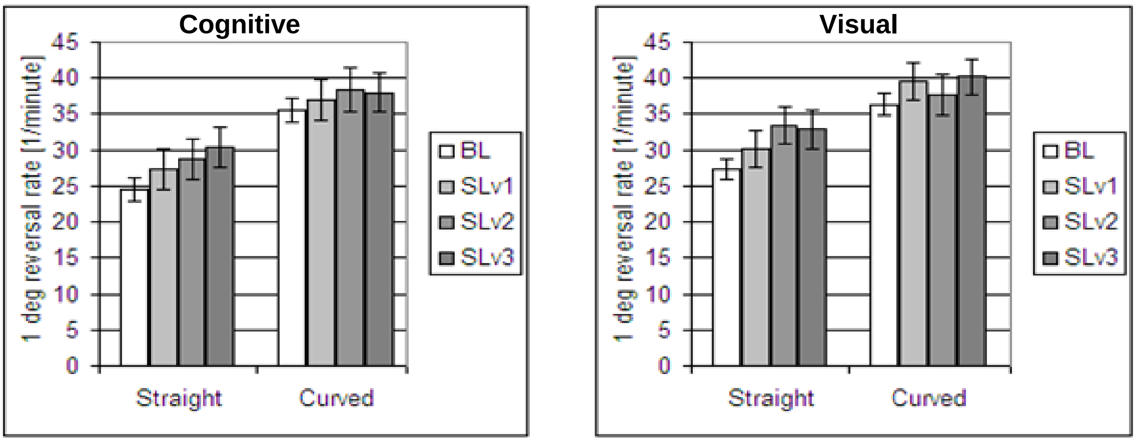


Figure 13 – Effect of the two S-IVIS on steering reversal rate (simulator rural road)

ii. Static versus dynamic performance of S-IVIS

Generally, the studies found that performance on the S-IVIS task deteriorated from the static baseline condition to the dynamic condition when S-IVIS was performed during driving. This supports the HASTE approach of requiring the driving context to be considered in assessing an IVIS. Static performance did not predict dynamic performance.

iii. Simulator vs. Field

The field studies had the propensity to highlight somewhat different effects of the systems, when compared to the simulator studies, and this was perhaps primarily due to the observation ratings. For example, inappropriate speed choice was identified ahead of zebra crossings, especially for the cognitive task (Figure 14). Additionally, simulator sickness meant that it was not possible to test elderly drivers with the visual task in the simulator. This shows the value of the field tests, but also suggests that the incorporation of some additional scenarios or tests in the simulator roads should be considered. These could perhaps take the form of detecting objects in the periphery or detecting changes in the peripheral scene.

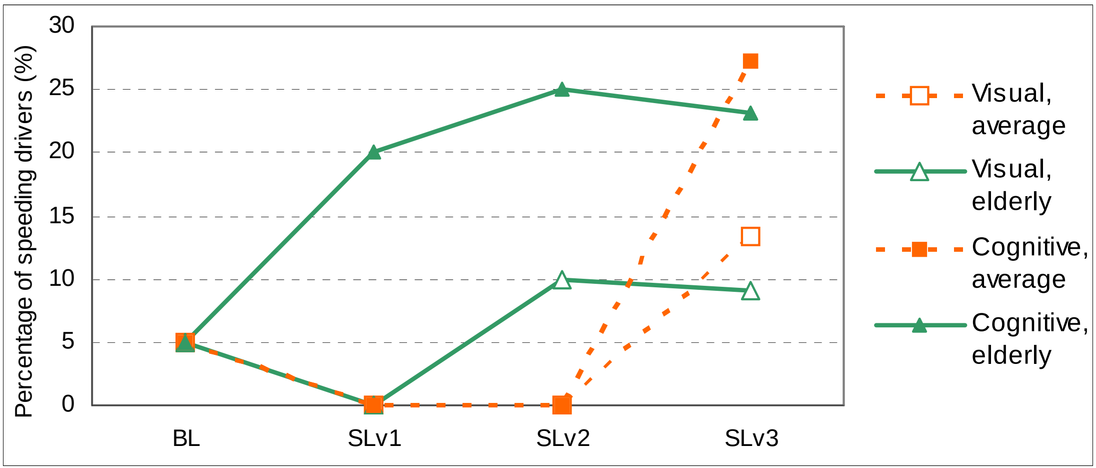


Figure 14 – Observed speeding behaviour (field urban road)

iv. Simulator Type

The broad conclusion was that the type of simulator or laboratory used in the experiments did not have an effect on the included measures in the chosen scenarios.

v. Road Category

The meta-analysis suggested that, in the simulator studies, the rural road was the most diagnostic, i.e. the effect sizes from the rural road were generally the largest. The same was true for motorway in the field experiments. The other combinations of road type and assessment method did not pick up any additional information that was not provided in simulator rural road and field motorway. This means that, in the further refinement and validation of the HASTE test regime, only these combinations were kept.

vi. Road Complexity Level

Road level was found to greatly affect driving behaviour and secondary task prioritisation, and was thus kept as a factor for the WP3 experiments. The critical events in the rural road were however found difficult to use due to learning effects and also difficulties with designing identical events in terms of impact on driving behaviour. Very few effects were found in reaction time or relevant headway measures. Critical events were thus excluded from further investigation in the WP3 experiments.

vii. "Average" vs Elderly Drivers:

The findings have confirmed the hypothesis proposed in Deliverable 1 (Roskam et al., 2002), that there would be severe problems for elderly drivers in using IVIS while driving, particularly at higher levels of task demand. Not only was the impact of task demand greater for the elderly drivers, but there were also indications that they had fewer mental resources available for managing attention between primary and secondary task. Evidence for this was found in the fact that there were fewer signs of a ceiling effect in secondary task performance for the elderly drivers than for the younger drivers, especially when the driving task was most difficult.

viii. UK vs Portugal

The controlled comparison of the British and Portuguese showed the expected effect: the Portuguese drivers exhibited riskier driving behaviours. But, reassuringly, the analysis revealed there was no interaction effect of the "country" factor. In other words, results obtained with Portuguese drivers should be as reliable as those obtained with drivers from northern Europe.

2.6.11 Methodological discussion and conclusions

As regards methodology**,** the results obtained from this very large set of studies confirm some of the initial decisions made in formulating the HASTE approach. There was clear value to the focus on *dynamic* evaluation, i.e. of looking at interaction with an IVIS while driving and of identifying the effects of that interaction on driving. Static testing cannot predict how an IVIS will affect steering behaviour or interaction with other road users. The different road levels proved their worth, particularly levels 2 and 3 of the rural road where it was very difficult to manage both driving and the S-IVISs.

Results from this Workpackage also confirmed the value of using a very large number of indicators. Some of these indicators turned out to be non-diagnostic and were therefore abandoned for the next phase of the project. Others turned out to be superfluous in that what they revealed overlap with the diagnosis provided by other indicators. The meta-analysis helped to sift through the indicators and test environments to identify the most powerful ones. The self-reported driving performance was found to be the most sensitive and reliable measure of driving performance. Other sensitive and reliable measures were mean speed of travel, lateral position variation, fast steering actions (high frequency steering), mean headway and headway variation.

The following main conclusions can be drawn from the results of Workpackage 2:

- The effect of visual load on driving was very clear: increased distraction leads to problems in lateral control.

- The effect of cognitive load was more complex, in that some driving parameters, particularly related to steering control and lateral position appeared to improve. However, this improvement seems to be an artefact of greater concentration on the road straight ahead at the expense of information acquired from the periphery. Thought needs to be given to tasks or tests that might capture this loss of information acquisition from the periphery.

- The field studies and simulator/lab studies were complementary.

- Elderly drivers exhibited very risky driving while performing IVIS tasks.

- There were also national/cultural differences in driving style. Impact of IVIS can, however, be assessed with average drivers and generalised to all nationalities and extrapolated to elderly drivers.

- Simulator rural road and field motorway were the most diagnostic and were therefore used in the next stage of experiments in WP3.

2.7 Workpackage 3 – Validation of the HASTE Protocol Specification

2.7.1 Overview

The main purpose of WP3 was to test the suggested methods and measures from WP2 on a number of tasks within real in-vehicle information systems. The overall aim of this Workpackage was to suggest a test regime for the assessment of IVIS. In particular, the intention in HASTE has been to recommend a test regime which is as cost effective as possible, and can be used both as a *pre-deployment* regime and for *final verification* of IVIS tasks. It is the intention of HASTE that this regime be usable by government organisations, consumer groups, equipment manufacturers and OEMs.

Four real navigation and traffic information systems were assessed in WP3. In addition, the TRL checklist (Stevens, Board, Allen and Quimby, 1999) was used to evaluate the selected systems, in order to assess the utility of using such checklists as part of the test regime. The result of this evaluation was then compared to the results of the HASTE experiments and the two methods were seen as a useful complement to each other.

One main outcome of this Workpackage was the results of a meta-analysis which highlighted four behavioural measures which are considered sufficient for evaluating a particular system in terms of its safety-related effects on driving. These are: subjective ratings, high frequency steering, minimum headway, mean speed. Percentage of gazes to road centre and reaction time to a Peripheral Detection Task are additional candidates for this list. The HASTE test regime also recommends that testing can be done on as little as 15 participants, and that the most diagnostic tool is a full-scale simulator exhibiting rural road driving.

2.7.2 Systems and tasks

The objectives of this WP were to assess ‘tasks’ rather than ‘systems’. This is because a system can consist of both ‘good’ and ‘not so good’ tasks, and the intention of HASTE was to have a research focus which identified the ‘bad’ tasks in particular. A final test regime would then need to weight scores across tasks.

Nine tasks were chosen for each system, on the basis of their common features across the four systems (for an example of tasks for System A see Table 5). In the field runs, three tasks were removed due to their complexity and therefore unsuitability in real traffic.

The four systems can be broadly described as follows:

System A is available on the market and consists of a removable 6.5 TFT colour display in 16/9 format with a remote control and hand controls. Route guidance information is provided with symbols, a map and voice output. A variety of displays (simultaneous arrow/map, large map or arrow) and map alignments (north, automatic, zoom on junction) are possible with this system. Examples of on-board computer functions for this system include: display of arrival time and remaining distance from destinations, current speed, distance already travelled, total journey time and average speed. Route options can be pre-set (e.g. fast/short route; avoid motorway/ferry/toll). System A included tasks with both auditory and visual/manual output.

Table 5 – Example of tasks used for System A

Task	Description	Input	Output
1	Route guidance message, incl. arithmetic information	Auditory	None
2	Route guidance message, incl. arithmetic information – more information than 1	Auditory	None
3	Route guidance message, incl. spatial information (turn by turn instructions)	Auditory	None
4	Route guidance message, incl. spatial information (turn by turn instructions) – more information than 3	Auditory	None
5	Alter volume	Visual	Visual/Manual
6	Change one item in map setting	Visual	Visual/Manual
7	Change several items in map setting	Visual	Visual/Manual
8	Destination entry – City*	Visual	Visual/Manual
9	Destination entry – City*, Street*	Visual	Visual/Manual

**Not included in field trials*

System B is available on the market and consists of a PDA and a GPS unit. The system is attached to the windscreen coupled to the vehicle power via a cable. The PDA has a colour touch display and data entry is made either by a stylus for most functions or by hardware keys. The approximate cost is €750. Data is mainly presented as visual information in a range of ways (icons, text etc) but also via voice output (e.g. route guidance information). The user has a range of possibilities to alter settings and enter information. System B had a range of tasks which were both visual and visual/manual.

System C is a system simulation of a traffic information system used in Finland (with mobile phones). In this simulation, the display consists of a removable black and white touch screen. Traffic information is provided with written messages, together with an auditory presence sign. The information is presented in menus that allow the reading of the message and a search for previous messages in a menu. This search can be made using the number of the message, or the road name. The system requires a manual action by the driver. Drivers can accept and select messages by pressing a button on the touch screen and use a scroll function to read the entire message. All tasks have a maximum presentation period of 60 seconds. The major difference between the message types is the use of different menus and the order of presentation of the messages. All tasks for System C included visual/manual components.

System D is available on the market and is the first PDA to include integrated GPS technology. It consists of a 54 x 81mm display in a 72 x 128 x 20.3 mm PDA unit. It is operated using a stylus. The approximate cost is $750 CAN (approx €465). Route guidance information is provided with symbols, a map and voice output. A variety of displays and map alignments (north, automatic, zoom on junction) are possible. Other functions include an MP3 player, appointments and contacts information, voice recorder and an SD expansion slot for flexible memory and additional software. The on-board computer functions include a display of arrival time and remaining distance from destinations, current speed, distance already travelled, total journey time and average speed. Route options can be pre-set (e.g. fast/short route; avoid motorway/ferry/toll). System D had a range of tasks which were both visual and visual/manual.

Task difficulty for each system was determined a priori, based on factors such as time on task, number of button presses and level of manual complexity.

2.7.3 Participants and sites

Experiments included average drivers (as described in WP1 and WP2). Laboratory, simulator and instrumented vehicles were used to investigate the interaction between driving and real IVIS tasks.

2.7.4 Dependent measures and analysis

Based on the results from WP2, the most sensitive dependent measures were selected for this Workpackage. In addition, since the cognitive S-IVIS used in WP2 showed a reduction in peripheral gaze, two sites included a Peripheral Detection Task (PDT) to examine drivers' reaction time to stimuli presented in the visual periphery. Performance on the PDT was measured using percentage of hits, reaction time, percentage of misses and percentage of cheats.

Since the duration of tasks varied greatly, some of the driving related measures used in this WP were shown to be biased by task length. To overcome this bias, a 'sliding window' standardisation technique was used for tasks which lasted longer than 15 seconds (see Johansson et al., 2005, for further details). The measures used in this Workpackage were grouped into the following categories:

- Lane-position and time-to-line-crossing measures
- Steering wheel measures
- Speed and headway-related measures
- Eye Movements
- PDT measures
- Subjective ratings

The analysis method designed for the WP2 experiments was also adopted for the WP3 experiments.

2.7.5 Results

The results from WP3 confirmed the results from the WP2 experiments. The effects were more pronounced for the visual and visual-manual tasks compared to the auditory tasks. As in WP2, it was found that there were somewhat different effects on driving from the auditory versus visual IVIS tasks. However, one drawback in the WP3 experiments was that there were only a few auditory tasks with a reasonably high cognitive load.

The most sensitive driving measures were found to be subjective rating, headway, mean speed and high frequency steering. In each case, effects from the secondary task on the particular measure were plotted according to the a priori classification of task difficulty, defined for a particular system (see Figure 15 for an example). As shown in Figure 16, subjective rating of driving performance was found to be a good indicator of changes in driving difficulty with task. In general, subjective ratings were found to fall as the complexity of the tasks increased.

Task number	1	2	3	4	7	8	9	5	6

Easy → Difficult

Figure 15 – A priori classification of Tasks for System A.

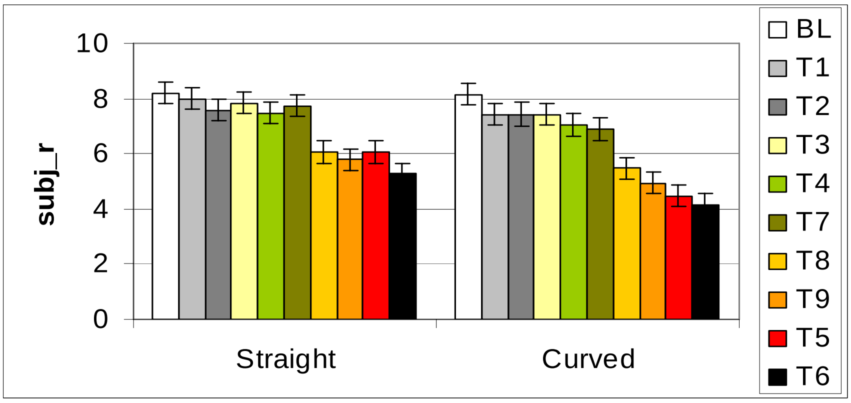


Figure 16 – Subjective rating by task length (system A).

Overall, subjects were seen to increase their distance with the lead car as difficulty of the secondary task increased. Figure 17, shows an example of this from the Leeds simulator studies. This was coupled with a reduction in average speed as difficulty of the secondary task increased (Figure 18).

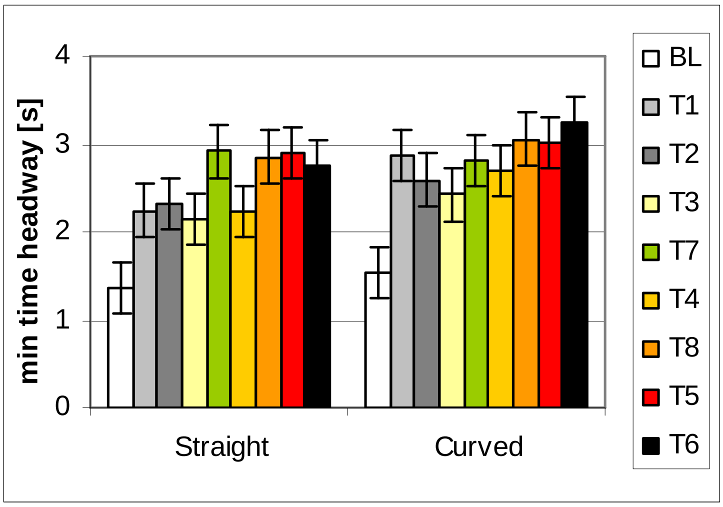


Figure 17 – The effect of each task on minimum time headway (System B)

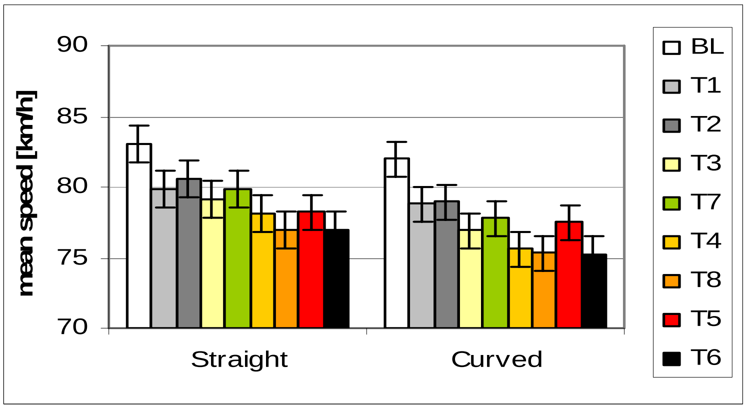


Figure 18 – The effect of each task on mean speed (System B)

The high frequency component of steering was also seen to increase from baseline, with larger values seen for the more visual tasks (T7, T8 and T9) than for those with a mainly auditory output (Figure 19).

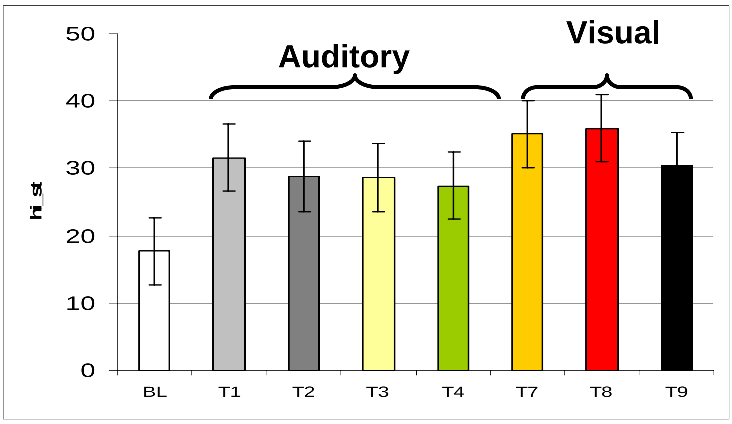


Figure 19 – The effect of each task on high frequency component of steering (System A).

The effect of task difficulty on eye movement measures such as Percent Road Centre was quite interesting, in that there was a clear distinction in this measure between auditory and visual tasks (Figure 20).

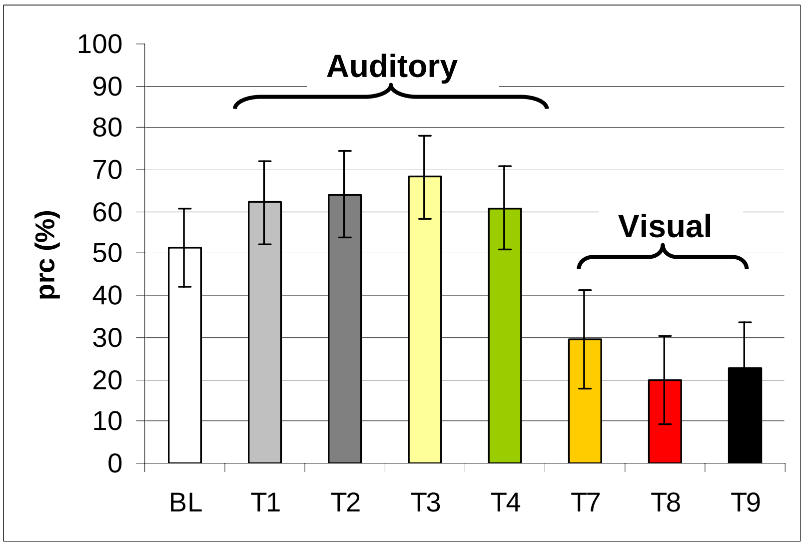


Figure 20 – Effect of each task on Percent Road Centre (System A).

Finally, reaction time to a peripheral detection task was also found to be quite sensitive to secondary task difficulty. This value was seen to rise as workload imposed by subjects increased as a result of secondary task difficulty.

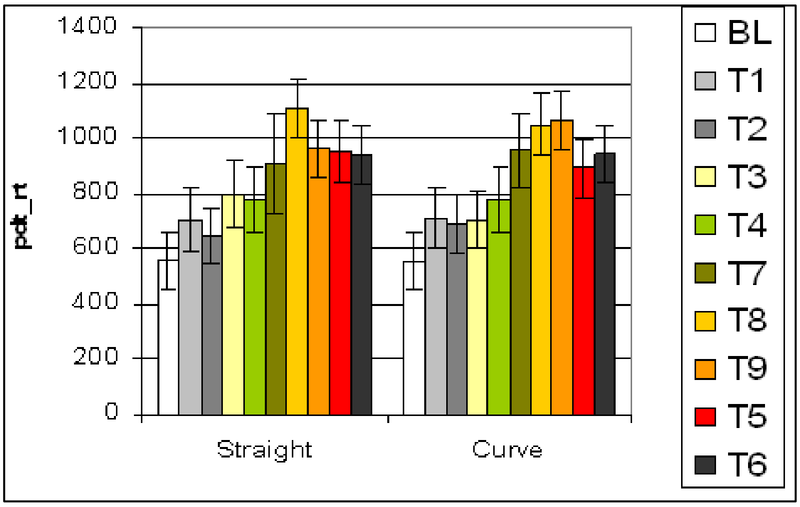


Figure 21 – The effect of each task on reaction time to the PDT (System D)

2.7.6 Meta-analysis

A meta-analysis of the results of all studies in WP3 was also performed in order to (i) bring out and grasp the common patterns in all experiments, (ii) to identify and select the most powerful parameters for detecting these patterns, and (iii) to check whether the conclusions drawn from the meta-analysis of the earlier S-IVIS studies (WP2) would hold up for real systems. If so, it could be concluded that results of the present studies provided the most important ingredients for a test regime for an IVIS. This analysis mainly focused on the mandatory measures, i.e. the measures used by all partners. The Peripheral Detection Task measures and those reflecting eye movements (Percent Road Centre) were also examined since they appeared to be promising on the basis of earlier results. However, since these measures were not tested in all experiments due to lack of resources, results should be viewed with some caution.

To qualify for meta-analysis, each measure had to pass the following four criteria:

i. Check for significant main effects on a .10% level.
ii. Check for consistency of effects (i.e. if the dependent measure varied in the same direction as task difficulty)
iii. Effect sizes (Cohen's d) and range of effects (i.e. detailed check of how strong the effects actually were and if they were sufficiently diverse over the task levels)
iv. Obtaining final index of parameters' discriminative power (i.e. adding a parameter's average effect size and its range, and multiplying the sum by its consistency value)

Based on the meta-analysis of the mandatory measures, subjective rating of driving performance, mean speed, high frequency steering and minimum headway were the measures which made it through the above criteria. In addition to these measures, the Peripheral Detection Task measures (reaction time and hit rate) as well as Gaze concentration (Percent Road Centre) were analysed. The number of hits and reaction time to PDT were both found to be significant (at the 10% level) in 5 out of 6 studies. Formally, the PDT parameters therefore failed the very first criterion for including them in the further steps of the meta-analysis, which is that they should have been shown to be significant in at least 90% of the studies. Nevertheless, in the studies in which they were significant, the two parameters showed good discriminative power. Percent Road Centre (PRC) was measured in the two VTEC field studies and at Leeds. The meta-analysis showed that the measures were significant in all three, and consistent in two studies. However, this parameter is a special case in that it specifically captures visual activity rather than driving performance. Thus, we would have to look into the modality of the underlying tasks in order to see what 'consistency' would mean in this case. That is, visual and cognitive tasks affect PRC values in different directions (visual tasks show lower PRC values and cognitive show higher).

The meta-analysis also indicated that, for most parameters, task modality does not make a difference. The exceptions are subjective rating and PRC. The latter is maybe not surprising, since PRC specifically captures visual activity (in case of cognitive tasks PRC even goes the opposite way, explaining the inconsistency in the results mentioned in the preceding paragraph). The former appears not to have an easy explanation.

Results from the meta-analysis also suggested that a full-scale simulator would be a more useful tool for safety assessment of IVIS, when compared to a less complex laboratory set-up, for instance. This was partly due to more uncertain patterns from the laboratory experiments in terms of steering and car-following measures. When comparing simulator and field experiments, the effects seen on driving performance in simulators were clearer than in the field. However, the inclusion of field studies in this context is thought to be important since this naturalistic setting allows the observation of scenarios which are not always possible/considered in simulation.

2.7.7 Expert assessment with TRL checklist

While one of the main aims of HASTE was to explicitly focus on driver-behaviour related parameters for the safety assessment of systems, there are clearly alternative approaches, specifically those that rely on expert assessment. The TRL checklist (Stevens, et al., 1999) is one of the most prominent of these. The systems used in HASTE WP3 were therefore also tested with the TRL checklist, to establish whether the two approaches could be compared. Since the checklist yields judgments that are naturally qualitative, they are relatively difficult to compare with the quantitative results from the HASTE experiments. However, the TRL results could be seen as a good complement to explain the results gained in these experiments.

2.7.8 Towards a draft test regime

A *full size simulator* is considered to be sufficient to capture differences between tasks and baseline (no tasks), as well as establishing the effect of different tasks. The road environment suggested is a *rural road* with straight sections[6]. Normally, *15 participants* are considered sufficient for testing the effect of an IVIS task on driving performance, and the following metrics should be sufficient to provide an indication of the effects of a particular system on driving: s*ubjective ratings, mean speed, high frequency steering, minimum headway, Percent Road Centre and PDT reaction time.*

The strengths and weaknesses of each of the measures can be summarised as follows:

Subjective rating (average effect size in experiments 2.02). This parameter was found to be powerful in both WP2 and WP3 experiments. The use of this parameter is quite easy and also inexpensive. However, it works best following appropriate instructions by the experimenter and adequate understanding by participants.

Mean Speed (average effect size in experiments 0.99). This parameter appeared to be the best from among the collection of those describing speed behaviour. Mean speed is measured quite easily by most simulators and instrumented vehicles. However, caution must be taken when using this parameter: generally, IVIS use resulted in reduced speed but a reduction in

[6] The WP2 studies did not find sufficient interaction between the straight/curved sections of roads and IVIS task level.

speed is not necessarily a safe behaviour in all circumstances (e.g. fast lane of a motorway).

High frequency steering (average effect size in experiments 1.08). The measurement of this signal is reasonably easy in both the simulator and the field. However, individual differences in driving may affect interpretation of this measure.

Minimum headway (average effect size in experiments 0.96): This is an easy measure to achieve, although it does require the presence of a lead vehicle.

Percent road centre (effect size not available): This measure is easy to calculate but its use is only possible with the right tool.

Peripheral Detection Task (effect size not available): It is quite easy to use this measure, although the right tool is required for its administration.

When a test regime has been defined, it can be used in a range of approaches, depending on purpose and preference. For example, the idea in HASTE is that the test regime should be used both for pre-deployment phases (e.g. throughout early design phases and iteratively during system development) as well as for the final safety validation of a system. In order for the test regime to be as constructive as possible, it is best used as a tool by vehicle and system manufacturers. If the test regime is to serve as a tool during the design phase, the focus of the test regime should ideally be on tasks rather than systems.

Almost regardless of whether assessment of an IVIS is summative or formative (or both), the regime could still be presented in a range of different ways. Some examples are code-of-practice, ISO standard, Pass/Fail criteria or as a testing procedure providing the general public with information on a product.

2.8 Workpackage 4 – HMI Safety and Risk Analysis

2.8.1 Overview

The objective of WP4 was to examine how a preliminary hazard analysis of an IVIS concept or design could be carried out and to examine how other IVIS safety hazards, including those related to reliability, security and tampering could also be identified. This WP initially reviewed the current state-of-the-art concerning the techniques developed for the preliminary safety analysis of automotive systems in general to identify those that had the most potential for application to an IVIS HMI. The applicability of these techniques for IVIS was assessed in the context of applicability within an industrial design and development lifecycle. A potential evaluation methodology was then defined for the hazard identification and risk analysis of an IVIS concept or design. The effectiveness of this selected methodology was then validated by applying it to the HMI of existing IVIS system designs and the risk assessment results compared with that given by in other HASTE analysis performed in

WP3. The results from this comparison process then guided the definition of a final IVIS HMI assessment methodology called the Driver Operability Procedure (DOP). Additional guidance was also given as to how this procedure could be applied within an industrial design and development process.

2.8.2 Initial issues identified

The objectives of other WPs within HASTE were to develop a test methodology that could be used to evaluate the impact of IVIS operation and use within the context of driving. This test methodology was evaluated in a range of real and virtual settings. In all cases this required a real or a detailed simulation of an IVIS and functionality that the subjects could operate within a driving task setting. However it is acknowledged that such a level of detailed design completion is only achieved after considerable development from an initial concept definition phase. It was also acknowledged that it is necessary to consider human operability of an IVIS at earlier stages in development to ensure that an acceptable product is eventually delivered to market. The preliminary safety assessment procedures are intended to provide such guidance and the HASTE analysis should augment existing procedures and processes within industry.

It was also acknowledged that the design of an IVIS in an automotive application is potentially affected by a wide range of product standards, design regulations and other legislation. This noted that risk assessment methodology already makes an important contribution to the design, development and manufacture of an automotive product to ensure that products are “safe” and do not contribute to increased injury and/or accident risk. It was therefore necessary to examine the context within which the safety of an industrial product is established.

2.8.3 Industrial design process

Standards such as IEC 61508 are based on a safety lifecycle that is intended to be conducted in parallel with the overall engineering process for a system. The standard was developed against the background of industrial process control. In this context, there is an item of “equipment under control” (EUC). The EUC may have a control system. Safety functions are added separately to mitigate against hazardous states of the EUC and/or its control system. The safety functions are implemented either in the EUC control system, or in a separate safety system. IEC 61508 and its safety lifecycle applies to these safety functions when they are implemented in an electrical system, an electronic system or a “programmable electronic system”. While many aspects of IEC 61508 are applicable to the engineering of vehicle systems, the safety lifecycle does not align well to the traditional vehicle engineering model; in particular:

- Both vehicles and their electronic systems are developed on the basis of a number of iterative cycles and “samples”;
- Final validation is performed before the products are released to sale (e.g. through Type Approval) rather than during installation and commissioning.

Figure 22 shows a diagrammatic representation of the safety lifecycle from ISO 61508.

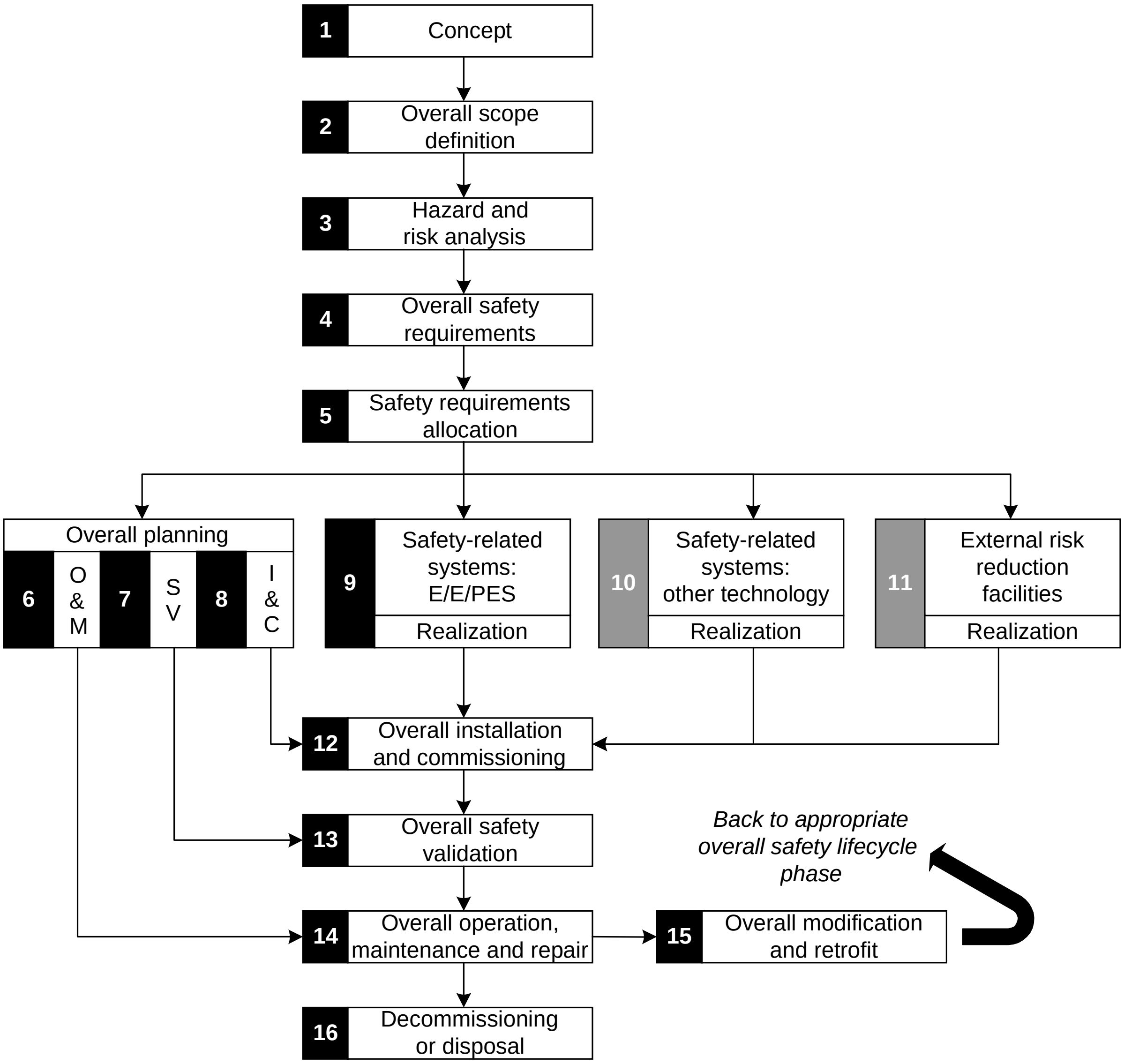


Figure 22 – IEC 61508 safety lifecycle

2.8.4 Safety and safety cases

The overall objective of HASTE was to provide criteria whereby the safety of an In-Vehicle Information System (IVIS) can be assessed for its potential use by a driver while driving. Since IVIS are complex devices it is unlikely that any assessment/certification will be done using the classic "pass/fail" techniques of Statutory Type Approval (STA). Instead it will be necessary for the developer, or importer (if the device originates from outside the EU), to create a Safety Case for its intended use. This approach is already common in other industry sectors, and has recently been added to the STA regulation for braking [Annex 18].

A Safety Case is a formal presentation of evidence, arguments and assumptions aimed at providing assurance that a system has met its Safety Requirements and that the Safety Requirements are adequate. At the beginning of a project consideration needs to be given to the logical argument that will be used to demonstrate that the final IVIS is safe to use. This can be structured using Goal Structured Notation. Objectives, or *goals*, are sub-divided into sub-goals until a *means* of demonstrating those goals can be identified. These means will then form the safety validation part of the system development process. Goal Structured Notation can also be used to present a Safety Case, though an alternative method is to use a Claims-Argument-Evidence analysis. Using this method an item of *evidence*, e.g. the results of some tests, created during the development process is used to support a sub-*claim*. These sub-claims are then brought together in an argument to demonstrate the validity of the top claim.

This raises the question of how a safety case process exists within industry. Previous work in safety-related systems assessment (e.g. DRIVE Safely, PASSPORT, MISRA) has developed a process called “Preliminary Safety Analysis” (PSA) that can be used to identify the safety properties of a concept system. The HASTE project has asked whether is it possible to define a “PSA”-like process that can be applied to analysis of the human factors aspects of a concept, specifically those related to an IVIS (In Vehicle Information System).

2.8.5 Application of assessments to IVIS

The lifecycle identified above indicates the sequential progression of initial concepts to mock-ups, engineering prototypes and eventual manufacture ready approved design. They indicate that in an industrial context design processes have to operate within a complex procedure that includes incremental development of systems and integration to refine a design from an “idea” to a finally accepted defined design. If no relevant HMI evaluations are carried out within this process then it is possible that HMI operability risks may become built-in to the design and difficult or impossible to remedy close to manufacture. It was therefore within the objectives of HASTE to consider how such a risk assessment or operability study can be delivered within a concept development process and a procedure called a Driver Operability Procedure (DOP) was proposed and its applicability is illustrated in Figure 23 below. In this context, “concept” is understood to mean an idea or a requested feature. “Prototype” means any kind of pre-production sample. “Product” means production-intent samples, volume production and also covers in-service issues.

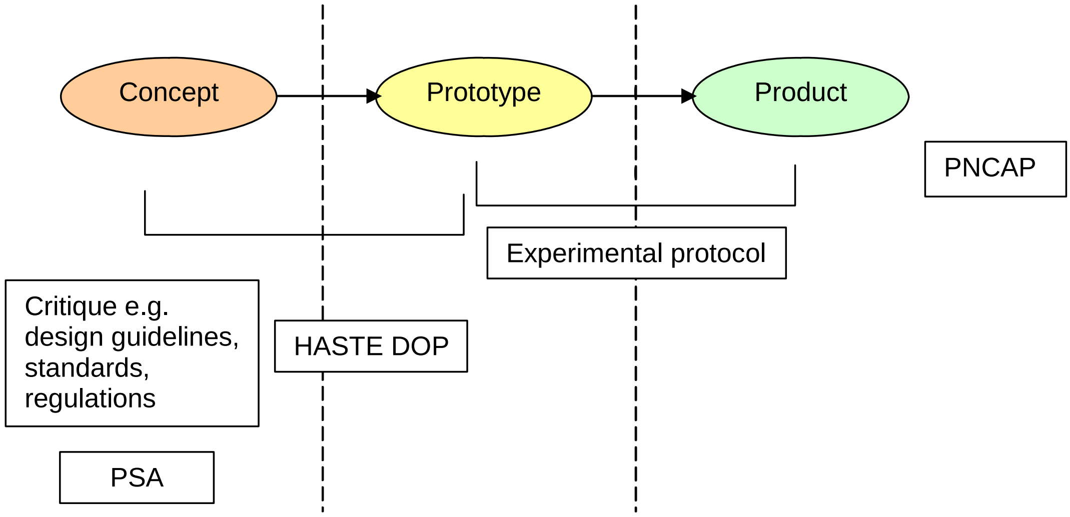


Figure 23 – Scope of HASTE DOP

This figure shows that very early concept stages may not contain enough detail of HMI design to enable meaningful analysis to take place. At this initial stage concept development should take appropriate note of published design guidelines, standards and regulations to guide development. When a more detailed concept specification has been developed prior to prototype development then a DOP can be applied. The figure also shows that a Preliminary Safety Assessment analysis (PSA) should also be undertaken on the IVIS concept. This is because, even if no functional safety hazards are obvious, it is necessary to demonstrate explicitly that they do not exist. This could include issues such as system security and tamper proofing. The PSA is therefore complimentary to the HASTE DOP.

2.8.6 Comparison of risk assessment procedures

In general, any risk or hazard analysis process consists of the following basic steps:

- Identify the risks or hazards associated with a system or process
- Classify them in some way
- Record the results of the analysis to permit review at a later stage.

The content and application of various risk assessment procedures were reviewed.

PASSPORT

The PASSPORT process for preliminary safety analysis was developed during the eponymous DRIVE II project. It was originally developed for analysis of what where then called “road transport telematic” systems, and has subsequently been adopted for in-vehicle systems by the MISRA Guidelines. A PASSPORT PSA consists of the following stages:

- Model the system under evaluation using a modified form of context diagram

- Carry out a "what if" analysis on scenarios to determine potential hazards of the system
- Carry out a "what causes" analysis on these potential hazards
- Determine top-level safety requirements for the system.

PASSPORT PSA can be applied when a system is only at the concept stage, and has the advantages that there does not need to be a design for it to be applied and that safety requirements can be considered for all stages of a system specification and design. It provides a way to apply a structured approach to what are essentially informal analyses of informal ideas or designs. The approach has to be applied up to the system boundary, i.e. the system is treated as a "black box" and any failures are assumed to occur at the "interfaces" or "boundary elements", namely the point at which information enters or leaves the system. However it is difficult to see how this technique could be applied to any form of preliminary analysis at the concept stage of an IVIS. At the concept stage, an IVIS is likely to exist only in the form of a stated requirement to have such a system, probably from a marketing department. Any analysis of failures at the system boundary is likely to lead to the same answers no matter what the system (e.g. driver misreads display, display blank, …), though the classification of these hazards may differ, and this will affect the degree of rigour needed of the development process.

A parallel recommendation is for detailed safety analysis (DSA), which is essentially a formal framework for the application of techniques such as FMEA and FTA. The PASSPORT DSA recommendations are not widely available. MISRA is developing a guidance document on automotive safety analysis that will provide a similar framework.

FMEA

Failure mode and effects analysis (FMEA) is a process widely applied in the automotive industry to identify potential failures and their consequences. It can be applied to the design of a component or system, and also to a process such as production. FMEA requires that there is a design or similar mature set of information on which the analysis can be based. NB in strict terms FMEA should be referred to as "fault mode and effects analysis". Generally the deviation of systems or processes from their design intent follows this sequence:

- There is a **fault** in a component or part of the system
- This leads to an **error** in the state of the system
- This leads in turn to the **failure** of the system to perform to specification.

FMEA is therefore, strictly speaking, concerned with identifying faults and determining what failures could result. A further issue that has to be considered is the system boundary and the point at which the effects (failures) are manifest. There are usually three boundaries that have to be considered:

- The boundary of the "target of evaluation" – the system, subsystem or component on which the analysis is being performed;

- The system boundary (usually the point at which the systems sensors and actuators observe and act on the “thing” under control);
- The hazard boundary at which the hazardous occurrence will be observed (usually at the external “skin” of the vehicle).

FTA

Fault tree analysis (FTA) is a process applied to the same set of data used for FMEA, but the process is run “in reverse”, starting from a specified failure and exploring the faults that could lead to it. Essentially each failure is decomposed into an hierarchy of lower-level events that could cause it, with the analysis following down to the level at which a basic event occurs (e.g. a wire breaks) or a fault is identified in an item for which a separate analysis is available. FTA is usually presented in a tree-like structure (see Figure 15), with the failure at the top of the tree and the combination of events leading to it presented underneath. Multiple events can be combined with “AND” gates (i.e. they must all occur for the next level event to occur), or with “OR” gates (i.e. if one or more occurs, then the next level event will occur). FTA is particularly useful for calculating predicted failure rates for systems, as individual low-level fault probabilities can be combined

HAZOP

Hazard and operability study (HAZOP or sometimes HAZOPS) is another form of hazard analysis that was originally developed in the chemical engineering industry but has now found wider applications [6, 7]. This has found HAZOP applied successfully to many sectors and to systems based upon various types of technology (electrical, hydraulic, etcetera) and to many different types of systems. A HAZOP analysis starts with a postulated deviation from design intent (effectively the “error” in the 3-step event sequence described above) and examines both what could have caused the error (i.e. the fault that caused it) and the hazard it could lead to (i.e. the failure resulting from it).

HAZOP is based on a series of entities, attributes and guidewords, and the hazard analysis is conducted by asking questions in the form:

What if [entity].[attribute] = [guideword] ?

The **entity** is the lowest level of component, system or function that will be examined in the analysis.
The **attribute** is an identifiable state or property of the entity.
The **guideword** describes a deviation from the intended design behaviour. There is a basic standard set of guidewords although these need to be interpreted in the context of the analysis being undertaken.

HAZOP can be applied to a concept (although it requires some sort of design to exist) and also to operational conditions. It is considered to be particularly effective for new systems or novel technologies. The relationship between FMEA, FTA and HAZOP may be summarized in Figure 24 below.

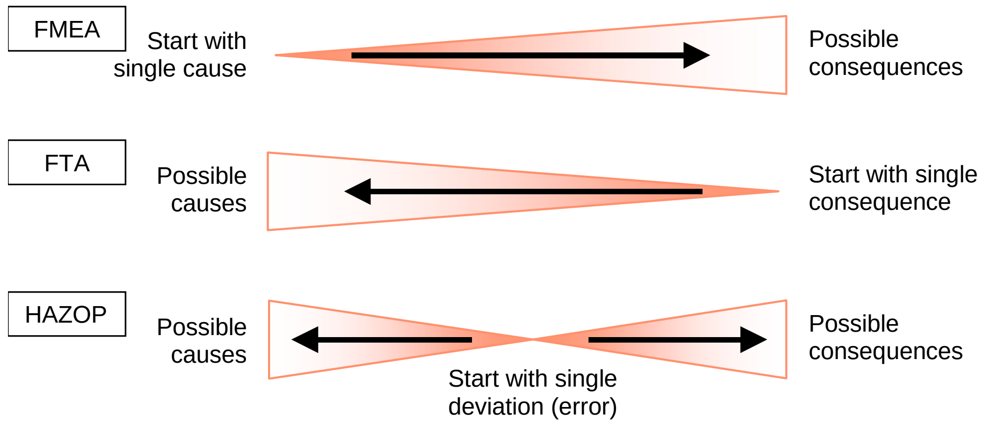


Figure 24 – Comparison of FMEA, FTA and HAZOP

The possible approaches for providing a framework for a Driver Operability Procedure for the assessment of an IVIS HMI were considered. Based upon experience of applying existing techniques to automotive products and functions for PASSPORT PSA, DSA, FMEA and FTA techniques and exploratory examination of HAZOP in the context of an IVIS HMI, an overall evaluation was made. This assessed the comparative potential for each of these techniques. The following table is a summary of results of this analysis by the expert group within the HASTE project.

Table 6 – Comparative Assessment of Technique Applicability to Lifecycle Phase

	Lifecycle phase			**Notes**
	Concept	***Prototype***	***Product***	
PASSPORT PSA	✓	✗	✗	
PASSPORT DSA	✗	✓	✗	Full details are not widely available
FMEA	✗	✓	?	Used for analysis of production processes
FTA	✗	✓	?	Used for generating service trees
HAZOP	✗	✓	✓	

This analysis therefore suggested that a HAZOP based technique is promising as a basis for the IVIS HMI analysis and was the subject of further development within WP4.

2.8.7 Validation of the HAZOP based DOP procedure

The development of a HAZOP based DOP within HASTE was carried out in two sequential stages. Firstly a draft procedure was defined based upon an IVIS specific application of the basic HAZOP process. This was then applied to an IVIS concept that was developed within HASTE and modelled on a generic form of IVIS technology already on the market, namely a speed camera location warning device. The analysis of this generic IVIS form by the draft DOP was then assessed and some procedural modifications were identified and guidance developed on the application of the guidewords for analysis.

The resulting modified DOP was then applied to a single real-world IVIS that had been used in earlier HASTE work to validate the approach. System B defined in section 3.3.2 above was chosen for this evaluation. This was a GPS equipped PDA running a route guidance application. The IVIS selected had manual data entry functionality and both visual and auditory outputs to the driver.

After familiarisation with the system the modified DOP was applied. Firstly an **IVIS definition** was developed to specify the hardware realisation of the IVIS and its likely installation and usage; this replicated that level of detail that would be available at an early development stage. Secondly a formal **functional definition** diagram was derived. This defined the states and state changes that would define the IVIS functionality. From these two inputs a **data flow** diagram was defined that defined the interface interactions between the IVIS and the user/driver. Finally a systematic application of the DOP guidewords was applied and the results noted. This was carried out as a team assessment exercise with software, electronics and human factors experts involved. This resulting definitions and analysis diagrams are illustrated in the figures below.

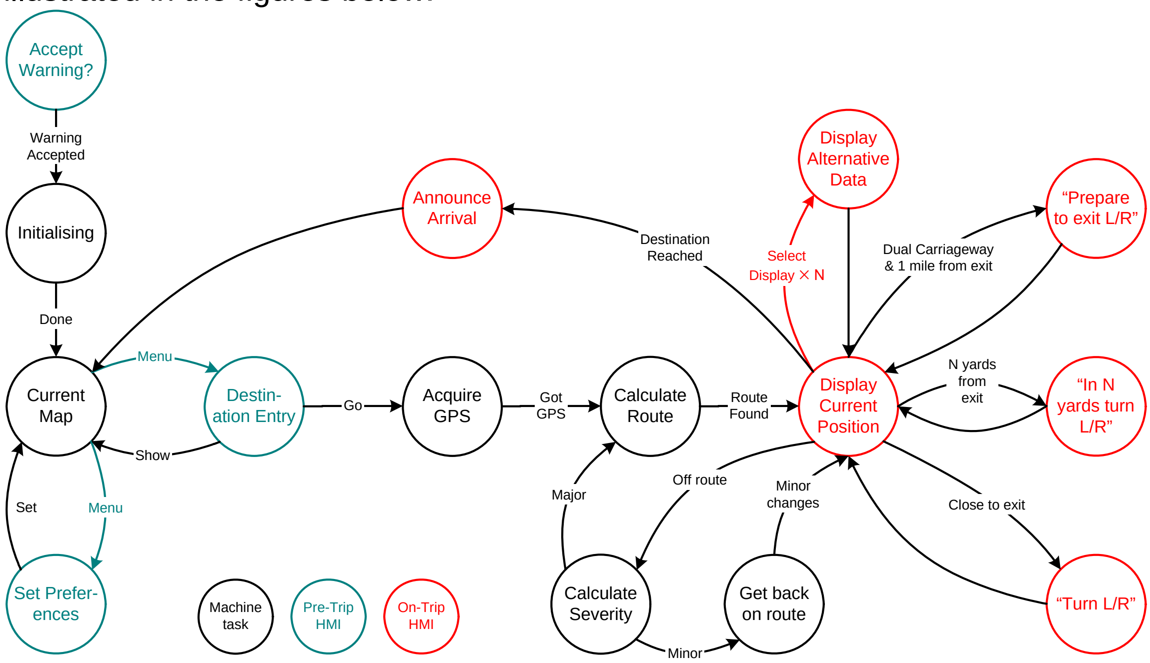


Figure 25 – Functional Definition Diagram of System B

Following this stage a data flow diagram was constructed and the data flows were evaluated using the HAZOP and DOP guidewords. The figure shown below illustrates the data flows for the real IVIS function.

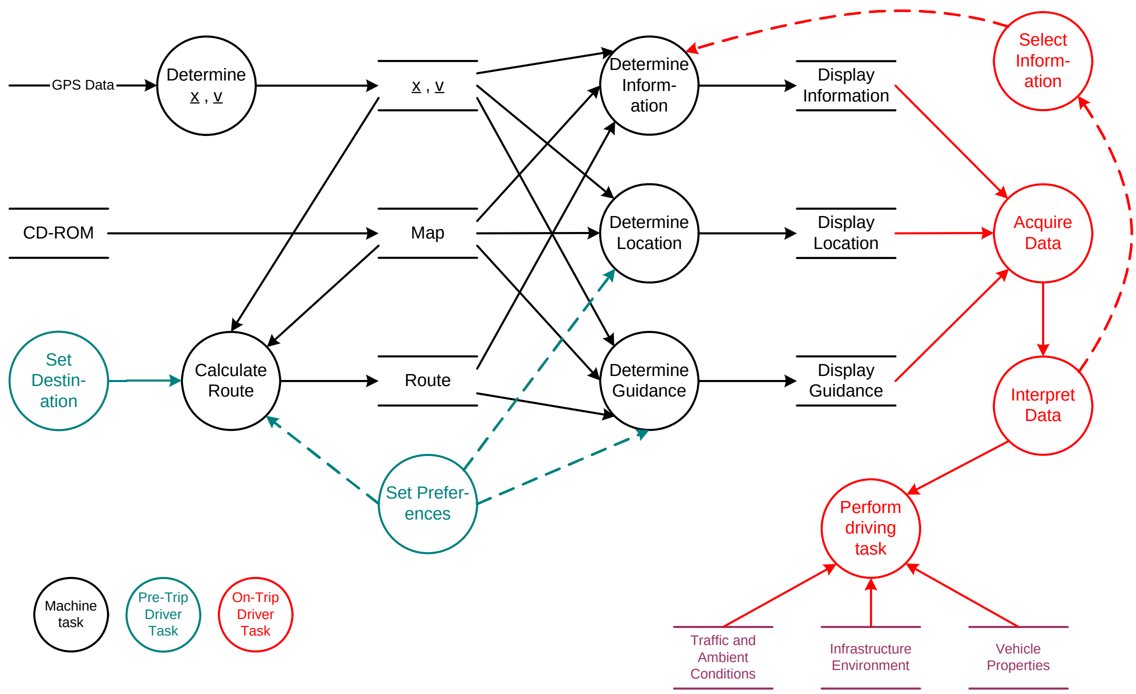


Figure 26 – Data Flow Diagram of System B

The application of the guide words resulted in the results table shown in Table 7.

Table 7 – Results from HASTE DOP applied to System B

Entity	*Attribute*	*Guide word*	*Interpretation*	*Cause*	*Consequence*	*Recommendation*
IVIS display	General image	Less	Driver doesn't (can't) see display	Ambient lighting conditions	Inappropriate driver reaction	Design – shading/contrast protection
Local map	Graphic image	Less	Not enough info	Inappropriate scale	Inadequate guidance Distraction	Ensure functionality has appropriate default
Local map	Graphic image	More	Too much info	Ditto	Ditto	Ditto
Local map	Graphic image	Less	No relevant info	Map out of date or off map	Ditto	
IVIS display	General image	More	Display too bright for ambient conditions	Backlight too bright for ambient conditions	Distraction and glare	Implement day/night or background lighting options
IVIS display	General image	Other than	Display interruption by another application	e.g. diary reminder pops up	Temporary loss of IVIS function	IVIS function should be capable of being set as the priority

Entity	Attribute	Guide word	Interpretation	Cause	Consequence	Recommendation
						application
Ditto	Ditto	Ditto	Ditto	Ditto	Requires additional interaction with interface to cancel and return	Ditto
Local map	Graphic image	Late	Map scale does not change in time	e.g. delay in GPS position update	Inadequate guidance Distraction	
Turn instruction	Auditory message	No	Driver does not receive message	Low signal level compared to ambient conditions	Driver not advised of imminent turning	1. If possible, control the volume of IVIS 2. If base device not loud enough provide additional amplification
Turn instruction	Auditory message	More	Instruction to take turn when there is no turn to take	Incorrect interpretation of mapped links	Driver could be confused and/or distracted	Ensure navigation algorithm is robust
Turn instruction	Auditory message	Less	No instruction to take turn when there is potentially a turn to take	Incorrect interpretation of mapped links	Driver could be confused and/or distracted	Ensure navigation algorithm is robust
Turn instruction	Auditory message	Other than	Message interruption by another application	e.g. diary reminder interrupts	Temporary loss of IVIS function	IVIS function should be capable of being set as the priority application
Turn instruction	Auditory message	Late	Message is not given in time	e.g. delay in GPS position update	Inadequate guidance Distraction	Use map image as backup

2.8.8 Conclusions

This analysis of potential hazards for system B was compared with the results generated by the application of the TRL checklist in WP3. This indicated that identified concerns with auditory output audibility, system response time and display size were indicated in both analysis. However the TRL checklist identified specific input functionality and display location and rigidity issues that were not specifically noted by the DOP. However it should be noted at a concept stage of development that not all such design aspects may be defined and/or known. The DOP did however identify issues that the checklist did not. These related to factors such as possible deterioration of the display due to ambient lighting conditions and IVIS function priority that the checklist did not identify. There therefore seems to be a place for both assessment tools, or a hybrid version incorporating elements of both within appropriate stages of the industrial product lifecycle.

The application of a HAZOP derived DOP to an IVIS therefore seems to be a useful additional assessment methodology to investigate the potential risks of a proposed IVIS HMI at an early stage in product development. The use of the DOP in association with a Preliminary Safety Assessment is also encouraged to develop a Safety Case for a new IVIS product. It will also assist in the identification of design issues that will need subsequent attention and re-evaluation as the design process proceeds prior to eventual assessment using a full HASTE experimental protocol.

The setting within which a DOP is performed, guided and recorded is also a relevant issue within an industrial context and further guidance on these issues are given in Deliverable 4 of HASTE. This includes guidance on the development of the various system definition stages required by the DOP, and illustrated above, and the use and application of IVIS related guidewords.

3 List of deliverables

Deliverable No	Due (month)	Output WP	Nature of Deliverable and brief description	Issue Date
D1	6	WP1	Specification of experimental protocol	July 2002
D2	23	WP2	HMI and safety related driver performance	August 2004
D3	32	WP3	Validation of the HASTE protocol specification	June 2005
D4	33	WP4	Recommended Methodology for the preliminary safety analysis of the HMI of an IVIS concept or design with supporting Case Studies	June 2005
D5	35	WP5	Brussels Workshop Proceedings	April 2005
D6	36	WP6	Final Report	June 2005

In addition to the above deliverables, the following articles have been published during the lifetime of the project.

Author(s)	Title	Journal/Conference	Date
Antilla, V. & Luoma, J.	Surrogate in-vehicle information systems and driver behaviour in an urban environment: A field study on the effects of visual and cognitive load	Transportation Research Part F: Traffic Psychology and Behaviour	2005
Carsten, O.	Implications of the first set of HASTE results on driver distraction.	Proceedings of the 14th DfT Seminar on Behavioural Research in Road Safety, pp. 100-109.	2004
Carsten, O. & Brookhuis, K.	Issues Arising from the HASTE Experiments	Transportation Research Part F: Traffic Psychology and Behaviour	2005
Carsten, O. & Brookhuis, K.	The relationship between distraction and driving performance: Towards a test regime for in-vehicle information	Transportation Research Part F: Traffic Psychology and Behaviour	2005

	systems		
Engstrom, J., Johansson, E. & Ostlund, J.	Effects of visual and cognitive load in real and simulated motorway driving, *Transportation*	Transportation Research Part F: Traffic Psychology and Behaviour	2005
Jamson, A.H. & Merat, N.	Surrogate In-Vehicle Information Systems and Driver Behaviour: Effects of Visual and Cognitive Load in Simulated Rural Driving.	Transportation Research Part F: Traffic Psychology and Behaviour	2005
Jamson, A.H. & Mouta, S. (2004)	More bang for your buck? A cross-cost simulator evaluation study.	Proceedings of the Driving Simulation Conference DSC2004. Paris.	2004
Merat, N.	Loading Drivers to their Limit: The Effect of Increasing Secondary Task on Driving.	Proceedings of the International Driving Symposium on Human Factor in Driver Assessment, Training and Vehicle Design, Park City, Utah, pp. 13-18.	2004
Merat, N., Anttila, V & Luoma, J..	Comparing the Driving Performance of Average and Older Drivers: The Effect of Surrogate In-Vehicle Information Systems.	Transportation Research Part F: Traffic Psychology and Behaviour	2005
Östlund, J.	Design of Critical Events for Assessing IVIS in Simulators - a VTI Driving Simulator Study.	Proceedings of the Driving Simulator Conference DSC 2004, pp.105-113.	2005
Santos, J.A., Merat, N., Mouta,S., Brookhuis, K.A. & de Waard, D.	The interaction between driving and in-vehicle information systems: comparison of results from laboratory, simulator and real-world studies.	Transportation Research Part F: Traffic Psychology and Behaviour	2005
Trent, V., Harbluk, J. & Engstrom, J.	Sensitivity of eye movement measures to In-vehicle task difficulty.	Transportation Research Part F: Traffic Psychology and Behaviour	2005

4 Comparison of initially planned activities and work actually accomplished.

All of the activities initially planned for the HASTE project have been achieved. The project officer to HASTE also approved a three-month extension to the project. This extension was required to manage a number of additional activities in the project, which were considered essential to the overall success of HASTE. These activities were as follows:

(a) Extensive piloting of the secondary tasks in Workpackage 1
This work was required because there were no suitable off-the-shelf tasks that could be used for the experiments in WP1 and 2. Specifically, the aim of this

work was to create tasks that placed load on one particular modality and could be manipulated in difficulty in a systematic manner.

(b) An additional study in WP2 comparing UK and Portuguese drivers
This work was done to establish whether the results from WP2 were applicable across different nationalities in Europe.

(c) A meta-analysis of the experimental results in Workpackage 2
A total of 14 driving simulator and three field studies were achieved in WP2, with 527 participants in all. This created the need for a mechanism to integrate the results, in order to draw conclusions about the types of road environments and behavioural measures that are most sensitive to changes in workload. In order to do this, one of the partners, TNO, volunteered to undertake a meta-analysis, which is a quantitative statistical procedure that yields overall estimates of effect sizes that are more accurate and reliable than that of any separate study.

(d) A Stated Preference/Conjoint analysis
This work was done using the expertise of the Economics and Behavioural Modelling group at ITS, Leeds. Whilst the experimental results in WP2 and WP3 informed us of changes in behaviour when driving and interacting with an IVIS, there was no procedure for weighting the importance of these changes. For example, is a reduction in minimum headway more unsafe than an increase in lateral deviations? To answer such questions, the scientific procedure of Stated Preference modelling was applied to a number of safety indicators. This is a novel approach and, as far as we are aware, has never been attempted in this context. The purpose of the work is to obtain importance weightings for a variety of safety indicators and to quantify how trade-offs are made between the indicators.

5 Management and co-ordination aspects

5.9 Meetings

Regular project and Workpackage meetings were arranged throughout the lifetime of the project on a regular basis, and these were always attended by all partners. All of the partners involved in HASTE were very dedicated to the project and worked hard to ensure a successful outcome to the project.

5.10 Dissemination

The work conducted in HASTE has been discussed in a variety of settings. A selection of the dissemination activities of the project are summarised below:

- A HASTE Expert Group of 15 members from different relevant organisations was formed in the beginning of 2002. These included representatives from international bodies and networking organizations (ACEA, CLEPA, FIA), national governments (UK, Sweden, The Netherlands, Finland, Spain), and independent organizations (ETSC). The HASTE project coordinator and the leader of WP5, which dealt with '*Outreach, users and dissemination*", have met the Expert group on two occasions. On these occasions, the results from HASTE were outlined and the Expert Group's views on Deliverables 1 and 2 of HASTE were sought. Furthermore, in the course of the project, experts were often consulted individually, on an ad hoc basis.

- An interactive www site was set up within a few months of the project and all public documents were made available on it immediately after they were approved by the commission. This website was also used for activities such as: distribution of information across the HASTE consortium, announcement of the HASTE final workshop and its outcome, and transfer of the Stated Preference questionnaire (see section 5(d)).
- Results of the project have also been presented at a number of relevant conferences and symposia. These include special sessions at the 2002 and 2004 ITS World Congress, and also the Human Factors Europe Chapter meeting in 2004. Numerous presentations of the HASTE project have also been given by several project participants, including meetings with partners' own national experts, ministries, and other interested parties.
- A joint Workshop with the US CAMP (Crash Avoidance Metrics Partnership) project members took place in Detroit, in 2004. Apart from exchanging views and results, this was also an opportunity to demonstrate that HASTE is a major European initiative with respect to establishing an HMI-based test regime for IVIS.
- HASTE has also been approached by several individuals and organisations that have been keen on using the S-IVIS tasks and simulator scenarios developed by the project. The consortium agreed at an early stage of the project that these should be free for others to use, as long as an official acknowledgement was made to the HASTE project.
- As outlined above, results of the HASTE project have contributed to a Special Issue of 'Transportation Research Part F', which will be out some time in 2005.
- On March 22, 2005, HASTE organised its own Final Workshop in Brussels. This was open to the public, and its aim was to present the HASTE results and its implications for a European test regime, as well as receiving feedback from a knowledgeable audience. This Workshop was considered successful, and a report outlining the main findings can be found on the HASTE website.

Details of the person to be contacted concerning the follow-up of the project are as follows:

Professor Oliver Carsten
Institute for Transport Studies
University of Leeds
Leeds LS2 9JT
UK

Tel: +44 113 343 5348
Fax: +44 113 343 5334
Email: O.M.J.Carsten@its.leeds.ac.uk

6 Results and Conclusions

The major conclusion of HASTE is that it is possible to devise an efficient and effective test regime for assessing the safety of interaction with an in-vehicle information system (IVIS) while driving. The major constituents of a **recommended test regime** have been defined. They are:

- Driving in at least a medium-level driving simulator with a relatively small number of subjects (15 subjects are thought to be sufficient)
- A rural, two-lane road, driving situation and a duration of approximately one hour
- Assessment needs to take place at the level of specific tasks on the IVIS, since an IVIS may have a combination of comparatively easier and relatively harder tasks
- A small number of dependent variables (indicators) are sufficient. At the moment, a set of 6 indicators are recommended. They are: subjective ratings, mean speed, high frequency steering, minimum headway, Percent Road Centre and PDT reaction time.

The test regime therefore meets a number of important criteria, as established early in the HASTE project:

- It is technology-independent (does not depend on any particular use of hardware or technology in system design);
- It uses safety-related criteria;
- It is cost effective;
- It is appropriate for any system design; and
- It is validated through real-world testing.

The test regime is applicable to any IVIS, including nomadic devices, and is even applicable to the use of mobile phones while driving. It is this totally generic. Once finalised, this test regime can be used by system developers (suppliers) and vehicle manufacturers both with system prototypes as a way of assisting the design process and with final products as a way of ensuring that marketed products are safe. It can also be used by consumer organisations and government authorities as a means of either (1) informing the public about the relative merits of different products or (2) ensuring that products meet minimum safety performance criteria.[7] The discussion at the end of project workshop, held in Brussels in March 2005, focussed on using the test regime as one element for scoring vehicle performance in primary safety, along the lines now being developed in Euro-NCAP.

The final, detailed specification of the test regime has not been fully defined. There are still some substantive issues to be examined in order to fully specify a test regime. Issues that require further investigation or definition are:

1. **Scoring and weighting Issues**: A final protocol should include a scoring system whereby the hazard implications of an IVIS and its

[7] The project has not defined any such minimum thresholds. Instead it has created *relative* criteria which could subsequently be used to create minimum performance thresholds.

individual tasks can be determined for particular road scenarios and driver types. Some time is required to achieve such a scoring system by combining and comparing the results of the experimental work packages with the recommendations from the meta analysis of the WP2 and WP3 results. The Stated Preference/Conjoint data collection and modelling that was undertaken as an extra element in WP2 can also help to inform the recommended scoring. The major outcome of this Stated Preference exercise will be guidance from experts on the relationship between various behavioural effects, such as changes in speed and headway, both alone and in combination with each other on safety (i.e. risk of an accident). The scoring system will help replace the task-based approach (such as that adopted in HASTE WP3) with a more generic method for IVIS rating.

2. **Test re-test reliability:** Due to time and financial limitations, only a select number of IVIS were tested in Workpackage 3. Clearly, it is important to ascertain if the HASTE protocol can be applied to any in-vehicle HMI, independent of its technology and design. It is necessary to whether the test regime produces repeatable and reliable results when it is used on the same system time and time again. This needs to be checked both within laboratory (where measurement error might be an issue even when using the same group of drivers and between laboratories. The results here would help in defining, for example, how the test group of drivers should be selected and precisely what is the minimum specification of a "qualifying" simulator. This work would therefore involve undertaking repeated measurements of the HASTE protocol on a much wider range of available IVIS, across a range of laboratories.

3. **Applying the HASTE protocol to the older driver:** Results from Workpackage 2 of the project demonstrated that older drivers are inclined to suffer from severe motion sickness when attempting to complete a visual IVIS whilst also driving the simulator. Therefore, for older drivers, studying the interaction between visual in-vehicle systems and driving is best achieved in the field. Some preliminary observations on this area have been made in the HASTE project by partners at VTT. However, due to vehicular limitations, most of this data is based on experimenter observations. A more detailed investigation that includes the same kind of parameters as collected in the simulator studies (in particular, speed, headway, steering behaviour, PRC and PDT reaction time) is therefore warranted. Such work on the older driver would ascertain whether the scoring system used by the HASTE protocol for the ‘average driver’ is applicable to the older driver. For instance, it would be useful to ascertain if the classification system described above might declare that a ‘three star system’ for average drivers would in fact only be a ‘two star system’ for elderly drivers.

This test regime can, as stated earlier, be applied once a design has become solidified so that at least a prototype is available. But it is also necessary to consider human operability of an IVIS at earlier stages in development to ensure that an acceptable product is eventually delivered to the market. Here, preliminary safety assessment procedures are intended to provide such guidance and HASTE has developed a procedure that can augment and complement such existing procedures as formal Preliminary Safety Analysis or the use of checklist such as the TRL checklist. Here HASTE has created an HMI assessment methodology called the Driver Operability Procedure (DOP), applying techniques used in hazard analysis. The methodology can be used to evaluate the impact of IVIS operation and use within the context of driving. The DOP has been evaluated in a range of real and virtual settings. The contention is that the use of this procedure would assist in developing safe and effective designs. Thus, by applying it, costly mistakes could be prevented — mistakes that would otherwise only be revealed in final testing and/or approval.

There are a number of other important conclusions from the HASTE work. One is that visual distraction and cognitive distraction from the use of IVIS have very different impacts on the primary task of driving. Visual distraction leads, not unexpectedly, to poor steering behaviour and degradation of lateral control of the vehicle. By contrast, with cognitive distraction the major negative effect is more on longitudinal control, particularly in car following, rather than on lateral control. In addition, with cognitive distraction, there was the phenomenon of an apparent ‘improvement’ in lateral control with increased cognitive task load, as shown, for example, by a reduction in the standard deviation of lateral position. The eye movement analysis, carried out in some of the studies, provides a possible explanation. With increased task load there was greater concentration of glances on the road straight ahead as opposed to the periphery, i.e. greater visual funnelling. This greater concentration of gaze and the accompanying “improvement” in lateral control has two possible explanations. One is that it is a conscious adaptation by drivers to the presence of distraction: aware of the increased risk, they focus on road ahead to maintain stable control. The other possible explanation is that the change in the concentration of gaze is autonomic and accounts for the improved tracking in that the drivers are then subconsciously aiming for the point at which they are gazing.

Another finding is that static performance on an IVIS, i.e. performance interaction with a system as a *single* task, does not reliably predict dynamic performance. Generally, the studies carried out in Workpackage 2 with the surrogate IVISs found that there was an interaction between S-IVIS performance across the baseline (static) and three levels of dynamic situation (i.e. the three levels of road difficulty). This advocates the HASTE approach of requiring the driving context to be considered in assessing an IVIS.

Finally, the studies in HASTE have confirmed the hypothesis proposed in Deliverable 1 (Roskam et al., 2002) that there would be severe problems for elderly drivers in using IVIS while driving, particularly at higher levels of task demand. "Average" drivers were not always able to manage the trade-off between primary and secondary task, and there were many indications of driving performance being poorest when the secondary task demand was the highest. But elderly drivers were particularly poor at this task management, so that there was more interference from IVIS use with their driving performance and safety, particularly in terms of higher-order aspects of driving such as managing interaction with pedestrians at crossings while subjected to cognitive load from an IVIS. This has important design and policy implications in that elderly drivers are unlikely to be able to handle even moderate load from an IVIS in more demanding road and traffic situations.

7 Acknowledgements

The following individuals were members of the HASTE team:

Anders Agnvall
Virpi Anttila
Peter Blomqvist
Anne Bolling
Karel Brookhuis
Rino Brouwer
Oliver Carsten
José Domingos Carvalhais
Dick de Waard
Johan Engström
Mark Fowkes
Jan Giezen
Joanne Harbluk
Marika Hoedemaeker
Tony Horrobin
Samantha Jamson
Hamish Jamson
Wiel Janssen
Peter Jesty
Emma Johansson
Frank Lai
Christian Larsson
Mats Lidström
Juha Luoma
Vincent Marchau
Natasha Merat
Sandra Mouta
Lena Nilsson
Joakim Östlund
Lily Read
Albert-Jan Roskam
Dennis Saluaar
Heidi Sandberg
Jorge Santos
Jan Tornros
Maria Tevell
Trent Victor
David Ward

HASTE also benefited from a panel of expert advisors:

Neil Bowerman
Martin Gelderen
Floortje Hanneman
Mike Hollingsworth
Willy Maes
Valerie Moutal
Luis Nunes
Chris Patten
Matti Roine

Valuable suggestions and feedback were also provided by the two European Commission project officers who were responsible for HASTE:

John Berry
Bipin Radia

8 References

AAM (2002). *Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems.* Version 2.0. Alliance of Automotive Manufacturers. URL: http://www.autoalliance.org/archives/driver_guidelines.pdf (2005-08-01)

Battelle Research Group (1998). Identification of human factors research needs. Publication number FHWA-RD-98-178.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Lawrence Erlbaum Associates.

Department of Transport (1994). *The design of in-vehicle information systems code of practice and design guidelines.* Revision D. Unpublished report, Transport Research Laboratory, Crowthorne.

de Waard, D. (1996). *The measurement of drivers' mental workload.* PhD thesis. University of Groningen, Traffic Research Centre.

DRIVE Safely (1992). *Towards a European Standard: The Development of Safe Road Transport Informatic Systems*. DRIVE I Project DRIVE Safely (V1051).

Dingus, T.A. et al, (1986). *Human Factors tests and evaluation of an automobile moving-map navigation system, Part 1, attentional demand requirements.* IEOR Department Report 86-03, Virginia Polytechnic Institute.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1): 65-84.

ECMT (1995). Statement of principles of good practice concerning the ergonomics and safety of in-vehicle information systems. In: *New Information Technologies in the Road Transport Sector: Policy Issues, Ergonomics and Safety,* pp. 35–42. The European Conference of Ministers of Transport.

European Commission DGXIII (1999). *Recommendation of the Commission of the European Communities to the Member States and Industry on safe and efficient in-vehicle information and communication systems: a European statement of principles on human machine interface.* European Commission, Brussels.

European Commission DG Information Society (2005). *European statement of principles on the design of human machine interaction (ESoP 2005), Draft*. European Commission, Directorate-General Information Society and Media. June.

European Transport Safety Council (1999). *Intelligent transportation systems and road safety.* Brussels.

Farber, E. (2000). Visual attention design limits for ITS in-vehicle systems: The Society of Automotive Engineers Standard (SAE) for limiting visual distraction while driving. TRB 2000.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: P.A. Hancock & N. Meshkati (eds.), *Human Mental Workload* (p. 139-183). Amsterdam: North Holland.

Hobley, K.M. et al. (1995). *Framework for Prospective System Safety Analysis Volume 1 — Preliminary Safety Analysis*. Deliverable N° 9a, DRIVE II project PASSPORT.

IEC 61508. *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems*. International Electrotechnical Commission, 1998–2000.

JAMA (2004). *Guideline for in-vehicle display systems.* Version 3.0. Japanese Automobile Manufacturers Association. URL: http://www.jama.or.jp/safe/guideline/pdf/jama_guideline_v30_en.pdf.

Johansson, E., Carsten, O., Janssen, W., Jamson, S., Jamson, H., Merat, N. Östlund, J., Brouwer, R., Mouta, S., Harbluk, J., Anttila, V., Sandberg, H., J. Luoma & Victor, T. (2005). *Validation of the HASTE protocol specification.* (HASTE Project Deliverable 3): European Commission.

MISRA (1994). *Development Guidelines for Vehicle Based Software.* Motor Industry Research Association, Nuneaton, UK.

Morton, A.H. & Angel, H.A. (2005). *Evaluation of industry safety principles for in-vehicle information and communication systems.* Final Report. Humansystems Incorporated, Guelph, Ontario, Canada.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H., Jamson, S., Mouta, S., Carvalhais, J., Santos, J., V. Anttila, H., Sandberg, Luoma, J., de Waard, D., Brookhuis, K., Johansson, E., Engström, J., Victor, T., Harbluk, J., Janssen, W., & Brouwer, R. (2004). *HMI and Safety-Related Driver Performance* (HASTE Project Deliverable 2): European Commission

Risser, R. Ed. (1997). *Assessing the Driver*. Contribution to a workshop on Driver Diagnostics and Selection, March 17, 1997, Basle. Der Mensch im Verkehr. Braunschweig: Rot-Gelb-Grün.

Roskam, A.J., Brookhuis, K., de Waard, D., Carsten, O., Read, L. Jamson, S. Östlund, J., Nilsson, L., Bolling, A., Anttila, V., Hoedemaeker, M., Janssen, W. H. Harbluk, J., Johansson, E., Tevell, M., Fowkes, M. Engstrom, J. & Victor, T. (2002). *Development of an experimental protocol.* (HASTE Deliverable 1): European Commission.

Ross, T., Vaughan G., Engert, A., Peters, H., Burnett, G. & May, A. (1995). *Human factors design guidelines for information presentation by route guidance and navigation systems.* Deliverable 19 of Project HARDIE. HUSAT, Loughborough.

Ross, T. et al. (1996). *HARDIE Design Guidelines Handbook.* Deliverable 20 of project HARDIE. HUSAT, Loughborough, UK.

Smiley, A. & Brookhuis, K.A. (1987). Alcohol, drugs and traffic safety. In J.A. Rothengatter & R.A. de Bruin (Eds.), *Road users and traffic safety* (pp. 83-105). Assen, The Netherlands: Van Gorcum.

Stevens, A., Board, A., Allen, P. & Quimby, A. (1999). *A safety checklist for the assessment of in-vehicle information systems: A user's manual.* Project Report PA3536-A/99. Transport Research Laboratory.

Treisman, A. (1988). Features and objects: The fourteenth annual Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 40*, 201-237.

Veltman, J.A. & Gaillard, A.W.K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics, 41*(5), 656-669.

Wierwille, W.W. (1988). Strategic use of visual resources by the driver while navigating with an in-car navigation display system. SAE Paper, 885180, Society of Automotive Engineers.

Wirtschaftsforum Verkehrstelematik (1996). *Vereinbarung zu Leitlinien für die Gestaltung und Installation von Informations- und Kommunikationssystemen in Kraftfahrzeugen.* (English translation: Steering Group on the Economic Forum on Telematics in Transport. Agreement on guidelines for the design and installation of information and communication systems in motor vehicles.) Bonn, Germany.

Zijlstra, F.R.H. & Van Doorn, L. (1985). *The construction of a scale to measure perceived effort*. Delft, The Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.

Zwahlen, H.T., Adams, C.C. and Schwartz, P.J. (1988). Safety aspects of cellular phones in automobiles. 18^{th} International Symposium on Automotive Technology and Automation (Paper 88058).