Automation of the driving task

Final report

R-98-9
Tom Heijer, SWOV Institute for Road Safety Research
Karel Brookhuis, Centre for Environmental and Traffic Psychology
Wim van Winsum, TNO Human Factors Research Institute
Lies Duynstee, Transport Research Centre of the Dutch Ministry of Transport and Public Works
Leidschendam, 1998
SWOV Institute for Road Safety Research, The Netherlands

Report documentation

Number: R-98-9
Title: Automation of the driving task
Subtitle: Final report
Author(s): Tom Heijer, SWOV Institute for Road Safety Research
Karel Brookhuis, Centre for Environmental and Traffic Psychology
Wim van Winsum, TNO Human Factors Research Institute
Lies Duynstee, Transport Research Centre of the Dutch Ministry of Transport and Public Works
Research manager: Siem Oppe
Project number SWOV: 54.512
Project code client: HVVL 97.501
Client: This research was funded by the Dutch Ministry of Transport and Public Works.

Keywords: Driver information, safety, traffic, driving (veh), telecommunication, data processing, simulation, stress (psychol), evaluation (assessment), design (overall design), psychology.

Contents of the project: The fast development of all sorts of telematic devices to support or partially substitute a drivers tasks has also led to some concern for the possible detrimental effects that such devices may have on the safety of driving. This report summarises the final results of a three year project aimed at the development of criteria to assess the effects on road safety of various applications of Advanced Traffic Telematics (ATT systems) intended to support the driver in different aspects of the driving task.

Number of pages: 66 pp.
Price: Dfl. 25,-
Published by: SWOV, Leidschendam, 1998

SWOV Institute for Road Safety Research
P.O. Box 1090
2260 BB Leidschendam
The Netherlands
Telephone 31703209323
Telefax 31703201261

Summary

The fast development of all sorts of telematic devices to support or partially substitute a drivers tasks has also led to some concern for the possible detrimental effects that such devices may have on the safety of driving.
A three year research project has been carried out to develop criteria and procedures that can be used to assess the nature and extent of these detrimental effects. A further aim was, to make the assessment procedure as simple as possible, so that it can be used by relatively non-expert users. Moreover, these procedures aim to limit the number of otherwise necessary (expensive) field tests.

A general conclusion is, that current knowledge on this subject is not yet sufficient to provide a comprehensive set of checks and this has resulted in the following compromise for the testing procedure:

- A checklist for safety characteristics of telematic devices is proposed based upon known safety effects of task load changes: overload and underload: this checklist can be employed by non-scientists to provide a first screening of unsafe characteristics.
- A laboratory test has been developed that uses an ordinary Personal Computer. The test emulates a simplified driving task and can accommodate a functional simulacrum of a telematic device. The user is also provided with a set of criteria to produce an assessment of the safety effects. This test can also be used by non-experts.

Furthermore, recognising that in the current state of affairs field testing will still often be necessary, an attempt was made to formulate guidelines and criteria for the setup of these tests.

Contents

1. General introduction

This report summarises the final results of a three year project commissioned by the Transport Research Centre of the Dutch Ministry of Transport and Public Works (TRC) to the following research institutes:

- TNO Human Factors Research Institute (TNO-HFRI);
- Centre for Environmental and Traffic Psychology (COV);
- SWOV Institute for Road Safety Research (main contractor for this project).

The project is aimed at the development of criteria to assess the effects on road safety of various applications of Advanced Traffic Telematics (ATT systems) intended to support the driver in different aspects of the driving task. Such ATT systems are being developed (or are already on the market) e.g. to provide up to date route information, to maintain a constant speed and headway, to adapt the maximum speed to the local limit or to prevent collisions.

Although many of those support systems are intended to make driving easier or safer, they can also interfere with or modify the driver's tasks in such a way that safety is impaired. This leads to the conclusion that acceptability of ATT systems should be determined by a careful consideration of both the intended beneficial and the unintended detrimental effects on safety before any application is given a 'green light' by the government. Preferably, such a consideration should be conducted by way of standardised procedures and criteria, but these do not yet exist. This project has been initiated to provide at least a preliminary set of guidelines and methods to identify potential safety hazards that single or multiple applications of these ATT systems may produce.

2. Setup of the project

The project was initiated to develop an efficient procedure to assess the acceptability (for reasons of safety) of marketable in-car telematic devices. Efficient in this respect means, that as much as possible of the assessment procedures can be carried out quickly and cheaply behind the desk and that more expensive and time consuming field testing will only be used to resolve remaining questions or doubts.
Ideally, the end result of this project would be a comprehensive checklist that can be applied to all sorts of features of ATT applications and that will immediately render a verdict in terms of safe or unsafe. Furthermore, this checklist must be formulated in such a way that it can be applied by (relative) non-experts like policy makers. In this way, the participation of expensive experts or researchers could be avoided and the evaluation can be made very quickly.
Acknowledging that such a procedure is, as yet, unfeasible, we have attempted to develop a procedure in which the work of evaluating safety effects can be shared between policy makers and experts/researchers in a more or less efficient way. This procedure then consists of the following steps:

1. Preliminary screening of a ATT device by a non-expert with the aid of a checklist: this checklist will only rarely result in an absolute verdict but will mostly generate points of attention.
2. Possible extension of the screening by non experts, employing a simple laboratory test that uses a desktop PC and a limited amount of additional equipment: this form of testing can be applied to certain points of attention generated by the checklist. Such a test will, for the time being, only apply to a limited number of types of ATT applications.
3. If the previous two steps do not produce a conclusive result because the attention points that remain unresolved seem too severe or numerous, a field test will be indicated to render a final verdict. This test, or rather a test series, is set up around well documented testing procedures and uses well defined testing parameters, criteria and assessment protocols.

These steps have been worked out in a research programme that was executed in the past three years. The project itself contained the following phases:

- An inventory of existing knowledge on the subject by means of a literature study.
- Execution of preliminary experiments on well known ATT systems.
  The results of these first two phases haven been described in chapter 4 of this report.
- The formulation of practical criteria and procedures necessary to well based experimental safety tests: chapter 5 contains the results of this part of the work.
- The first formulation of an evaluation method suitable for use outside a laboratory consisting of :
- a preliminary checklist described in chapter 7;
- a limited experimental testing method suitable for a desktop PC as described in chapter 6.
- An evaluation of the PC test using the results of the preliminary experiments; the results of this part can also be found in chapter 6.

Finally, in chapter 8 the conclusions per phase are summarised as well as the general conclusions and recommendations.
This programme was not the only research concerning safety effects of ATT under way during this three year period: in parallel other, more specific projects have been carried out by some of these same research institutes. The results of one of these projects, the IVIS project primarily aimed at RDS-TMC, have also contributed to this project, specifically to the criteria for testing, the checklist and the PC test.

3. Results of the literature study and experiments

3.1. The literature study by TNO-HFRI and COV

The discussion in this paragraph is based on the report *Safety effects of in-vehicle information systems* of Verwey, Brookhuis & Janssen (1996).

3.1.1. *Introduction*

A growing number of traffic safety studies shows that human error is a major contributing factor in traffic accidents. For instance, Treat et al. (1977) showed that difficulties with perception, attention, distraction, etc. are important causes in 30 to 50 percent of traffic accidents. Countermeasures have to be devised and introduced to prevent behaviour contributing to these accidents, without eliciting undesirable side-effects. Modern electronic systems in- and outside the car, also indicated by telematics applications, could be such countermeasures but their usage may be accompanied by side-effects such as distraction, overload, insufficient attention for the driving task, and counterproductive adaptation as a consequence of (misleading) feelings of safety evoked by these measures. So, an important question is how can one decide whether a specific telematics application affects traffic safety.

A literature study was performed to delineate how a standardised test methodology for assessing safety effects of in-vehicle information systems, also indicated by telematics applications, should be designed. Since it is virtually impossible to cover all safety aspects that telematics applications could have in the broadest sense, a limitation to the effect of applications on the driving task in a narrow sense has been chosen in this report. An analysis has been carried out of how the use of in-vehicle telematics applications could affect driving performance of individual drivers.

A major reason to develop driver support systems is the reduction of traffic accidents. One of the problems with these systems is that it is very difficult, if not impossible, to forecast the savings of death and disability that might result from the introduction of such systems. Although there is an urgent need to know what the effects are of introducing a specific system before it enters the market, no data exist on which estimates of the risks caused by these systems can be based. However, the effect of individual systems *can* be studied on a low level, i.e. safety aspects of the driving task per se. Hence, each individual telematics application should be subjected to a test for behavioural safety effects before marketing in order to pinpoint unwanted side-effects at the behavioural level. However, in order to determine whether a system has unacceptable side-effects, criteria must be developed for what exactly constitutes unacceptable.

Even though the need for a general, preferably standardised methodology for assessing safety effects of telematics applications has been generally recognised, there have only been a few indecisive attempts to come to a methodology at a level on which empirical studies can be easily based. Knowledge about how actual safety effects should be assessed is only beginning to emerge from a handful of DRIVE II projects. One of the reasons is that a lot of empirical research has been devoted to the comparison of different versions of a group of telematics applications in order to improve the human-machine interface rather than to absolute safety effects. However, at some point in time it still has to be clarified whether a specific telematics application should be allowed on the market or not and the project reported here was specifically aimed to provide instruments for such a clarification. To that end, the study focused on the safety effects of three mechanisms: workload and distraction, reduced attention, and counter-productive behavioural adaptation which will now be treated briefly.

3.1.2. *Overload*

In driving, overload refers to the situation that the driver is unable to process all relevant information for executing the driving task. This results in increased error rates and later detection of other traffic participants and, hence, to reduced traffic safety (e.g., Rumar, 1990). The role of overload in accident causation is supported by studies showing that high levels of workload on specific routes were associated with the probability of traffic accidents on those routes (MacDonald, 1979; Taylor, 1964) and the finding that drivers with less information processing capacity are more likely to have accidents (Lim & Dewar, 1988).

Whereas overload may occur while no telematics application is being used, the introduction of such applications makes overload more likely because interaction with these applications involves a task additional to normal driving.

A common model of the driving task assumes that driving tasks can be categorised into tasks at three levels (Allen, Lunenfeld & Alexander, 1971; King & Lunenfeld, 1971; Michon, 1985). The first level, the *control level*, is concerned with elementary vehicle handling functions like lane-keeping and handling of the controls which allows one to follow the road and to keep the vehicle on the road. Time constants at this level are usually below one second and the tasks at this level usually cause little mental workload.
The *manoeuvring level* deals with reactions to events in the traffic environment. These reactions have to do with interactions with other traffic like overtaking, intersection negotiation and the like. Time constants are normally between one and ten seconds and mental workload is usually higher than that associated with the control level. The *strategic level* regards choice of transport modality, route planning, and route following. That is, how drivers choose their destination, the route and the modality of travel. Time constants related to the processes at the strategic level are typically more than ten seconds and workload is generally high.
In general, mental overload may be caused by tasks at the manoeuvring and strategic level, especially for tasks with low time constants. Visual workload may be high for control tasks too. So, if one considers the occurrence of overload caused by telematics applications, the type of task that is likely to be carried out and the type of workload it incurs may be important to consider before workload is actually measured.

The major methodologies for assessing driver workload are summarised with an emphasis on the techniques sensitive to various types of workload at the same time.

3.1.2.1. *The secondary task technique*

The *secondary task technique* involves the addition of a low priority task to the primary - here, driving - task. It is assumed that an increase in workload associated with the primary task is indicated by performance reduction of the secondary task. Depending on the type of secondary task, the load on separate resources can be assessed. Applying a task which has many common characteristics with driving would provide a general indication for driver workload. With the utilisation of this technique, Verwey (1993a,b) could show in which driving situations overload is more probable. However, the secondary task method seems less suited for performing safety assessments of telematics applications because it is likely to affect driving performance in any case (e.g., Noy, 1987). This makes it hard to directly assess safety effects from driving performance.

3.1.2.2. *Physiological measures*

A number of *physiological measures* is also used for assessing workload. Some measures are primarily sensitive to load on specific processing resources whereas others index overall mental workload. The most common ones in driving research are various measures based on heart rate. Less common nowadays are responses in skin conductance (SCR) and eye blink frequency.

The advantage of physiological measures is that they affect the driving task only in minor ways and, once the electrodes are attached, are fairly unobtrusive. Their disadvantage is that they usually have a limited temporal resolution (Verwey & Veltman, 1995). Therefore, these measures appear primarily useful when overload occurs for periods of at least a few minutes. Another disadvantage is that analysis of physiological data is fairly labourious because these data are often contaminated by other physiological signals and noise. One promising index of workload for assessing effects of telematics applications on driver workload is eye movement registration. Eye movements can be registered by way of electrodes located round the eye, which is why eye movements are often categorised under physiological measures, but the more common procedure nowadays is to use external registration by way of one or more video cameras.

A number of studies have measured eye movements in terms of glance frequency (the number of glances toward a display) and glance duration (the time the driver looks at a display). Wierwille (1993) found that glance duration towards an in-car display will generally not exceed about 1.5 s. If more time is required glance frequency increases. Glance duration would be associated with the time required for chunking parts of the display (e.g. reading words) while glance frequency would indicate overall complexity of the display. Zwahlen et al. (1988) state that more than three glance times of 1 s each are unacceptable. These studies give a general idea of what is acceptable in terms of glance times and frequencies.

However, it remains hard to conclude any thing about safety if it is not clear when, in which situations, drivers will actually look at the display. For example, glances of only half a second may be fatal in complex driving situations whereas much longer glance times are harmless on straight roads. Therefore, safety evaluations should also consider when telematics applications increase workload and whether the increase will actually affect safety. For example, future work might focus on the possibility to compare recorded glances times toward in-vehicle displays with glance times that are theoretically allowed (given current headway, lane position etc.).

A method frequently applied to assess driver workload is simply asking the participants in a study how loaded they are (*subjective measures*). Relatively simple rating lists have initially been used which involved only a single rating, that is a one-dimensional list. In the last decade the NASA-TLX (Hart & Staveland, 1988) and the SWAT (e.g., Reid & Nygren, 1988) have been used more frequently. These techniques involve several subscales and would give more accurate indications for workload than univariate lists.
A disadvantage of subjective ratings is that they involve an additional task and, as such, are less reliable when workload exceeds working memory limitations (Yeh & Wickens, 1988). This pleads for the use of more simple one-dimensional lists. In fact, recent studies found that one-dimensional lists give as reliable results as the more complex SWAT and TLX (Hendy, Hamilton & Landry, 1993; Verwey & Veltman, 1995). Care should be taken that it is always clear to the participants to what periods the estimates refer and whether the estimates should refer to peaks or to averages (Verwey & Veltman, 1995). Another disadvantage of subjective lists is that subjective scores may be affected by opinions on system characteristics. With respect to safety evaluations it should be noted that subjective rating lists are probably sensitive to workload in general rather than to specific aspects (resources) of workload (Verwey & Veltman, 1995)

3.1.2.3. *Performance on some driving subtasks*

Performance on some driving subtasks can be regarded as indicators for driver workload because these have no direct implications for safety (e.g., steering wheel frequency) whereas others are more directly associated with safety (e.g., headway). As drivers are usually able to distinguish the safety of various subtasks, they will give safety related tasks a higher priority at times of elevated workload. Hence, high priority aspects of driving are less likely to be affected by increased levels of workload than lower priority tasks. Consequently, the effect of interacting with a telematics application will affect low priority subtasks earlier and more often than high priority subtasks.
For example, Verwey (1991a) found that glances at the interior mirror reduced with increased task load. Obviously, this has no direct consequences for safety but it might still affect safety in that the driver is less aware of what is going on around the car. This leads to the inference that reduced performance of high priority parts of the driving task are clear indications that safety is affected but this will occur rather infrequently, whereas deterioration of low priority task performance is more likely and, hence, a reasonable indicator for driver overload, but a worse indicator for reduced levels of safety.
With respect to evaluating safety effects of telematics applications, Zaidel (1991) proposed to assess the quality of driving in a real driving study in which an expert observer, most likely a driving instructor, gives detailed judgements on the quality of driving with and without the telematics application.
Such a method is presented in De Gier (1980). However, objections against this method have been raised. The major problem is that subjective ratings are sensitive to the individual rater's opinions and to whether or not the rater is able to register all relevant information.

As there are many ways to assess (driver) workload, it is not always obvious which ones should be used in a specific experiment. There appear to be a few criteria. First, the technique should be sensitive to relatively short and acute changes in workload. Second, the measures should not affect driving behaviour as, in fact, aspects of this behaviour should say also something about safety effects. Third, the measure should be sensitive to general workload and show overload when only one processing resource is overloaded. Fourth, the method should preferably not be too labourious.

Given the large variety of techniques available for assessing workload and driving performance, it is not at all obvious how actual safety effects of telematics applications due to information overload can be assessed.
The relation between driving performance, workload and safety is a complex one because important parts of driving will usually not suffer and are, therefore, not sensitive indicators for reduced safety. Still, in practice they may sometimes be affected and cause accidents anyway. If it can be shown that driver workload does not increase at all when one is using a telematics application one may infer that safety will not be affected by driver overload. But if there is an increase in driver workload it is not clear what level of increase can be considered acceptable.
The view that workload may not increase at all is too simple because many telematics applications may increase workload in some situations and reduce workload in other situations. Also, a minor increase needs not affect traffic safety, especially when the telematics application does not present information that is distracting and if message presentation happens in situations in which the driver is not heavily loaded. In order to assess safety effects, one might analyse to what extent anticipatory behaviour (including looking out) is affected. Before workload is assessed, it is important to determine in which driving situations and during what type of tasks, interaction with telematics applications is likely to occur so that one can hypothesise which type of overload can be expected at which moments in time. There is a need to cross-validate popular objective workload and driving performance measures with safety assessments of experts in order to determine which objective measures are most sensitive to safety effects and what their absolute threshold values are from a safety point of view. To find out how reliable subjective rater opinions are, inter-rater reliability (i.e., the consensus of various raters) should be assessed too.

3.1.3. *Underload*

Driver underload is defined as indicating the situation that the driver gets into a state of limited attention to driving, due to either diverted attention (e.g., no specific driving task demands) or deactivation (e.g., the driver dozes off). Driver state is not some unitary, fixed phenomenon, not even within an individual. It varies with time-of-day, age, subjective feelings (mood), but also with time-on-task and all kinds of external influences, such as traffic environment and situational task-load, alcohol and (medicinal) drugs.
Certain telematics applications are developed to support the driver, for instance, by taking over parts of the driving task whereafter less attention to aspects of the driving task is strictly needed. One approach in the field of safety assessment is locating the optimum with respect to driving performance, and then detecting signs of deviations from the optimum (cf. Wiener et al., 1984), analogous to the concept of arousal, as expressed with the inverted-U-hypothesis. This hypothesis states that there is a level of arousal, or activation, that yields the highest level of performance. Ideally, the driver acts on top of this performance curve (cf. Wiener, 1987), gaining maximum results. Deviations from the top (Wiener's optimum) could be

caused by a decrease in arousal according to the inverted-U-hypothesis or diminished workload, or task demand, according to Wiener's analogue. The concomitant phenomena can be registered better from the individual's physiology, being the precursor of behavioural decrements, than from the relevant behavioural parameters themselves.

Changes in driver state are reflected in changes in relevant physiological parameters such as electroencephalogram (EEG), electrocardiogram (ECG), galvanic skin response (GSR), Electro myogram (EMG) (see Brookhuis & De Waard, 1993). In order to estimate the contribution of the changes in driver state to accident causation, one would want to be able to continuously measure these physiological signals in real life driving. However, it is highly unlikely that measuring physiological parameters while driving a motor-vehicle will be acceptable ever. Therefore, measuring physiological parameters of underload can only be used for assessing safety effects of various telematics applications in experimental conditions.

After prolonged driving, a driver's activation tends to diminish rapidly, such as can be measured by means of spectral analysis of the EEG. Alcohol and medicinal drugs mostly aggravate these effects (Brookhuis et al., 1986) although this is not necessarily the case, depending on dose and type of drug. Petit et al. (1990) demonstrated a relationship between the handling of the steering wheel and the occurrence of alpha waves in the EEG. The state of vigilance, as indicated by the power in the alpha range, was found to correlate highly with specially developed steering wheel functions in most cases.

De Waard & Brookhuis (1991) also found effects on steering wheel behaviour with time-on-task (150 minutes of continuous driving). The standard deviation of the steering wheel movements increased and the number of steering wheel reversals per minute decreased, both highly significant. Subjects' activation, measured by a relative energy parameter [(alpha+theta)/beta], gradually diminished from the start of the experiment.

With respect to the assessment of safety effects of telematics applications, it is important that underload may well be caused by the presence of telematics applications that take over (part of) the driving task. Until telematics applications are capable of reducing the driving task fully to a mere supervisory task, introduction of telematics applications taking over part of the driving task must be considered very carefully and maybe even reluctantly.

In contrast to the matter of overload, where the relationship between objective measures of driving performance and safety is yet largely unresolved, the DRIVE I V1004 project (DREAM) cross-validated physiological and performance data (Thomas et al., 1989). Because safety effects of underload are fairly clearly related to performance data there is less need for a further cross-validation with subjective opinions on safety. Safety effects of underload are largely confined to reduced lane keeping performance and increased reaction times to events in the traffic environment and as such fairly easy to measure.

On the basis of a literature study and experimental research De Waard & Brookhuis (1991) concluded that the combined measurement of the driver's physiology and behaviour should allow the development of a monitoring device on the basis of unobtrusive vehicle parameters alone. The driver actions, as measured from vehicle parameters, that seemed most promising were aspects of steering wheel movements, variability in lateral position, time-to-line crossing, reactions to lead vehicles, as measured by reaction time and time-to-collision, and perhaps speed management. Furthermore, subjective measures of, for instance, psychological impairment, effort and acceptance should be taken into account.

Similar to the problem that high workload and driving performance measures have not been linked to safety measures directly yet, there is no clear relationship between safety and measures of underload in all respects. Although for some driving parameters a first attempt has been made to establish safety criteria, a complete picture of the relationship is still far away. In fact, studies in various DRIVE projects have enabled us to relate a few driving parameters to physiological measures and based on that a limited series of absolute and relative criteria have been proposed. These criteria facilitate the evaluation of underload effects introduced by telematics applications. However, based on driving performance alone it is hard to predict when and where things go wrong in the absolute sense.

3.1.4. *Counterproductive behaviour adaptation*

Counterproductive behavioural adaptation is the phenomenon that drivers start behaving in riskier ways exactly because they are supported by a safety-raising device. It may thus manifest itself, first of all, in the behavioural parameters that are typically used in driver behaviour studies. In a wider sense counterproductive adaptation may also be taken to comprise drops in attention, or a reduced level of general alertness that is induced because the device acts as a guardian angel. In a still more general sense counterproductive adaptation may occur at the strategic level of the driving task, that is, with respect to conditions under which trips are undertaken at all or with respect to changes in overall mobility (mileage).

The effect of counterproductive adaptation, when it occurs, is to make the net safety benefit less than would have been expected on the basis of the effect of the device by itself. As a rule of thumb counterproductive adaptation can be expected to occur in the fastest possible way for devices that are clearly intended to raise safety, that are conspicuously present within the vehicle, and that act relatively frequently, i.e., under conditions that are not necessarily highly critical. It will take longer for drivers to compensate, if at all, for safety features that are not immediately making their presence clear, since the feedback loop here is more indirect and diffuse.

A problem that is typical for counterproductive adaptation effects is how to assess what the net effect of a (safety-raising) device is, that is, how a counterproductive adaptation effect should be 'subtracted' from the device's original-beneficial-effect. The latter, the so-called 'engineering estimate' of a device's expected safety effect, is the accident reduction that would be achieved if 100 percent of the relevant population had the device and if that population showed no behavioural adaptation to the new situation (Janssen & Van der Horst, 1992).

The basic notion in making an engineering estimate is that expected safety benefits are given as an extrapolation or an implication of a rather straightforward engineering calculation. By doing so, physical changes to the system are considered without initially addressing possible induced user change. For example, if design changes to some roadside device would be calculated by engineering methods to reduce the probability of a driver death on impact by 10 percent, then the engineering estimate is that a 10 percent reduction in driver deaths from collisions with the modified device will occur. The most common way to obtain an engineering estimate is indeed from accident data. A device that, according to accident statistics, causes x percent of deaths is expected to yield a safety return of x percent upon its removal. Alternatively, if the absence of a device would lead to y percent of all deaths, then the implementation of that device would be expected to reduce deaths by those same y percent.

In other cases an engineering estimate can be made on the basis of laboratory results, e.g., for hardware devices tested under crash conditions representative of those occurring in reality. In still other cases the engineering estimate can be no more than the expectation of a beneficial safety effect, or an order of magnitude thereof. The estimation of (net) safety effects can never be better than the engineering estimate permits. That is, each and every safety measure needs an estimate of its effect per se when its implementation is being considered, and against which the effect that is ultimately realised must be evaluated. It should be considered to be the obligation of those proposing a device's implementation to support the safety claim that is being made by a hard engineering estimate originating from either accident or laboratory studies. In case the device is not directly safety-directed, but aims to support other driver functions (like navigating), the engineering estimate is naturally close to 0.

Another factor that should be taken into account when assessing the net safety benefit of any device is the device's use rate within the population. For safety-directed measures which for their effectiveness rely on the acceptance of the population there is the complicating and complex issue of selective recruitment, meaning that the use rate per se as well as the effect that is achieved are affected by self-selective processes in the population. The hypothesis is that those who opt for some device differ from those who do not in respects that are essential to its effectiveness. Useful quantitative expressions describing the implications of self-selective processes for driving behaviour as well as for resultant accident involvement rates have been derived by Evans (1987a,b).

It should be noted that counterproductive adaptation may also manifest itself in ways that are not restricted to driving behaviour per se. There are several ways in which this may occur. The availability of a device that reduces risk per kilometre driven may thus lead to any or all of the following:

1. The participation in traffic of segments of the general population that did formerly not dare to do so because they considered the risks on the road too high.
2. An increase in mileage driven (overall Vehicle Miles Travelled), both because of:
   - a shift in modal split (from other modes of transport to the automobile);
   - a direct increase per vehicle.
3. An increase in mileage driven under more unfavourable conditions and/or under decreased levels of personal fitness.

All these effects will have as a consequence that the exposure to risk will increase, which by itself will generate more accidents even when the risk per kilometre has been reduced.

It is presently not possible, given the state of traffic safety science, to anticipate upon the occurrence of these effects in a quantitative form. It is, nevertheless, unwise not to consider the possibility that this type of consequences could occur. The rule of thumb must be that, if the average automobile driver can imagine a way of getting a mobility profit out of a safety-raising device, the average experimenter should surely be capable of doing so and must admit that this can be a consequence.

3.2. Experimental results by TNO-TM

The discussion in this paragraph is based on the following reports:

- Verwey, W. (1996). *Evaluating safety effects of in-vehicle information systems; A detailed research proposal.*
- Verwey, W. (1996). *Evaluating safety effects of in-vehicle information systems; Testing the method.*
- Verwey, W. (1996). *Evaluating safety effects of in-vehicle information systems (IVIS); A field experiment with traffic congestion information systems (RDS-TMC) and preliminary guidelines for IVIS.*

3.2.1. *Introduction*

In recent years, there has been a considerable boost of research and development of modern technology in road transport. From the early start on, many people have expressed their concern that this technology, known as in-vehicle information systems (IVIS), advanced transport telematics (ATT) or intelligent transport systems (ITS), might jeopardise traffic safety rather than that it would improve safety as is claimed by others (e.g., Hancock & Parasuraman, 1992; Parkes & Ross, 1991).
One reason for this concern is the distraction and overload the driver may be confronted with. An ordinary driver may well be able to perform additional tasks while driving on a quiet motorway. However, in dense city traffic and while negotiating complex intersections behaviour may become unsafe. With attention attracting information in demanding driving situations, there is the danger that drivers are not able to ignore the message entirely. When messages are not conspicuous, drivers may choose to attend to the information because they think they can handle it, which is not necessarily the case. Finally, even in quiet driving situations, drivers may take risks by giving too much attention to the IVIS.
Many studies have shown effects of IVIS on driving performance. These effects suggest that safety is affected as well. However, given the difficulties to assess safety in experimental situations in a reliable way, there is still limited proof for negative safety effects. There is an urgent need to develop guidelines and standards for the design of the man-machine interface of IVIS based on safety research.

The experiment aimed at testing safety effects of three major types of driver-system interaction with a specific IVIS, that is a system giving on-line traffic information, and relating the characteristics of human-machine interface to these safety effects. The results of the study and guidelines from the literature are used to propose preliminary guidelines for the driver-vehicle interaction with IVIS systems and to propose a framework for the development of guidelines aimed at preventing negative safety effects of the workload and distraction caused by use of IVIS.

3.2.2. *Method*

Twelve experienced drivers drove a route through the city of Amersfoort. At a set of predetermined driving situations, which included right turns, intersections and straight driving, congestion information was presented by means of either a map display or a speech message, or subjects were instructed to program a filter on an RDS-TMC system. This filter determines which particular subset of all available RDS-TMC information will be passed to the user (e.g. the information for a particular road). The task involves quite elaborate manipulation of buttons on the receiver and reading a display.
Driving performance and looking behaviour in these situations were analysed in terms of their safety ramifications by comparison with absolute safety limits obtained from the literature (Verwey, 1996), safety judgments by an experienced driving instructor, and comparison with a control condition.

3.2.3. *Major results*

The driving instructor opinions showed that, across all situations and types of judgments, driving safety was significantly impaired by each of the three IVIS tasks. In the filter programming task these effects were in large part caused by poor course keeping and braking/decelerating (mainly looking in advance). More detailed analyses showed that the safety reductions concerned looking, course keeping, and braking (mainly anticipation) when turning right, course keeping when approaching general rule intersections, braking (mainly anticipation) when approaching priority intersections, and course keeping when driving straight. Steering wheel frequency increased at the straight urban sections with filter programming but not with map and speech.

Performing the IVIS tasks in this study did not affect the ratings with respect to adapting speed to other traffic (straight driving), distance to heading traffic (straight ahead), anticipation in general (straight ahead, general rule and yield intersections), braking and deceleration (straight ahead), giving priority (general rule and yield intersections), watching priority traffic (general rule and yield intersections), and course keeping (yield intersections). Neither were any effects found on the objective measures looking behaviour as scored from video (right turns, general rule intersections), the hypothetical occurrences of high decelerations (right turns, yield intersections), exceeding the critical minimum TTI during each approach of an intersection, steering frequency (map and speech conditions), the proportion high decelerations (all situations), and the standard deviation of speed and TLC (straight driving).

3.3. Experimental results by COV

The discussion in this paragraph is based on the report:

- Brookhuis, K.A., Waard, D. de (1996). *Limiting speed through telematics; Towards an Intelligent Speed Adaptor (ISA).*

3.3.1. *Introduction*

Through recent legislation in the Netherlands, the maximum driving speed is restricted by a speed limiter in the heavier types of lorries and coaches.
The effect of these devices on fuel consumption, noise, air pollution, wearing of the tires and traffic safety is expected to be mainly positive (e.g., Almqvist, 1991; Van der Mede, 1992; Wilbers, 1992).
The obvious restriction of the speed limiter as mandatory now, is that it only prevents driving above *the* maximum allowed driving speed of heavy goods haulage vehicles, and is independent of local limit in a specific road environment.
An intelligent speed adapter (ISA) takes into account local restrictions, and adjusts the maximum driving speed to the posted maximum speed. When it comes to restriction of driving speed of *private* vehicles, the use of intelligent speed limiters is to be preferred due to further differentiation of speed limits for private cars as compared to heavy goods vehicles.
A non-intelligent speed limiter is set at the maximum allowed driving speed for motorways (120 km/h), while the majority of a speed limiting system's safety benefits can be attained on 'A'-class roads (limit 80 km/h) and in built-up areas (50 km/h).

In general, a standard speed limiter is an intrusive system that restricts speed control, i.e. the device sets the maximum possible driving speed.
An intelligent speed limiter is able to set this maximum speed in accordance with local posted legal limits. A less intrusive device is a system that provides the driver with feedback about local limits, for instance, on the gas pedal. An active gas pedal increases the counterforce if the driver is driving too fast (Godthelp & Schumann, 1991). In principle, such a speed limiter leaves the driver in control, while the feedback provided in case of a speed violation is highly compelling. Moreover, the feedback is provided in the tactile modality, i.e. the same modality through which action has to be undertaken to observe the rules again.
Feedback could also be presented in the visual modality, e.g. a warning light or message in the dashboard, or auditory, an acoustic signal or vocal message. On the one hand this type of feedback seems less intrusive than the feedback an active gas pedal provides because these warnings can easily be ignored. On the other hand, it might be that the social effect of being warned in the presence of other passengers is a more severe chastisement and therefore less preferred. Anyway, acceptance of the feedback type systems can be expected to be higher than of a strict, standard speed limiter, because behaviour is less restrained.

Observation of behaviour at the level of driver reactions to these systems is of primary importance. However, apart from individual reactions, interaction with other traffic that is not equipped with speed limiters is also important. In such a 'mixed traffic situation', cars with a speed limiter could easily annoy drivers of cars that are not restricted and vice versa (Almqvist et al., 1991, Persson et al., 1993). These type of interactions deserve at least some attention in behavioural studies as well.

3.3.2. *Development of a prototype intelligent speed adaptor*

An effort is now undertaken in the Netherlands to develop a prototype intelligent speed adaptor that leaves the driver in control. For a start, this resulted in the development of a continuous feedback display in close proximity of the speedometer indicating the current speed limit, quite similar to the CAROSI system (Nilsson and Berlin, 1992). Central part of the CAROSI (CAr ROadside SIgnalling) system is the instrument panel, which includes not only standard displays such as the speedometer, but also contains sections on which roadside information is displayed. Amongst these is the posted speed limit, which is displayed below the speedometer. Major advantage of giving feedback by displaying the speed limit inside the car is that this information remains continuously visible instead of only being visible at the moment a sign is passed. This might reduce speeding because of general unawareness of the limit, which is not uncommon in the Netherlands (e.g., Steyvers et al., 1992; De Waard et al., 1995).

A special version of the latter type of feedback display is developed for implementation in the COV experimental test-vehicle. Whenever the speed limit is exceeded the colour in which the speed limit is displayed changes from green ('normal/neutral') to amber, or yellow, ('warning'). In case the speed limit is exceeded by 10%, the colour changes from amber to red ('violation'), and then an additional, auditory warning message is issued (see also De Waard et al., 1994; De Waard & Brookhuis, 1995a,b). The systems are integrated in the existing DETER system (see De Waard & Brookhuis, 1995a), which is developed as an open system to integrate driver monitoring and feedback (sub)systems. In the present experiment this set-up is tested, letting subjects drive the COV vehicle with and without the feedback systems.
Additionally, an active gas pedal is tested as a medium for haptic feedback in case of speed limit violations, exceeding by 10%, in the COV driving simulator with the same subjects, in a cross-over design. All modes of feedback are studied to effects on behaviour, mental workload and acceptance.

3.3.3. *Method*

Twenty-four subjects were paid for their participation in the test on effects of feedback concerning speed restrictions and violations in the institute's instrumented test vehicle and driving simulator.
Half of the subjects drove an instrumented test vehicle over a fixed route and then performed a simulator test, half of the subjects vice versa. Each of the test-rides consisted of two parts, first the baseline measurement, then after a short break, either the test ride with feedback or the control ride. Half of the subjects received feedback, half were in the control condition.
The test rides in the instrumented test vehicle were in normal traffic, under various conditions. Subjects were guided by sampled vocal route guidance messages that were triggered by the investigator for reasons of proper timing. They were led over a varied route that included sections of motorways, A-roads and built-up areas, with speed restrictions of 50, 70, 80, 100 and 120 km/h.
After each of the (four) test rides, subjects were requested to complete questionnaires concerning perceived workload and subjective driving quality. At the end of the whole test, subjects completed a general questionnaire again, asking for their ideas with respect to speed restricting systems again.

3.3.4. *Results and conclusions*

Although the number of detected violations during the second (feedback) series of trials is lower in the experimental group, this effect did not attain statistical significance. The difference in the number of speed violations between the two test facilities was significant, drivers more frequently violated the speed limit on-the-road. The extent to which the speed limit was exceeded was higher in the simulator. During the feedback-trial the extent to which the limit was exceeded was on average lower (with the exception of the second on-the-road trail for the control group).

A new parameter, the proportion of time violating the limit, does not discretely sum up the number of times the limit is exceeded, but reflects the time the driver is not complying. Two parameters were determined, the proportion of time driving above the limit, i.e. the time the display was or would have been amber or red, and the proportion time driving above the limit + 10%, i.e. the time the display was or would have been red and an auditory message was or would have been issued. The 'would have been'-condition is for the control group, the experimental group actually received the described feedback. As much as 20 to 25% drivers were speeding in the strict juridical sense. Between 5 and 10% of the time they are driving faster than the speed limit plus a 10% margin. The effect of the feedback system is highly significant for the latter parameter.

From the acceptance data it followed that acceptance very much depends upon feedback system, continuous feedback (on display) was accepted best of all means of feedback by far. The ratings for the continuous visual feedback were unusually high and can maybe even considered as (highly) appreciated.

An new, unexpected effect of the compound feedback was found, a significant reduction in speed variation. One of the reasons is the earlier mentioned use of the amber to stay in the margin of 'limit to limit+10%'. The implication of this finding is that less variation in driving speed could help to harmonise traffic, which is one of the candidate tools to reduce the number of accidents (see also Brookhuis & Brown, 1992).

No effects on workload were found in this study, again contrary to the first two experiments as mentioned. However, in the latter studies the (slight) effects were marginally significant, while in the present data the (slight) effects demonstrated in either of the two measures of mental load did not attain significance. The implication of these findings, in line with Verwey, Brookhuis & Janssen (1996), is that before implementing telematics systems, in principle workload effects should be measured, just to be sure, but the type of systems tested so far are not implying alarming effects.

3.4. Results and conclusions with respect to guidelines and criteria

The discussion in this paragraph is based on the following report:

- Verwey, (1996a-d). *Evaluating safety effects of in-vehicle information systems (IVIS); A field experiment with traffic congestion information systems (RDS-TMC) and preliminary guidelines for IVIS.*

Verweij (1996c,d) proposed several guidelines for IVIS. Guidelines and standards for IVIS are still under development. The currently existing and directly relevant standards for in-vehicle system design come largely from the office and manufacturing environment, such as human computer interaction but efforts are momentarily undertaken to develop guidelines and standards for IVIS within the ISO framework (TC/SC13/WG8/9).

Guidelines can be categorised as procedural guidelines, product guidelines and performance guidelines (Sherwood-Jones, 1990; Parkes, 1995). Procedural guidelines indicate how usability and safety should be assessed, performance guidelines specify acceptable user performance levels (maximum looking time) and product guidelines specify the physical aspects of the system (e.g. eye-display distance). Basically, a manufacturer needs product guidelines. However, Parkes (1995) argues that standards and guidelines for IVIS should be expressed as performance standards.
This makes the guidelines product independent and takes the interaction between various IVIS into account. This implies that standardised methods are to be developed which allows the translation of performance standards into product guidelines to which manufacturers can adhere.

Because safety should be the ultimate criterion for IVIS interface design, guidelines for IVIS should be tested against safety criteria. Furthermore, effect of IVIS on the traffic flow should be taken into account. So, guidelines for IVIS should involve procedural guidelines indicating how safety effects can be assessed, performance guidelines, indicating which performance levels indicate unsafe situations, and for certain types of IVIS, product guidelines indicating which system characteristics are likely to invoke unsafe behaviour. This procedure can be carried out for traffic flow effects as well.

In the following section, a framework is presented which allows determining which types of guidelines are required for preventing effects of IVIS on traffic safety. This leads to a limited set of procedural and performance guidelines and criteria. Also ergonomically oriented product guidelines will be discussed. The actual product guidelines themselves are presented in Verweij (1996).

It should be emphasised that the guidelines presented here are preliminary. They are derived from the experimental results and from existing guidelines and standards which have partly been tested with respect to traffic safety effects. If the text refers to 'the present experiment' the experiments discussed in paragraph 4.2 is implied. The guidelines hold for overload and distraction situations. That is, the possibility that the IVIS may also have positive safety effects, such as with anti-collision systems and detection of driver drowsiness, is not considered.

3.4.1. *A framework for developing safety related guidelines and criteria*

The development of guidelines for preventing negative safety effects requires a theoretical analysis of the driving task and the IVIS task.
The distinction between control and maneuver tasks can be extended into part tasks for which individual guidelines should be proposed. Then the relevant control tasks are course keeping and speed control. Relevant maneuver tasks are car following, intersection negotiation, and obstacle detection. The control task 'speed control' refers to the task of keeping the vehicle on the road. This is especially relevant with respect to curves. Speed

control in interaction with other vehicles are considered maneuver tasks (e.g., car following).

For quantitative criteria, values may differ for various types of road. For example, criteria for glance times based on course keeping criteria may differ for freeways and residential areas because of the differences in lane width and speed. Furthermore, the consequences of not obeying these criteria will differ for different roads. Safety effects will be much greater when obstacle detection is affected in residential areas than at freeways where obstacles are rare.

Finally, criteria will have to be developed for different types of human-machine interaction. With respect to IVIS, these criteria should indicate the degree of glance time for *visually* demanding displays, *mental* workload for complex messages, *manual* complexity in system control tasks, and the extent that interaction with the system is *paced* by the driver or the system. Future work might also take effects of haptic and kinesthetic information into account.

3.4.2. *Safety assessment and performance guidelines*

This section proposes preliminary performance guidelines which are most crucial for safety evaluation. Later sections will present procedural guidelines which will have to assure that the appropriate methods are being used. Finally, product guidelines will be discussed.

3.4.2.1. *Visual messages*

The safety effect of visually presented messages lies quite obviously in the fact that looking at a display interferes with looking at the road environment. Even though peripheral detection of objects outside the car is possible when glancing at an in-car display, this possibility should not be taken too seriously.
Guidelines to prevent driver overload by visual information presentation should preferably be expressed in terms of total glance time, time of individual glances, and glance frequency. Even though people have some liberty in choosing the frequency/duration ratio when extracting information from a single display, it appears that the total glance time is relatively constant across driving situations. This suggests that total looking time is a reasonable measure for expressing visual workload criteria.

Visual in-vehicle displays should not require more than three glances of 1 s and it should be possible to acquire useful chunks of information in at least one second. Messages should always be driver paced in that the information will remain to be presented for a relatively long time (e.g., 1 min) and the driver is free to decide when to look. For the display of complex information, the possibility should be considered to present information in relatively simple portions which are presented only when requested by the driver. Only with heads-up displays longer glance times are allowed because course keeping and car following can then be carried out with peripheral vision.

The following issues still need to be resolved:

- Do drivers adapt headway safely when they are looking at an in-vehicle display in a car-following situation?
- Do they take following traffic into account when braking in front of an intersection or will the probability on rear-end collisions increase due to sudden hard braking?

3.4.2.2. *Mentally demanding messages*

On the assumption that drivers should not fully attend to other tasks for longer periods of time even though their eyes are directed at the roadway, auditory (speech) messages should not last longer than a certain period of time unless the messages do not require full attention, such as with highly familiar messages or messages that do not provide crucial information.
For the time being, it is proposed that loading messages of more than five seconds should be segmented and the driver should indicate when each segment is to be presented. A repeat function of the last message should be available and be easily activated. Obviously, it is necessary to determine how long highly loading speech messages may take before safety is affected.

Notice that repetition and segmentation are characteristics which are usually inherent in normal (and informal) phone conversations. In case of conversations requiring much attention or more formal telephone conversations, in which the driver might hesitate to interrupt and ask for information again, safety might be affected because the driving task gets insufficient attention (cf. Briem & Hedman, 1995; McKnight & McKnight, 1993; Parkes, 1991). Similarly, when drivers listen closely to the normal traffic messages the safety of driving reduces (Akerboom, 1989). In other words, it has been demonstrated that driving performance and safety are affected with tasks that are legally allowed now and a test for safety effects should be able to show this.

3.4.2.3. *Manual system control*

Manual system control may affect driving safety due to visual, mental, and manual demands of the task. Task analysis should estimate which of these demands are most detrimental to safety. For now, the following guidelines can be presented. Controls requiring visual feedback during their use, such as controls which are small, close together, or which function is visually indicated, should be avoided. The need to look at the movement is allowed only when reaching for a control. Hand support and tactile cues should facilitate control without looking at the movements. The direction of movement of a control should take account of the location and orientation of the driver relative to the control. Critical and frequently used controls should be close to the predominate position of the hands and should be relatively big. Make errors difficult but be forgiving. All controls should be in easy reach of the driver. Single handed operation should always be possible. Returning the hand to the steering wheel should be possible immediately.

3.4.2.4. *Sensory modality and pacing*

Visual or auditory information?
The present results show that visual as well as complex auditory messages might reduce safety. This indicates that the choice between visual and auditory information presentation is one that requires careful consideration.

Auditory signals may be used to inform the driver that important visual information is being presented. This signal should not be such that the driver startles and the nature of the visual message should not force the driver to look immediately which would change the task in a system-paced task.

Also, auditory information is recommended for signals of acoustic origin, for warning signals, to draw attention to visual indicators, when information must be presented independently of the orientation of the head, and when vision is limited or impossible. Tonal signals can be used when the message is extremely simple, the signal designates a point in time, the message calls for immediate action, speech signals are overburdening the driver, and conditions are unfavourable for speech messages (e.g., high noise levels). Speech can be used when flexibility of communication is necessary, rapid two-way exchanges of information are necessary, and the message deals with a time-related activity.

Pacing

Even though the aspect of pacing has been mentioned a few times, this characteristic can be considered of primary importance. Drivers are able to perform complex tasks and process complex information as long as they are able to indicate when they are able to do so. This makes them responsible for negative safety effects. Therefore, the system should be designed such that pacing, though determined largely by the driver, is still limited by the system. So, the IVIS should not present more information than can be processed in a limited time (visual 1 s, speech 5 s) and the rate at which next chunks are presented should be limited with sufficient inter-chunk intervals to allow the driver to redirect attention to the driving task.

3.4.2.5. *Testing safety and performance guidelines*

This section presented guidelines and criteria for the human-machine interface of IVIS. These guidelines were presented with respect to the visual, mental, and manual demands of the driver-IVIS interaction. Visual demands should be limited to three or less glances of up to 1 s on the average. Any visual information should be presented sufficiently long so that the driver has ample time to scan the display at a moment that the driving task allows display scanning. Smart display design should prevent individual glances of more than 1 s.

Since instrumented vehicle studies are time and money consuming, there is the need to test new IVIS with respect to their safety effects in a relatively simple setting. The visual workload associated with IVIS displays could possibly be assessed by a laboratory test in which subjects watch the IVIS display for successive 1 s intervals. The duration of these intervals are system controlled but the onset of each glance should be controlled by the subject. Visual occlusion can be created with spectacles or another device occluding part of the visual scene. To mimic the temporal and mental demands of course keeping, subjects should also perform a tracking task with time characteristics similar to those found in real course keeping. In this setup, visual workload is equated to the number of 1 s intervals given that tracking performance was acceptable. The IVIS display is considered safe when no more than three 1 s glances were required for understanding the information.

Mentally demanding speech messages should not last longer than a fixed period of time. Only when speech messages are not very loading, for example, when they are familiar or redundant, they may be longer. Loading messages that last longer than five s should be segmented and the driver should indicate explicitly whether the next or the last segment should be presented next.

The safety effect of speech messages should be determined in a laboratory setting too. Such a test should assess both the duration and the mental workload of the message. A possible metric for mental demand of speech messages can be developed with the Continuous Memory Task (CMT) technique (Boer & Jorna, 1987). Then the IVIS safety test would involve responding to speech messages of the IVIS while at the same time counting the number of target letters that are displayed visually. Higher workload and longer IVIS message durations will reduce CMT accuracy. Given the lack of data on safety effects of mentally demanding messages, driving studies should indicate which levels of mental demand are acceptable. The speech messages used in the present experiment might be used as a first approximation of unacceptable mental workload. Similar tasks of varying complexity can be used to determine safety as a function of mental workload.

The present results clearly show that manually controlling a system might affect safety. The cause may lie primarily in the visual, the mental, and the manual workload of the interaction. For visual and mental workload, the above criteria should be used: no more than three 1 s glances and no high mental demands for longer than a certain period of time. Operating complexity should be limited too in order to limit the visual and mental workload caused by movement control. Again, driver workload should be tested in a laboratory setting in which interference with a secondary task indicates the workload of the interaction.

Now, this secondary task should be sensitive to all types of workload associated with manual system control. For example, subjects will have to perform a tracking task with timing characteristics which are comparable to lane keeping (indicating visual and manual workload). Certain simple discrete actions will have be made in response to stimuli which are presented at various locations (indicating mental workload). The timing and location of some of these stimuli can be anticipated, others cannot. Next, the effects of a couple of different manual control tasks should be tested with this task as well as in a safety study in real traffic in order to determine the relationship between workload and safety effects. Then, the degree of performance reduction on this laboratory task will indicate the degree of safety reduction by controlling the IVIS.

Finally, no matter the type of task the driver has to perform with an IVIS, explicit attention should be given to the timing of the driver-IVIS interaction. The driver should always be able to determine when he or she is able to pay attention to the IVIS and should always be able to interrupt an ongoing interaction. Furthermore, interactions that require much attention may never last longer than a certain period of time (visual: 1 s, speech: 5 s for the time being) until the driver indicates that a next part of the task can be performed. For visual displays and manual control this implies that the state of the system does not automatically change. Speech messages should be of limited length and repeatable. All IVIS should comply with these guidelines on pacing

3.4.3. *Procedural guidelines for safety testing*

Eventually, each IVIS should be tested on its safety implications by assessing the safety effects caused by its support to the driver (positive safety effects) and its effects due to workload and distraction, and its effects

of exposure (negative safety effects). It is not possible to give concrete procedures for safety effects due to driver support and changes in exposure for all IVIS as these effects are largely dependent on the system at hand.
For example, with respect to safety effects caused by changed route choice, potential users may be interrogated how they would adapt their route choice given certain types of IVIS messages.
Effects can also be established in fleet-studies but this is usually not feasible. If negative safety effects can be expected due to improper route choice, system design may be changed such that they guide drivers more along roads that are known to be relatively safe. Also, when certain roads are known to be used more frequently due to some types of IVIS, measures may be taken to increase the safety of those roads.
Safety effects due to workload and distraction, which was the topic of the present report, can probably be tested by laboratory tests. With respect to visual workload, the guidelines are sufficiently clear to allow laboratory assessment of safety effects although more detailed analysis might still test the 3 times 1 s rule in more detail. Laboratory assessment of mental workload and the workload involved in manual system control is not yet possible. The laboratory procedures to investigate these types of workload require further detail.
Also, there is still the need to link these results to safety effects in actual driving. Tasks with various experimenter-controlled levels of mental and manual workload should be carried out both in an instrumented vehicle and in a laboratory setting. The safety effects obtained in real driving will indicate which are the criteria for acceptable levels of performance in the laboratory tests. It should be stressed that, as long as many of the criteria are unknown, assessment of these safety effects should be done both by way of by (subjective) expert opinion and by measurement of objective characteristics. Once these criteria are known, safety effects of IVIS can be determined in relatively simple laboratory tasks.

3.4.4. *Product guidelines*

In order to function appropriately, an IVIS should fulfill elementary ergonomics criteria. The product guidelines are derived from various proposals in the literature (Boff & Lincoln, 1988; Galer & Simmonds, 1984; Green et al., 1995; Sherwood-Jones, 1990). These guidelines do not guarantee safety but are aimed at coming to optimal designs for individual human-machine interaction. As such, they can be used for improving IVIS which are likely to affect safety. The guidelines are meant for information which is not directly alerting. That is, the messages are assumed to convey information that need not be perceived quickly by the driver. For alerting information which require immediate driver responses, more conspicuous messages may be allowed.

The product guidelines should follow some general principles: It should be possible to turn off all IVIS at any time with a single action in order to inhibit all potentially distracting messages. Be consistent across various subsystems or tasks and across time with respect to system characteristics that recur such as color coding and data entry. For example, use the same sequence of actions across subsystems for entering information and for presenting similar types of information. Comply with people's expectations. For example, increase volume and any another value by rotating clockwise. Minimise what the driver has to remember during interaction with the IVIS. Operations that occur most often or have the greatest impact on driving safety should be the easiest to carry out. For example, setting a destination should be easier than recalibrating a system. Also, present only information useful to the driver; always consider whether the driver needs the information that is presented.

Two types of ergonomics criteria can be distinguished. Criteria that will have a direct effect on a safety and performance evaluation and those that do not have such immediate effects because they are meant for special conditions such as low or high luminance or high noise conditions (driving fast). Characteristics of the IVIS that directly affect safety and performance will affect the measurements in the safety and performance test described in the previous section. No further test of these criteria is needed. Yet, they can be considered for improving the IVIS. The ergonomics criteria that are meant to improve the IVIS under specific conditions, require careful consideration in a safety evaluation too. For each IVIS, it should be demonstrated that that system can be used under the various conditions that occur in driving such as different light and noise conditions.

The list of ergonomics criteria and product guidelines (Verweij, 1996) forms a core list. On the basis of this knowledge, each IVIS should be tested on whether they will still convey their information and perform their task under adverse light and noise conditions. This will have to prevent that the demands of the IVIS increase under such conditions.

4. Criteria for experimental testing

The discussion in this paragraph is based on the previous chapter and on the report:

- Verwey, W., Brookhuis, K.A. & Janssen, W.H., (1996). *Safety effects of in-vehicle information systems (IVIS).*

In the present chapter it is discussed what type of experimental study is required for assessing safety effects caused by the use of telematics applications. It describes a series of general guidelines, or building blocks, of how potential overload, underload, or counterproductive adaptation effects should be investigated in relation to safety.

4.1. Summary of results

Criteria for experimental tests of safety basically consist of three parts:

- recommendations for the general design of the experiments;
- recommended parameters to be observed;
- recommended procedures for interpretation of the results.

The results on these items are summarised below.

4.1.1. *Recommendations for the general design of the experiments*

To get an overall safety assessment it seems useful to use a representative sample of the driver population in the sense that the number of older, middle-aged, and younger/inexperienced drivers should be proportional to the distance driven by these three groups. Since the resulting interindividual noise might reduce the reliability of measured effects, a more labourious but better method than having a representative sample is to have three groups of drivers and weigh their contributions according to their relative contributions to the annual distance driven.

After obtaining some experience with the operation of the ATT device to be tested, these subjects should drive in a number of crowded road situations for 45-60 minutes.

Furthermore it is recommended that safety assessment will be based upon three phenomena: task overload, task underload and counterproductive behavioural adapatation. This assessment requires at least the observation of the following parameters.

4.1.2. *Recommended parameters*

The following list of measurable parameters is based upon a number of sub-tasks and the task load of the driving task.

- lateral position handling represented by:
  - SDLP (standard deviation of lateral position);
  - steering wheel handling;
  - time-to-line crossing;
  - the proportion under the distribution of lateral positions of the vehicle that is actually outside the lane.

- headway control represented by:
  - value of time headway;
  - time-to-collision;
  - the proportion under the distribution of headways to preceding vehicles that is below some critical value, notably 1.0 s ('short headways') or 0.5 s ('critically short headways').

- speed management represented by:
  - actual speed;
  - characteristics of the speed- and acceleration distribution: standard deviation of speed;
  - excess of speed limit;
  - pedal use;
  - the 50th and 85th or 95th percentile of the driving speeds and 50th and 85th or 95th percentile of acellerations and decellerations may be used to characterise the distributions of these parameters. The changes in these values provide sufficient insight in the way these distributions are affected (usually shifted) as a result of task interference.

- general driving behaviour like:
  - running red/yellow lights;
  - reaction time to important events;
  - reaction time to events of secondary importance;
  - lane exceedence.

- task load represented by:
  - self reported taskload (according to scales like RSME, NASA-TLX);
  - heart rate variability;
  - mirror usage;
  - duration and frequency of glances at display;
  - steering wheel frequency;

4.1.3. *Safety assessment by expert opinion*

The parameters listed above can be observed and compared to, more or less, established criteria. However, individual differences that are hard to assess will still provide considerable difficulties in interpretation of these objective measures. It is therefore strongly recommended to include a subjective assessment of the safety of driver behaviour by an experienced driving instructor. So far the consistency and comprehensiveness of this subjective rating has been proven superior to the interpretation objective measures. In section 5.2.2, under the heading 'expert opinions', references are provided for the type of variables and assessment techniques that are to be used by such experts.

4.2. **Building blocks for studying overload**

4.2.1. *Assumptions*

The method evolving from the building blocks in this section involves a number of assumptions. First, the participants in the study are assumed to behave as they would do in normal driving. This is a *very* strong assumption, considering that subjects are driving an instrumented vehicle under surveillance, even if it is in real traffic. For driving on a closed circuit and for tests in driving simulators, this argument holds a forteriori. Therefore, care should be taken that driver performance and behaviour are as normal as possible.

To make the situation more realistic, participants will be offered bonuses to increase time pressure and to prevent over-cautious driving behaviour. Also, it can be stated that system performance is the topic of study rather than driving behaviour.

Furthermore, it is assumed that the participants are familiar with the telematics application at hand. Basically, their performance and workload should more or less have reached an asymptotic level when the experiment starts. This is more problematic as the application is more complex because it takes more time to train the participants. General guidelines in this respect can not be given. With respect to route guidance systems, Zaidel (1991) suggests a period of four hours. To check whether there is sufficient practice, the study should involve a pilot study in which repeated measurements show how long it takes the participants to get used to the telematics application.

In general, drivers should not exceed certain performance boundaries, should not undertake unsafe driving actions, and should respond sufficiently fast to predictable and unpredictable events.

4.2.2. *Assessing safety effects*

The major issue is how one can decide whether a telematics application jeopardises traffic safety or not. For example, even if it is shown that workload increases this need not necessarily affect safety as the workload increment may occur at places where workload is low anyway. There is a need for cross-validation between objective measures and subjective measures as given by driving experts such as driving instructors. That is, a study should be performed in which drivers follow a prescribed route with and without a telematics application.

The objective measures summarised in paragraph 5.1 should be employed which are related to aspects of driving safety and workload and which are likely to be affected by the use of a telematics application. Expert drivers should rate driving performance on a series of safety related aspects. Both subjective and objective measures should be assessed per driving situation. This procedure will allow cross-validation of the subjective and objective measures so that acceptable boundary values for the various measures can be established. Furthermore, to check the reliability of the expert opinions, inter-rater variability should be determined. If the inter-rater variability is high the ratings are of little use. In order to get an idea of how the safety effects relate to the interaction with conventional systems (radio, cassette, heating system), the study might have interaction with these systems as a control condition (apart from a non-interaction control condition).

The driving task involves a series of subtasks differing in their relationship with traffic safety.

Subtasks of driving that are closely related to safety *should not be affected in any case.* These are:

General driving behaviour

Lane exceedence frequency, running red/yellow traffic lights, and reaction time to important events should not be affected at all. Reaction time to events of secondary importance should not change more than a certain percentage.

Speed management
Speed should never increase and should not drop more than a certain percentage. Brookhuis (1991) found that drivers show speed drops whenever their attention is diverted seriously, which is demonstrated, for example, in case of telephoning while driving. How much change is allowed will depend on the driving situation and the degree that it relates to the opinions of experts. Munden (1967) asserted that speed should not change more than one standard deviation from the mean for else accident probability will increase too.
Furthermore, deceleration should not change more than a certain percentage.

Lateral position handling
Steering frequency and Time-to-Line Crossing should not change more than a certain percentage.

Headway control
TTC should remain above the 1.5 s criterion. The minimum following distance should not change more than a certain percentage.

Task load
Also, drivers should never look at a display for such a long time that there is the (theoretical!) possibility of hitting other traffic participants or leaving the road. Usage of mirrors should also not be affected more than a certain margin.

Expert opinions
The expert observers should provide detailed opinions about the quality of driving. Zaidel (1991) provides a list of driving quality variables. Use can also be made of a standardised observation technique for assessing erroneous and unsafe behaviour (e.g., Galsterer et al., 1990; Risser & Brandstätter, 1985). Evaluations should be aggregated and averaged across the units of evaluation (street, kilometre of road, junction, etc.) so that different trips can be compared. To prevent experts from knowing whether the driver is or is not interacting with a telematics application, their judgements should preferably be made from video so that they do not know whether a telematics application is being used.

General design
Often the effect of a telematics application is likely to be smaller than the differences between individuals so that within-subject designs are required. After sufficient experience with the telematics application, subjects should drive in crowded road situations, preferably in their own car, for 45-60 minutes.

Notice that it is important to give participants in an experimental study relatively much freedom so that they can compensate for the occurrence of overload. For example, if headway increases when interacting with a telematics application this indicates an increased level of workload but not necessarily unsafety. In this case, eye movement analyses may indicate whether the headway increase is sufficient for compensating for the visual attention paid to a display.

In short, assessment of driver overload should not affect driving performance. Also, the measures should have a high temporal resolution and allow one to determine rapid increases in driver workload while interacting with the telematics application. In this way, a close relationship can be established between rapid variations in workload and interactions with the telematics application. Basically, these criteria exclude the use of secondary tasks and subjective estimates which were precisely the ones mentioned as best indicators for workload peaks by Verwey & Veltman (1995). However, subjective ratings might be given immediately after a loading situation has been negotiated.

Workload indicators
Which workload indicators are chosen depends on the effects that are expected. If workload is expected to be higher over longer periods of time, heart rate variability can be used but for short (<60 s) workload elevations they seem less suited (Verwey & Veltman, 1995). Steering frequency, although not sensitive to 10 s peaks, turns out to be useful for longer periods of increased workload and an easy to capture indicator for (visual) workload. For subjective estimates, univariate lists are preferred such as the RSME. Finally, the driving task parameters advanced in *Table 1* should be assessed when effects of driving are expected. For direct effects on driving performance variables should be assessed which are closely related to safety (e.g. lane accedence frequency, running red/yellow, TTC). More indirect effects can be assessed by, for example, mirror usage, speed and steering wheel frequency but these measures are not so tightly related to safety. Hence, an effect at these variables need not indicate safety effects directly. Whether safety is affected depends on the type of situation.

4.3. Building blocks for studying underload

Safety effects with respect to underload are in fact a consequence of driver impairment, i.e. when driver performance deviates from the optimum of the U-shaped performance curve (Wiener et al., 1984). Similarly to what was stated for driver overload, thresholds or limits or criteria have to be established to decide when exactly performance should be considered as unsafe. It may be possible to assess on the one hand absolute criteria for a number of measures c.q. variables, on the other hand these criteria will be too lenient (or strict, depending on the point of view) in many instances for assuring actual safety. For example, if time-headways of less than 0.7 seconds are below the absolute red line, then it still might well be that a following time of 0.8 seconds is disastrously dangerous in heavy rain or after four hours of driving, or quite differently, for a certain driver of old age. For this reason relative criteria, i.e. dependant on situation, condition and perhaps even personal characteristics, will necessarily have to be incorporated in the test too. Below, each of the candidate measures as mentioned is shortly treated to its merits for inclusion in the check list of features and dispositions of the applications.

Lateral position handling
Since around 1980 lateral position statistics are developed and repeatedly applied in 'on-the-road' tests for measuring effects of drugs on driving performance (O'Hanlon et al., 1982; Brookhuis et al., 1990). In particular the ability of the driver to control weaving of the car, measured as the standard deviation of lateral position, is a very sensitive indicator of drug-induced sedation (O'Hanlon et al., 1982). From the relationship between BAC and the ability to control weaving, expressed as SDLP (standard deviation lateral position), it follows that the critical increase in SDLP should start in the range of 2.5 to 4 cm. It is also worth while to put

some effort in interpreting SDLP in terms of traffic safety. The probability of lane edge accedence can be taken as a (surrogate) measure of potential accident involvement. It follows from this relationship that starting at an absolute SDLP of around 25 cm the likelihood of lane accedence rapidly increases into a substantial proportion, inevitably leading to accidents in due course.

Steering wheel handling

A number of steering wheel parameters have been developed in the past, indicating decrements in driver performance (McLean & Hoffman, 1975), for instance, steering wheel reversals, number of zero crossings, standard deviation, amplitude range, integrated amplitude over time, jerk (rise time), frequency characteristics, and recently more sophisticated combinations of some of the earlier mentioned parameters. Examples of the latter ones are the Y-functions, developed by Renault (Petit et al., 1990) and the S-function, developed by INRETS (see Fairclough et al., 1993). Careful field experiments including physiological measurements for independent establishment of signs of decreasing vigilance lead to the determination of thresholds in these respective functions with regard to the turnover from vigilant into non-vigilant. The Nissan company combined lapses of not moving the steering wheel followed by a sudden jerk into a formula that indicated non-vigilance (Yaouta et al., 1985).

Time-to-line crossing (TLC)

While the SDLP and SDSTW mainly reflect performance at the control level, one level higher, at the manoeuvring level of performance, the Time-to-Line Crossing (TLC, Godthelp, 1984) is a measure of driver primary-task performance. TLC is a continuous measure that represents the time required for the vehicle to reach either the centre or edge line of the driving lane if no further corrective steering-wheel movements are executed. TLC reflects the time drivers can neglect path errors. Due to the measure's skewness, in general minimum, median or 15% TLC values are calculated (Godthelp et al., 1984, Godthelp, 1988). TLC is expected to reflect driving strategy and in particular occlusion strategy (time spent not looking at the road). With increases in mental load, smaller TLC values can be expected; a more demanding task is likely to decrease the amount of time spend looking at the road.

Speed management

In the past decades, several authors mentioned ‘deliberate’ speed variations as ways to suppress upcoming fatigue. Drivers increase their vehicle’s speed with the objective to increase their (feelings of) arousal, only temporarily of course, after which they let their vehicle’s speed drop again. In the series of prescription drug experiments carried out at the COV, significant increases in standard deviation of speed was demonstrated, predominantly in cases of severe sedation.

It seems reasonable (but not irrefutable) to suppose that drivers normally keep their speed steady at a level of their choice, at least while circumstances allow such freedom, which will in many cases be the local speed limit. Exceeding the speed limit by 10 percent, not surprisingly the margin the police ‘allows’ in most situations (except the 120 km/h highway speed limit), seems a reasonable criterion *at first shot* with respect to standardisation. The standard deviation of speed that accompanied high SDLP after some prescription drugs and was significantly different from placebo by itself, amounted to a 5 percent increase.

Headway control

A certain distance between cars has to be kept in an absolute sense to be able to react adequately, which is dependent upon the speed (Fuller, 1984; De Waard & Brookhuis, 1991). Therefore, a criterion for time headway, a measure of distance relative to speed, should be feasible to be assessed as absolute threshold for safe driving from the literature, whereas a relative criterion seems still to be determined. Recently the standard driving test, developed and in use by COV, has been expanded to include a car-following test to study the reaction of the driver to driving behaviour (i.e. speed management) of other road users with respect to distance keeping or headway control (Brookhuis & De Waard, 1993). Drug sedation or activities such as using a car telephone have been shown to impair these reactions. From several experiments a baseline criterion could be derived for preliminary (absolute) threshold, as well as a relative criterion based on alcohol and prescription drug affected increases.

Time-to-collision (TTC)

In general, the Time-to-Collision (TTC) measure is especially designed for and useful in studying conflicts. Conflict observation studies are in particular important at intersections and for red-light enforcement systems. In relation to speed management, conflict studies may be useful for estimating the speed differentials. Following the Time-To-Collision approach (TTC, Van der Horst, 1990), the detector data of speed and headway will be used. In the original TTC measure acceleration of lead and following vehicle should be taken into account. This is not feasible with detector data, which for the time being can only provide a momentary vehicle status. Therefore, TTC in the present test methodology will be of limited use.

Pedal use

Although pedal use is one aspect and sometimes the lever mechanism of speed management, in a sense, it is treated separately here. Fairclough et al. (1993) studied brake pedal use after alcohol and found significant differences after low amounts of alcohol (BAC <.05%), compared to sober trials. Subjects decelerated later and more forcefully as they approached junctions. Also approaching and taking bends is carried out differently after relatively low amounts of alcohol.

4.4. Building blocks for studying counterproductive adaptation

Design

The most appropriate way of doing an experiment that could detect counterproductive adaptation behavioural effects is to use a before-and-after, within-subjects design with a separate control group for estimating carry-over effects. Thus, each subject is employed in two sessions, the subjects in the control group being without the device in both sessions. The conditions in both sessions should be identical. For example, when a field experiment is performed the route driven and the traffic intensity should be the same, etc.

Long-term experimentation may be required in order to judge whether a device will result in counterproductive adaptation even if counterproductive adaptation did not occur in the initial measurements. Some studies have shown that counterproductive adaptation, even to relatively conspicuous safety measures, may extend over a period of at least a year (e.g., Janssen, 1994).

A final requirement in experimentation is that groups of subjects should be matched on all the relevant variables that one can think of. This may not only include such well-known variables as gender, age and driving experience, but also variables that have to do with risk-seeking and risk-avoidance. It may well be that the inclination to seek or avoid risks interacts with the inclination to indulge in counterproductive adaptation.

Measurement

The behavioural variables that should be measured and analysed in order to detect possible compensatory effects are not different from those used in other settings and that are summarised in paragraph 5.1.

However, because an interpretation in terms of (net) safety should always be attached to a counterproductive adaptation-type experiment, attention should be devoted to entire distributions of values within variables, and to changes in the 'risky' tails of distributions in particular.

Specifically, one should look for effects in:

- the 50th and the 85th or 95th percentile of the distribution of driving speeds
- the 50th and the 85th or 95th percentile of the distribution of momentary accelerations
- the 50th and the 85th or 95th percentile of the distribution of momentary decelerations
- the proportion under the distribution of headways to preceding vehicles that is below some critical value, notably 1.0 s ('short headways') or 0.5 s ('critically short headways')
- the proportion under the distribution of lateral positions of the vehicle that is actually outside the lane.

In all these cases theoretical distributions can be used to fit data that have been obtained in measurement sessions that, for one reason or another, may have lasted too short to produce exhaustive distributions. For example, while a subject may never have exceeded his lane boundary in the experimental session proper his probability of doing so by a given amount may be calculated from the theoretical distribution fitted to the sample that is available.

It may often be more convenient as well as more illuminating to work with generalised factors, derived from the separate variables, than with the factors themselves. MANOVA, or different available forms of factor analysis, may yield these more general descriptors, based on the patterns of relationships (correlations) that exist among the separate variables. These descriptors are indicative of driving styles, so that counterproductive adaptation-like findings can be expressed as changes in overall driving style. Whenever it is expected that counterproductive adaptation could manifest itself in the form of a drop in overall driver alertness any of a number of standard physiological parameters can be monitored.

4.5. Selecting participants and driving situations

It is likely that the effects assessed differ substantially for different drivers and different groups of drivers. It is generally found that older drivers (say over 65) are more likely to be involved in traffic accidents. Elderly drivers appear not only to be disproportionally involved in road accidents but are also more often legally at fault (Engels & Dellen, 1983; Fontaine, 1988). This shows that elderly are a group of drivers that should be included in any study of telematics applications that are likely to be used by elderly as well. In a review of the literature of navigation systems, Kaptein et al. (1993) found twelve studies which included age as independent variable, eight of which showed significant differences between younger and older drivers.

Another group of drivers which is more loaded by driving and which is also disproportionally involved in traffic accidents are inexperienced drivers. Whereas the very inexperienced are likely to be overloaded by the driving task in itself because of limited experience (Verwey, 1991), traffic accidents appear to peak some time after being licensed. For example, Pelz & Schuman (1971) showed that accident risk was highest two or three years after being licensed. It is unclear to what extent increased accident rates in the inexperienced driver is affected by reckless behaviour as often found in younger people or by inexperience per se. When testing safety effects of telematics applications, the effects of inexperience and age can be combined by taking drivers below 25 years of age and with one to three years of driving experience.
Hence, the recommendations as they are worded in the summary of this chapter.

5. Laboratory testing

5.1. Laboratory Test Results by TNO-HFRI

At present no standard procedure to test safety effects of IVISs is available. A standard test was developed that was aimed at evaluating safety effects of IVIS. A number of requirements of such a test are discussed in Steyvers, Van Winsum & Brookhuis (1997) where this test has been named IVIS-LabTest.
The experiments discussed in this chapter were aimed at validation of this test with two different IVIS applications. It is investigated whether it is possible to show effects of IVIS on driver behaviour and whether the test is sensitive enough to show differences between different types of IVIS applications. However, the relation between effects of an IVIS on driver behaviour and traffic safety is not always clear. Verwey, Brookhuis & Janssen (1996) discussed overload, underload and counterproductive adaptation as possible causes for negative safety effects of in-vehicle information systems.
In the experiments the effect of performing an attention demanding secondary task on primary task performance was investigated. Thus, the effect of overload induced by a secondary, non-driving related, task on driver behaviour and traffic safety is examined in some detail. For practical reasons the experiment was divided between two research institutes: TNO Human Factors Research Institute TNO-HFRI) and the COV (Centrum voor omgevings- en verkeerspsychologie).

5.1.1. *Introduction*

The discussion in this paragraph is based on the report:

- Van Winsum, W. (1997). *A validation study of a PC-based test of safety aspects of in-vehicle information systems; A test of a map display version of a RDS-TMC task.*

Two versions of an RDS-TMC (Radio Data System-Traffic Message Channel) application have been tested that are similar to the tasks discussed in Verwey (1996). Katteler (1996) has given an overview of a number of RDS-TMC systems that probably will be marketed in the near future. RDS-TMC systems provide information about the location of traffic congestions. Some systems provide information on a map display while others provide the information as a speech message. At TNO-HFRI the effects of a visual demanding RDS-TMC (map display) application was tested while the COV tested the effects of an auditive RDS-TMC application.

The IVIS-LabTest runs on a PC and it uses a steering wheel, an accelerator pedal, a brake pedal and two response buttons as input devices that give their signals to the computer program via a bus mouse or a game port.
The scene consists of the one-lane road with a lane width of 3.6 m.
The driving lane is coloured grey-blue and the are no lane markings. White poles with a distance of 30 meters between each pole are positioned next to both sides of the road. The road consists of alternating curved segments to left and to right. Each curved segment has a continuous radius of 2000 meters and deflection angle of 45°. The road is drawn within a green

surroundings. The horizon is drawn at half the screen height. Above the horizon white/grey clouds are drawn. Eye distance from the screen should be approximately 1 meter.

Speed is controlled by the accelerator and brake pedals and restricted to a maximum of 120 km/h. The present speed is indicated by a moving horizontal bar at the bottom of the screen. Lateral position is controlled by the steering wheel. The driver control actions are the input for a simple vehicle model that has an engine model with an automatic gear box and a model of vehicle dynamics. Engine sound, depending on engine rpm, is delivered to the driver via an external loudspeaker. If one wheel gets off the road, the driver receives auditive feedback by a change in the characteristics of the engine sound. If the vehicle moves off the road entirely not only the sound characteristics change but also the road friction. This results in a lower speed. The driver then has to push the accelerator pedal to the maximum and steer back into the road.

During a test run, the following variables are sampled and stored for eventual evalution, using a sampling frequency of 10 Hz:
x and y coordinate position, lateral position in meters (distance of centre front of vehicle from right lane boundary), vehicle speed in km.h, steering wheel position, accelerator pedal value, brake pedal value and response button.

5.1.2. *Method*

The RDS-TMC task was similar to the map condition described in Verwey (1996). In the version used in the present experiment every stimulus started with a double beep alerting tone presented via a headset the subject was wearing. After this, a computer generated voice presented the route the subject was driving, for instance "you are driving from Utrecht to Apeldoorn. Is the road congested?". Then 1 second after termination of the voice message, a computer generated map was displayed on a 15 inch monitor to the right of and behind the monitor for the primary task.
This required a small head movement of the subject in order to inspect the map. The map consisted of a set of major and minor roads and a number of cities. The cities and roads were not geographically in the correct position. This simulates a situation where the driver is not familiar with the region and has to actively search for the position of the cities on the map. All roads were drawn in green. Major roads were drawn with thick lines while minor roads were depicted as thin lines. Congestion was indicated as a red line instead of the usual green. Subjects were instructed to look at the map and search for the most direct major road connection between the two named cities. If there was a traffic congestion on the route the yes button (upper button right of the steering wheel) had to be pressed as soon as possible.
If there was no congestion the no-button (lower button right of the steering wheel) had to be pressed as soon as possible. If the response did not occur within 10 seconds after the map was displayed it was coded as a non-response.

All stimuli started at the beginning of a road curve to either left of right at a position that was the same for all subjects. Each IVIS trial contained ten stimuli of which five required a yes response and five required a no response in random order. A total of five different questions with a route description and eleven different computer generated maps were used to choose the stimuli from.

After the practice tasks four experimental tasks were executed. These consisted of a combination of two levels of IVIS, i.e. either with or without secondary IVIS task, and two levels of pacing, i.e. either self-paced or forced-paced. In the forced-paced condition the vehicle drove with a constant speed of 80 km/h. In the self-paced condition the subject was free to control the speed up to a limit of 120 km/h. The order of tasks was balanced.

Twenty-four subjects participated in the experiment. All were experienced drivers with at least five years of driving experience and driving more than 5,000 km per year. Half were male and half were female. Age ranged from twenty-two to fourty years. Subjects were selected from the institute's subjects pool.

5.1.3. *Results and conclusions*

The results show clear effects of the IVIS tested in the present experiment on driver behaviour. Interaction with a secondary IVIS application that induces a high visual load results in adaptation of the driver of primary task behaviour: the driver reduced speed. Although the speed variation within the subjects (SD of vehicle speed) failed to reach statistical significance, the speed reduction as a function of interaction with the system introduces a higher variation of speed across drivers if this would happen during real world driving. This would reduce traffic safety on an aggregated level.
In addition, the interaction with the IVIS application resulted in smaller TLC minima. The distribution of TLC minima shifted towards smaller values. In this the smaller TLC minima were more affected than the larger TLC minima. This basically means that the distribution gets broader towards smaller values when interacting with the RDS-TMC application tested here. This process occurs especially during forced-paced conditions where the subject has no opportunity to slow down.
However, it also occurs when the subject does have the opportunity to slow down. In terms of the aims of the project this would indicate that the map display version of the RDS-TMC needs to be tested more thoroughly in an advanced simulator or on the road before being introduced to the market.

5.2. **Laboratory Test Results by COV**

The discussion in this paragraph is based on the report:

- Brookhuis, K.A., Waard, D. de (1997). *A validation study of a PC-based test of safety aspects of in-vehicle information systems; A test of an auditory message version of an RDS-TMC task.*

5.2.1. *Introduction*

Below a first validation of the PC- test is described for an auditory version of RDS-TMC. During a part of the driving task, subjects are required to listen closely and respond to a few auditory messages. The most important indicators of driving performance used in the test are longitudinal control, i.e. vehicle speed, and lateral control, i.e. SD lateral position, minimum time-to-line crossing and actual lane boundary crossings

5.2.2. *Method*

The primary, driving, task runs on a 200 MHz Pentium PC with a steering wheel, an accelerator pedal, a brake pedal and two response buttons as input devices that give their signals to the computer program via a bus mouse or a game port. The images are presented on a 15 inch VGA monitor in colour. The scene consists of a one-lane road with a lane width of 3.6 m. The road consists of alternating curved segments to the left and to the right.
The horizon is drawn at half the screen hight.
Speed is controlled by the accelerator and brake pedals and restricted to a maximum of 120 km/h. The present speed is indicated by a moving horizontal bar at the bottom of the screen. Lateral position is controlled by the steering wheel. The driver control actions are the input for a simple vehicle model that has an engine model with an automatic gear box and a model of vehicle dynamics. Engine sound, depending on engine rpm, is delivered to the driver via an external loudspeaker.

An auditory secondary task (In Vehicle Information System, IVIS task) had to be combined with the primary task of lateral and longitudinal control.
The task was an adapted, i.e. computerised, version of the RDS-TMC traffic congestion information system. The IVIS-LabTest controls the interaction with the secondary IVIS task. It reads an input file that specifies the name of the output data file, the driving speed, the conditions of the secondary task (auditory RDS-TMC messages), the number of stimuli (congestion messages) for the secondary task, the position in the road environment where the stimuli for the secondary task are generated and the type of stimulus. If the specified driving speed is anything other than 120 km/h, then the speed is fixed at the specified speed. In that case the accelerator and brake pedal inputs are ignored by the vehicle model. This constitutes the forced-paced condition. If specified speed is 120 km/h, then the driver is free to choose the preferred speed. This constitutes the self-paced condition. If the number of stimuli in the input file is zero then the task is performed without interaction with an IVIS (no-IVIS condition). Otherwise the IVIS condition applies.

5.2.3. *Results and conclusions*

Thirteen male and twelve female subjects completed the tests.
The results show no effects of the auditory task tested in the present experiment on driver behaviour as measured by adopted speed and minimum time-to-line crossing. Interaction with the secondary task that induces an auditory load resulted in a difference of the minimum time-to-line crossing between force-paced and self-paced speed conditions only. In addition, the interaction with the IVIS application resulted in smaller SDs of lateral position. In the present experiment no manual operation was necessary, implying that subjects had to divert their visual attention from the 'road' only in the visual set-up (Van Winsum, 1997) and could keep their gaze fixed on the 'road' in the auditory set-up.

Subjects indicated higher mental workload while engaged in the secondary task, and also, rated the self-paced condition as more effortful.

6. A safety checklist

6.1. Introduction

In the literature study it was established that three major effects on driver behaviour may cause adverse safety effects: task overload, task underload and some forms of long term behavioural adaptation. As it turns out, most of the knowledge available concerns task overload and significantly less pertains to task underload. Both factors have yielded a mixture of practical criteria and more theoretical considerations that could be included in a preliminary checklist. About long term behavioural adaptation, practically no directly applicable knowledge was found however. For that reason 'counterproductive behavioural adaptation' has been treated as a special research topic to produce recommendations for a procedure to construct a checklist rather than a checklist itself.
Furthermore, an additional problem was considered that might arise from uncoordinated but simultaneous application of different ATT devices. Such an arbitrary combination of functions may produce confusing or even contradictory information that can also lead to unsafe driving situations. Because the interaction between various aspects of two simultaneously operating systems leads to a number of subsequent 'if- then' decisions, this has been set up as a flow chart rather than as a sequential checklist.
The results of this part of the project are summarised in the following paragraphs.

It should be emphasised that this checklist is a first attempt to organise knowledge about safety effects in a comprehensive way. Furthermore, this kind of knowledge itself is far from complete. We may actually have addressed only a limited part of all possible behavioural effects that ATT may evoke and the results should be regarded accordingly.

6.2. A safety checklist based on taskload considerations

This first version of a checklist distinguishes three types of tasks-loads:
a. mental taskload;
b. visual taskload;
c. physical taskload.

For each of these three types, some criteria have been formulated for both overload and underload and itemised into separate questions. These questions may sometimes seem strangely formulated but this has been done to maintain the same logical structure throughout the list which is: *an affirmative answer to any question signals a potential problem.*

6.2.1. *Mental taskload*

Overload checking

- are any messages exclusively system-paced an short-lived and can not be repeated or switched off at drivers request;
- do any messages require extended decision making;
- are any messages confusing or ambiguous;

- is control of the system context dependent => are there multi level menus;
- do any verbal messages not comply to:
  - the use of commonly familiar words only;
  - a limited set of phrases.

Underload checking

- does the system stimulate driving at night;
- does the system tempt the driver to abandon resting;
- does the system affect behaviour when the driver is in a unfavourable state (fatigue, drugged state).

6.2.2. *Visual taskload*

Overload checking

- do any visual messages require more than three glances of at most 1 s;
- can visual messages be seen well in extreme lighting conditions (at night, in heavy sun) => is there *no* automatic adaptation to external lighting conditions;
- does any visual display *fail* to comply with any of the legibility conditions:
  - viewing distance 70-75 cm;
  - character height 6,4 mm or larger;
  - minimum 5x7 matrix per character;
  - character width- height ratio 0,7-0,8;
  - horizontal character spacing 75% of character width;
  - vertical spacing 35%-100% of character height;
  - use only simple fonts without serifs and italics;
  - use only capitals on messages longer than 3 words.

Underload checking

- does the system take care of obstacle detection;
- does the system take care of signal input.

6.2.3. *Physical task load*

Overload checking

- is sometimes immediate manual control required (e.g. deactivating an alarm);
- is the loudness of the message outside the following limits:
  - 15-25 dB over background noise;
  - a maximum level of 115 dB;
- are the alarms used outside the following specifications:
  - frequency range 500 -2000Hz;
  - repetition rate 1-8 1/sec;
  - non-speech messages only;
- are some controls difficult to reach or to handle;
- are some controls difficult to identify;
- do some controls require visual feedback to operate (e.g touch screens);
- are any haptic messages confusing or ambiguous; is the sensory message always distinguishable from random environmental inputs.

Underload checking

- does the system take over pedal control;
- does the system take over part of manual control.

It should be stressed that any single affirmative answer means that an adverse safety effect is *possible*, not that it certainly will occur. This implies that the final decision on the necessity of further testing or even complete acceptation or rejection of the device can only take place after a judgement of the total outcome of all checks. Therefore the application of this (or any other) checklist still requires expert knowledge.

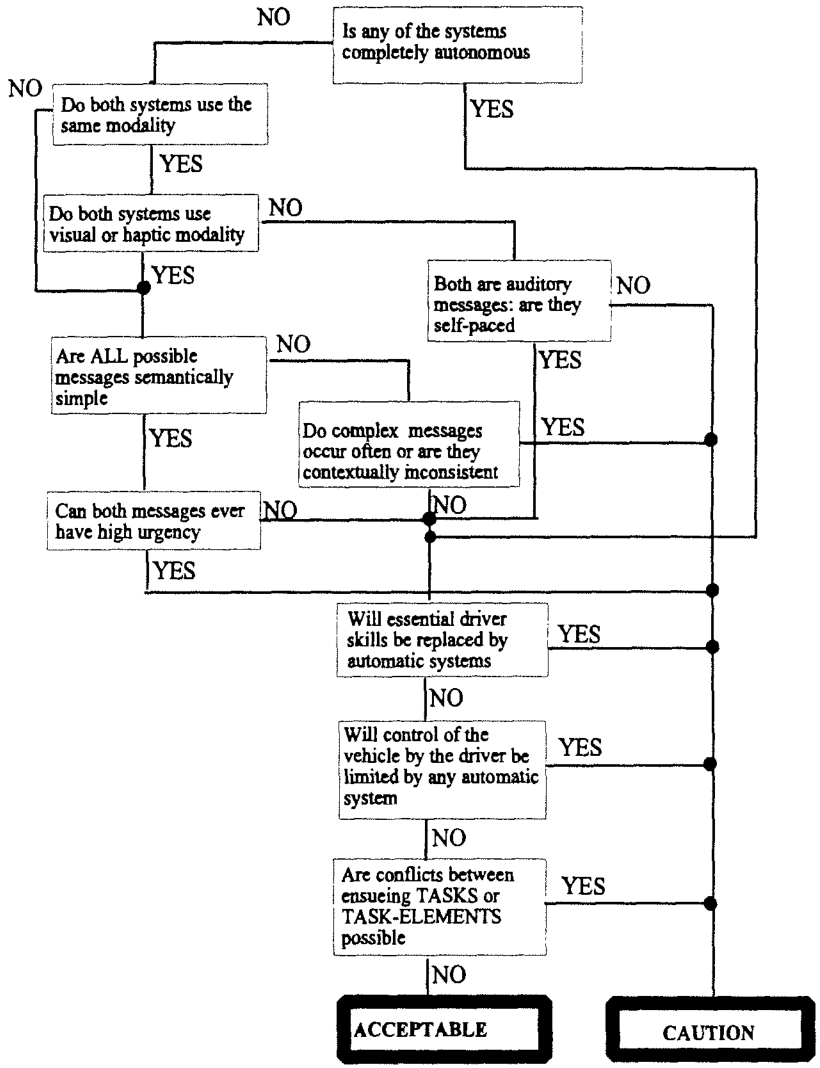

Figure 1. *A checklist for the assessment of interference of two systems.*

6.3. Checking system interference

If there exists a reasonable probability that the ATT under scrutiny may be applied simultaneously with an other ATT application that have not been designed as a single functional unit, the consequences of possible undesired functional interference must be assessed as well as possible. In the course of

this project a tentative checklist, pictured below, has been proposed to this end. In case this assessment leads to the verdict 'caution' special measures, like prioritising the messages of the separate systems, are recommended to avoid the interference.
As stated before, this checklist has consists of a number of conditional questions that cannot easily be accommodated by a simple list of questions. For this reason the form of a flow-chart has been chosen: this chart is depicted in *Figure 1.*

6.4. **Counterproductive adaptation**

6.4.1. *Introduction*

Behavioural adaptation to a changing environment is an everyday occurrence. Generally speaking this adaptation can be considered a survival trait of the species and as such cannot be termed 'counterproductive'. Still, under certain circumstances this counterproductivity may manifest itself, especially when the adaptation takes the form of losing rarely used (but sometimes necessary) knowledge or skills. Also, adaptation may cause compensatory behaviour in the sense that a perceived gain in safety margin is 'compensated' by more risky behaviour. In the initial study a few and criteria for possible counterproductive behavioural adaptation could be found to contribute to the checklist. They can be formulated as follows:

- is the device explicitly designed and presented as contributing to safer driving;
- is the device explicitly presented as something that watches over you;
- is the effect of the device continuously present in the driver's task environment.

These criteria a qualitative only and it is felt that these checks are still inadequate to account for a potentially important hazard. However, with in the time available for this project, only a small scale additional study could be implemented. As a consequence this small scale study did not produce an actual checklist on this item but only an inventory of possible components and their mutual relations. In order to prepare such an inventory, an expert meeting was instigated to provide a first basis from psychological principles. After the expert session an attempt was made to combine these principles with insights from the theory of Situation Awareness, to obtain more detailed descriptions of possible adaptation mechanisms.

6.4.2. *Results from the expert meeting*

The expert meeting resulted in a sort of formula for an approach to the problem. This approach can be formulated as follows:

1. Consider the characteristics of a specific ATT system: ATT system may differ vastly and different characteristics are expected to address or induce different forms of adaptation. Relevant characteristics may be:

- type of telematic system:
  - advisory;
  - suggesting/supervising;
  - controlling or autonomous;
- reliability of the system;
- the internal consistency of operation of the system;

- relevancy of the functionality of the system (for driver and road authority);
- nature and frequency of the feedback the system provides;

2. Consider the generalised behavioural adaptation mechanisms in the light of the ATT system's characteristics (this will be elaborated in the following section).

3. To assess a total safety effect, the consequences of the adaptations must be considered on several levels of the traffic environment:

- effects on mobility;
- effects on the traffic flow;
- effects on interactions of traffic participants;
- effects on the individual driving task, traditionally distinguishing:
  - strategic level;
  - tactical level;
  - operational level.

Step 2 of this procedure, a search for generalised behavioural adaptation mechanisms, has been the focus of the rest of this short study.

6.4.3. *Further results*

In order to assess the relevance of the characteristics in the previous paragraph for long term behavioural adaptation, a model will be applied. Such a model orders at least a large portion of the items in the list and provides insight into the (supposed) causal relationships between these characteristics. In our case, a model of Situation Awareness, that is proposed by Endsly (Endsley, 1988) *Design and evaluation for situation awareness enhancement*) was chosen as a reference. It is a model of the human controller as operator interacting with a well defined man-machine interface and participating in a process with a limited number of degrees of freedom. Under such conditions it is often assumed that the human operator's behaviour can be modelled on the basis of a set of internal models called schemata and scripts. The schemata are supposed to contain a simplified representation of dominant characteristics of the surrounding process and there causal relationships; here vehicle characteristics, general behavioural characteristics of other road user and the state of traffic in a limited vicinity. Scripts represent more or less automated sequences of actions.
Schemata interact with all other functions like perception, interpretation, comprehension, projection, decision making, action guidance and scripts. This interaction is reciprocal in the sense that schemata are considered on one hand to supply references for these functions and on the other hand to be modified or build by the results of these functions.
In this way, perception will be largely focussed on the characteristics that are most important in the currently active schema. The operator will than attempt to interpret the perceived characteristics in terms of the causal relations in the schema and if this succeeds (=comprehension) the schema will also be employed to generate short term predictions of the future state of the (traffic) system. Based on this prediction, actions and corresponding scripts will be chosen and executed.
If comprehension, prediction or action fail, the operator can choose another schema (if available) or implement an emergency script.

Both schemata and scripts are memory structures that can be learned, forgotten and changed. These modification processes will take place in every person but potentially in a somewhat different manner for each because various mechanisms that may be instrumental will always be based upon personal experiences and circumstances. Examples of those modification processes are:

- consistently and frequently occurring related events of various nature;
- events that occur rarely can hardly be learned by experience, but can be adapted to by inference by functional parallels or analogues or attribution.

Furthermore, part of the control of traffic behaviour depends on:

- rewards and punishments;
- emotional state.

It must be clear that individual behavioural adaptation will be very hard to predict since a significant part of those adaptations depend on personal circumstances. We may be able to define and predict a common part however by investigating the mutual relationships between schemata, scripts and perception. First, we can try to determine common elements in the schemata and scripts that road users employ. Subsequently, we can use these common characteristics to derive 'basically necessary' knowledge and skills that should not be impaired by adaptation as a template to define counter productivity. Finally we can try to categorise the influence of ATT devices according to:

- which of the common elements is affected and how (supported, enhanced, substituted, changed etc.) to establish the nature of the effect;
- does the effect occur rarely, frequently or continuously to establish the possible severity of the effect.

These three steps should at least lead to the a more elaborate set of checks than is currently available, but completeness cannot be claimed.

7. Conclusions and recommendations

7.1. General conclusions

The aims of the project were to construct comprehensive procedures and criteria that allow the prediction of possible adverse safety effects of various types of new or existing In Vehicle Telematics.
Given the state of the art of knowledge on this subject, these aims can be considered rather ambitious. Still, it can be concluded that this project has yielded useful and encouraging results that allow a more structured approach to testing safety effects of telematics in vehicles because:

1. A first version of a checklist has been established to obtain a first screening on safety effects; this checklist is predominantly based on safety criteria for taskload (over- and underload).
2. A laboratory test has been developed that enables practical testing of a range of In Car telematic devices. The test emulates a simplified driving task on a PC and that provides an interface to emulate, control and assess the safetyeffects of these telematics; this test is also set up to be used by relative non-experts
3. As was expected, both the checklist and the laboratory test do not cover all possible safety effects sufficiently and field testing of new Telematics will often still be indicated. This research has provided a useful structure for the setup and evaluation of such tests which allows comparable and consistent assessment.

It can be concluded that, in the light of still limited knowledge, the proposed prodedure of stepwise safety checking (first a checklist screening followed by laboratory testing and/or field testing of doubtful device characteristics or devices not included in the previous steps) seems an efficient approach. It should therefore be developed to the full extent of available knowledge.

7.2. Conclusions regarding criteria for field testing

The general finding is that, although an extensive array of objective measures for the safety of driving behaviour is available, subjective measures in the form of expert opinions still remain indispensible.
The main reason for this is that well defined criteria still lack for many measures and for some measures the research results seem contradictory.

7.3. Conclusions regarding the construction of a checklist

The results so far have yielded components for a checklist for safety effects on three items: task overload, task underload and counterproductive behavioural adaptation. Of the three, the first item 'task overload' has been investigated most and therefore produced the most substantial results. Task underload accounts for significantly fewer criteria and criteria for behavioural adaptation must be considered to be 'in statu nascendi'. Also, the problem is recognised that arises when two independent ATT systems operate simultaneously without coordination of their interactions with the driver. A separate checklist, based upon the present findings was conceived to determine possibly dangerous interference of such systems.

The total of the results cannot be considered definitive in any way: the checklists can be used to perform a first screening of proposed or existing ATT devices but certainly not to obtain a definitive verdict. Therefore, both laboratory tests and full scale tests will often still be necessary.

7.4. Conclusions regarding the laboratory PC-test

The PC laboratory test developed in the course of this project has been subjected to two validation tests: one by TNO-HFRI and one by COV.

7.4.1. *The results of the validation test by TNO-HFRI*

The results of the experiment suggest that the IVIS-LabTest in its present form is sensitive enough to show an effect of a specific IVIS application on primary task performance. Whether this result can be generated to real world driving depends on the validity of the IVIS-LabTest and the fidelity of the test. Fidelity can be distinguished into physical fidelity and functional fidelity. *Physical fidelity* is related to the extent to which the test resembles real car driving in terms of information presentation and how the system behaves.
Given the requirement that the test should be PC based, the physical fidelity is reasonable: the behaviour of the vehicle and engine model in the test results in control characteristics that resemble those of a real car, although this should be tested more systematically. Also the information presentation resembles the way the driver receives information in real world driving while performing a lane control task: the roadview is presented from the perspective of the driver in a realistic 3D-format. *Functional fidelity* relates to the similarity in behaviour in the test and while driving a car in the real world under similar conditions. This would have to be investigated more thoroughly but there is some evidence for behavioural similarities.
For example, on first sight the characteristics of the TLC signal measured in the test appear to resemble TLC characteristics as measured in advanced driving simulators and while driving on the road in terms of magnitude of minima and duration of phases to left and to right. Also, in real word driving a higher task load has been associated with a decrease of vehicle speed. This was also found in the present study. In addition, the behavioural requirements of the IVIS-LabTest resemble the lane following task in real world driving in terms of lateral and longitudinal control requirements.

Furthermore, whether it is valid to make safety inferences from this test is ultimately an empirical question. The theoretical view presented here has focussed on effects on operational performance (lateral control) and choice of speed. The first was associated with safety effects for the individual driver while the second was associated with effects on the traffic system surrounding the individual driver who uses the system. There are other possible effects of this specific IVIS application on safety that have not been examined in the present study, such as effects on route choice or on choice of means of transportation. These effects may have safety consequences that cannot be overseen at present.

7.4.2. *The results of the validation test by COV*

In the present experiment an auditory secondary task (In Vehicle Information System, IVIS task) had to be combined with the primary task of lateral and longitudinal control. The task was an adapted, i.e. computerised,

version of the RDS-TMC traffic congestion information system. Verwey (1996) reported effects of both visual and auditory messages on expert judgements of driving capacity. Contrary to the results in the companion experiment with visual messages of Van Winsum (1997), in the present experiment no effects of auditory messages were found on time-to-line crossing.
Two obvious differences between the experiment of Verwey (1996) and the present experiments' set-up could be held responsible for this finding.
The first is that in Verwey's experiment the effects were not on objective performance measures but only on subjective assessments by trained observers, and, subjects had to operate the RDS-TMC system manually in all cases. In the present experiments no manual operation was necessary, implying that subjects had to divert their visual attention from the 'road' only in the visual setup and could keep their gaze fixed on the 'road' in the auditory setup.

In terms of the aims of the project this would indicate that, perhaps contrary to earlier findings (IVIS-report by Verwey, 1996), the auditory version of the RDS-TMC is less intruding than the visual map version, however, this needs to be tested more thoroughly in an advanced simulator or on the road before being introduced to the market.

7.4.3. *Comparison of the results*

Where the results of the validation tests of visual IVIS systems conducted by TNO-HFRI seem to be in concordance with earlier research by Verwey (IVIS-report by Verwey, 1996) the results of tests of auditory IVIS systems seem to contradict earlier findings of Verwey. Closer examination of the experiments conducted by Verwey however, reveals that there is a, possibly very significant, difference in testing circumstances between the two experiments. Although in both tests exactly the same RDS-TMC modes and messages were employed, the setup that Verwey used required test subjects to respond to an RDS-TMC message by *manually selection from a menu display* whereas the test conducted by COV requiered *no manual response task.* Since a manual responsetask demands a considerable amount of attention, much of the adverse effects found could have been caused by the interference of this manual task.
This means that the apparent contradiction can be resolved and that the conclusion must be that auditory RDS-TMC displays are indeed less intrusive, and therefore safer, that visual displays.
It also means that the PC-test still can be considered a promising method for investigating safety effects of ATT devices.

7.5. **Recommendations**

As stated before, this project provides a procedural model for the evaluation of safety effects of ATT applications and some preliminary versions of tools to fit into that model. The procedure has three steps and in these recommendations the three steps will be examined separately for proposed future development.

7.5.1. *Step 1: a safety checklist*

Generally speaking the checklist contains criteria for the following areas:
1. interface ergonomics (Human Machine Interface);
2. direct effects on the human driving task;
3. effects on the human mental and physical condition.

As the checklist should contain conditions and effects that are relatively well investigated it is recommended that short term further development be aimed at incorporating as much of existing knowledge as possible. Since there are a number of ongoing international research projects that also address safety of telematics and in which some of the current partners participate, it is strongly recommended to first examine these projects for useful results.
For the longer term, we may have to intiate further research into the areas 2 and 3.
However, in order to do this, we first need further development of experimental evaluation tools, described in step 3.

In addition, the problems regarding counterproductive behavioural adaptation are only superficially covered by the current checklist and will need more substantiation. Several approaches seem to be available to this end:
- to conduct a long term observation experiment along with a current project like the ISA and AHS experiments;
- to conduct an intermedium term experiment in a tightly controlled environment like a simulator.

7.5.2. *Step 2: a PC test*

The current PC test, actually a tracking task which looks like a simplified traffic simulator, has certainly shown promise as a ow cost evaluation tool. A first attempt to validate the test against an existing field experiment has been carried out in two separate tests and seems to have produced satisfactory results. Though promising, it is still a preliminary version, and more extensive testing against field tests with a variety of ATT devices is an urgent necessity to ascertain the validity of the program in order to warrant general use.

Moreover, the current version of the test is only suitable to assess the effects, on driving perfromance, of ATT devices that somehow provide information to the driver. Devices that directly interact with functions of the vehicle, e.g. intelligent cruise control AICC, can not be tested yet.
Further development of the PC simulator test to incorporate also this class of devices is conceivable and certainly desirable.

7.5.3. *Step 3: criteria for field testing*

As has been described, there are many measurable parameters that may be relevant to the assessment of performance on the driving task. It has also been stated that the usefulness of many of these parameters is limited because reference values have only recently been started to be established. For that reason it has been advised to incorporate expert judgement in the observation procedure as well, at least for the time being. This makes (international) comparison of results difficult, however.

Therefore, further research in this field will certainly have to be aimed at establishing:

- what parameters must be measured in which type of experiment;
- what are the reference values/ranges for these parameters.

7.5.4. *Further recommendations*

There is a special aspect associated with behavioural adaptation that probably deserves special attention, i.e. complacency. Complacency is a serious potential problem in this respect, since the likelihood of system failure may be low for the moment, hopefully, but can not be neglected. Besides, a side-effect of complacecny might be loss of certain skills that could be necessary in other conditions.

Many of the presently developed ATT systems can take over tasks, which the driver will find useful and therefore will adapt driving behaviour to exploit the device. Sudden failure of the device, which is, given the increasing complexity of the ATT devices certainly not inconceivable in the near future, may give rise to acute dangerous situations. These safety effects and their possible remedies, need to be investigated further.

Literature

Akerboom, S.P. (1989). *The understanding and recollection of radio traffic information; Dissertation.* Leiden University, Leiden.

Akerstedt, T., Tborsvall, L. (1984). *Continuous electrophysiological recordings.* In: J.J. Cullen & J. Siegriest (Eds.), Breakdown in Human Adapatation to Stress. Towards a multidisciplinary approach, Vol.I. Martinus Nijhoff, The Hague, pp. 567-584.

Alexander, G.J., & Lunenfeld, H. (1986). *Driver expectancy in highway design and traffic operations.* Report FI-IWA-TO-86-1. Department of Transport, Washington, D.C.

Alink, G.M.M. (1992). *Snelheidsbegrenzer voor de vrachtauto.* In: Verkeerskunde 3(22), pp. 24.

Allen, T.M., Lunenfeld, H. & Alexander, G.J. (1971). *Driver information needs.* In: Highway Research Record 366, pp. 102-115.

Almqvist, S, Hyddn, C. & Risser, R. (1991). *Use of speed limiters in cars for increased safety and a better environment.* In: Transportation Research Record 1318. Transportation Research Board, Washington D.C.

Antin, J.F., Dingus, T.A., Hulse, M.C. & Wierwille, W.W. (1990). *An evaluation of the effectiveness and efficiency of an automobile moving-map navigational display.* In: International Journal of Man-Machine studies 33, pp. 581-594.

Bos, J.M.J. (1996). *Routewijzigingen en verkeersveiligheid; Schatting van het verkeersveiligheidseffect van verkeersinformatie-geïndiceerde routewijzigingen met behulp van enquêtegegevens uit het Rijn-Corridor-project.* R-96-25. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam.

Briem, V. & Hedman, L.R. (1995). *Behavioural effects of mobile telephone use during simulated driving.* In: Ergonomics 38, pp. 2536-2562.

Brookhuis, K.A. (1991). *Driver monitoring, het 'bewaken' van de toestand van de chauffeur.* In: Proceedings CVS-symposium DRIVE - De Nederlandse inbreng. Colloquium Vervoersplanologisch Speurwerk, Rotterdam.

Brookhuis, K.A. (1995). *De invloed van psychoactieve stoffen op de aandacht.* In: P. Eling & W. Brouwer (Eds.), Aandachtsstoornissen, een neuropsychologisch handboek. Swets & Zeitlinger, Lisse, pp. 217-229.

Brookhuis, K.A. & Brown, I.D. (1992). *Ergonomics and road safety.* In: Impact of science on society 16, pp. 35-40.

Brookhuis, K.A. & Waard, D. de (1993). *The use of psychophysiology to assess driver status.* In: Ergonomics 36, pp. 1099-1110.

Brookhuis, K.A., Louwerens, J.W. & O'Hanlon, J.F. (1986). *EEG energy-density spectra and driving performance under the influence of some antidepressant drugs.* In: J.F. O'Hanlon & J.J. De Gier (Eds.), Drugs and Driving. Taylor & Francis, London/Philadelphia.

Brookhuis, K.A., Volkerts, E.R. & O'Hanlon, J.F. (1990). *Repeated dose effects of lormetazepam and flurazepam upon driving performance.* In: European Journal of Clinical Pharmacology 39, pp. 83-87.

Brookhuis, K.A., Vries, G., de, Prins van Wijngaarden, P. & O'Hardon, J.F. (1985). *The effects of increasing doses of meptazinol (100, 200, 400 mg) and glafenine (200 mg) on actual driving perfomance.* Report VK 85-16. Traffic Research Centre, University of Groningen, Haren.

Brouwer, W.H. & Wolffelaar, P.C. van (1985). *Sustained attention and sustained effort after closed head injury: detection and 0.10 Hz. heart rate variability in a low event rate vigilance task.* In: Cortex 21, pp. 111-119.

Brown, I.D. (1982). *Driver fatigue.* In: Endeavour 6, pp. 83-90.

Coles, M.G.H. (1983). *Situational determinants and psychological significance of heart rate change.* In: A. Gale & J.k Edwards (eds.), Physiological correlates of human behavior, Vol.11: Attention and performance. Academic Press, London, pp.171-185.

Dewar, R. (1992). *Driver and pedestrian characteristics.* In: J.L. Pline (ed.), Traffic Engineering andbook. Prentice Hall, New Jersey, London, pp. 1-28.

Dingus, T.A., Antin, J.F., Hulse, M.C., & Wierwille, W.W. (1989). *Attentional demand requirements of an automobile moving-map navigation system.* In: Transportation Research 23A, pp. 301-315.

Draskoczy, M. (1993). *Mandatory Safety Quality Assurance.* In: Hopes Deliverable 11. University of Lund, Department of Traffic Planning and Engineering, Lund.

Endsley, M.R. (1988) *Design and evaluation for situation awareness enhancement.* In: Proceedings of the Human Factor Society 32nd Annual Meeting (pp. 97-101). Santa Monica. CA: Human Factors and Ergonomics Society.

Endsley, M.R. (1988) *Situation Awareness Global Assessment Technique (SAGAT).* In: Proceedings of the National Aerospace and Electronics Conference. IEEE, New York, pp. 789-795.

Engels, K. & Dellen, R.G. (1983). *Beitrag zur Quantifizierung des Altesrisiko von PKW-Fahrem.* In: Unfal- und Sicherheitsforschung Strassenverkehr 42. Bundesanstalt fiir Strassenwesen, Koln.

Evans, L. (1987a). *Belted and unbelted driver accident involvement rates compared.* In: Journal of Safety Research 18, pp. 57-64.

Evans, L. (1987b). *Estimating fatality reductions from increased safety belt use.* In: Risk Analysis 7, pp. 49-57.

Evans, L. (1991). *Traffic safety and the driver.* Van Nostrand Reinhold, New York.

Fairclough, S.H., Brookhuis, K.A. & Vallet, M. (1993). *Driver state monitoring system DETER (V2009).* In: Advanced Transport Telematics, Proceedings of the Technical Days. Commission of the European Communities, Brussel, pp. 330-335.

Fastenmaier, W. (ed.) (1995). *Autofahrer und Verkehrssituation.* Kdln: Tuv Rheinland.

Fontaine, H. (1988). *Wie krijgt wettelijk de schuld bij verkeersongevallen? (Who is legally blamedfor traffic accidents.)* In: Verkeerskunde 39, pp. 101-103.

Fuller, R.G.C. (1984). *Prolonged driving in convoy: the truck driver's experience.* In: Accident Analysis & Prevention 16, pp. 371-382.

Galsterer, H., Fasten-meier, W., & Gstalter, H. (1990). *Sicherheitsaspekte von LISB (Leit- und Informationssytem Berlin)-Fahrten auf Routinestrecken; Forschungsbericht.* In: Auftrag der BMW AG, Verkehrsforschung.

Gier, J.J. de (1980). *Evaluation of drugs in real driving situations.* Thesis, University of Utrecht, Utrecht. [in Dutch]

Godthelp, J. (1984), *Studies on human vehicle control.* PhD Thesis, Institute forPerception, INO, Soesterberg.

Godthelp, J. (1988). *The limits of path error-neglecting in straight lane driving.* In: Ergonomics 31, pp. 609-619.

Godthelp, J., Milgran P. & Blaauw, G.J. (1984). *The development of a time-related measure to describe driving strategy.* In: Human Factors 26, pp. 257--268.

Godthelp, H. & Schumann, J. (1991). *The use of an intelligent accelerator as an element of a driver support system.* In: Proceedings of the 24th ISATA International symposium on automotive technology and automation. Automotive Automation Limited, Croydon, pp. 615-622.

Haas, I. & Reker, K. (1976). *Einfluss von Lebensalter und Fahrerfahrung auf die Unfallsbelastung und Sicherheitsrelevante Einstellung.* In: Verkehrssicherheit 22, pp. 119-130.

Hagen, J.A.M. & Fokkema, H.J. (1990). *Snelheidsbegrenzers in personenauto's.* Rapport IT 90-4. Traffic Test, Veenendaal.

Hancock, P.A. & Parasuraman, R. (1992). *Human factors and safety in the design of Intelligent Vehicle-Highway Systems (IVHS).* Journal of Safety Research, 23, 181-198.

Hancock, P.A, Wulf, G., Thorn, D., & Fassnacht, P. (1990). *Driver workload during differing driving maneuvers.* In: Accident, Analysis & Prevention 22, pp. 281-290.

Harms, L. (1986). *Drivers' attentional responses to environmental variations; A dual-task real traffic study.* In: A.G. Gale, M.H. Freeman, C.M. Haslegrave, P. Smith, S.P.Taylor (Eds.), Vision in Vehicles. North-Holland, Amsterdam, pp. 131-138.

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. Meshkatie (Eds.), Human Mental Workload. North-Holland, Amsterdam, pp. 139-183.

Heijer, T. (1997). *A safety checklist for ATT devices.* R-97-19. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam.

Henderson, R. & Burg, A. (1974). *Vision and audition in driving.* Final report DOT HS801-265. National Highway Traffic Safety Administration.

Hendy, K.C., Hamilton, K.M., & Landry, L.N. (1993). *Measuring subjective workload: When is one scale better than many?* In: Human Factors 35, pp. 579-601.

Horne, J.A., Wilkinson, S. (1985). *Chronic sleep reduction; Daytime vigilance performance and EEG measures of sleepiness with particular reference to 'practice' effects.* In: Psychophysiology 22, pp. 69-78.

Horst, A.R.A. van der (1990). *A time-based analysis of road user behaviour in normal and critical encounters.* Unpublished doctoral dissertation. TNO Institute for Perception, Soesterberg.

Janssen, W.H. (1994). *Seat-belt wearing and driving behaviour: an instrumented-vehicle study.* In: Accident Analysis and Prevention 26, pp. 249-261.

Janssen, W.H. & Horst, A.R.A., van der (1992). *A framework for the prediction of the road safety effects of new technology.* report IZF 1992 C-37. TNO Institute for Perception, Soesterberg.

Joksch, H.C. (1993) *Velocity change and fatality risk in a crash - a rule of thumb.* In: Accident Analysis & Prevention 25, pp. 103-104.

Kampen, L.T.B. van (1988). *Analyse van de verkeersonveiligheid van jonge, onervaren automobilisten. (Analysis of traffic accident risk of young, inexperienced drivers).* R-88-45. SWOV Institute for Road Safety Research, Leidschendam.

Kantowitz, B.H. (1987). *Mental workload.* In: P.A. Hancock (Ed.), Human Factors Psychology. North-Holland, Amsterdam.

Kaptein, N.A., Verwey, W.B., & Horst, A.R.A. van der (1993). *Gedragseffecten van navigatiesystemen in relatie tot efficientie en verkeersveiligheid: Een literatuuroverzicht en een voorstel voor een standaard methode.* [In Dutch: Behavioral effects of in-car navigation systems in relation to efficiency and traffic safety: A literature review and an outline for a standardized testing method.] Report IZF 1993 C-39. TNO Institute for Perception, Soesterberg.

Katteler, H. (1996). *Uitvoeringsvormen van RDS/TMC.* Instituut voor Toegepaste Sociale Wetenschappen, Nijmegen.

King, G.F. & Lunenfeld, H. (1971). *Development of information requirements and transmission techniques for highway users.* National Cooperative Highway Research Program report 123.

Knippenberg, C.W.F., Rothengatter, J.A. & Michon, J.A. (1989). *Handboek Sociale Verkeerskunde.* Van Gorcum, Assen.

Kok, K. & Brookhuis, K.A. (1995). *Specifications of the integrated system.* Report V2009/DETER/Delivemble 19 (331B, 340B, 410A, 420A, 43OA) to the commission of the European Union. Traffic Research Centre, University of Groningen, Haren.

Laan, J.D. van der, Heino, A. & Waard, D. de (submitted). *A simple procedure for the assessment of acceptance of advanced transport telematics.*

Lim, C.S. & Dewar, R.E. (1988). *Cognitive ability and traffic acidents.* In: Proceedings of the tenth congress of the International Ergonomics Association, Sydney, pp. 593-595.

Lisper, H.O., Laurell, H. & Vanloon, J. (1986). *Relation between time to falling asleep behind the wheel on a closed track and changes in subsidiary reaction time during prolonged driving on a motorway.* In: Ergonomics 29, pp. 445-453.

MacDonald, W.A. (1979). *The measurement of driving task demand.* Corvina, Melbourne.

MacDonald, W.A. & Hoffman, E.R. (1980). *Review of relationships between steering wheel reversal rate and driving task demand.* In: Human Factors 22, pp. 733-739.

McLean, J.R. & Hoffmann, R. (1975). *Steering reversals as a measure of driver performance and steering task difficulty.* In: Human Factors 17, pp. 248-256.

Mede, P.H.J. van der (1992). *Snelheidsbegrenzer voor vrachtauto's in bedrijfseconomisch perspectief.* In: Verkeerskunde 3, pp. 24-27.

Michon, J.A. (1985). *A critical review of driver behaviour models: what do we do, what should we do.* In: L. Evans & R.C. Schwing (Eds.), Human behaviour and traffic safety. Plenum, New York, pp. 485-520.

Mulder, L.J.M. (1992). *Measurement and analysis methods of heart rate and respiration for use in applied environments.* In: Biological Psychology 34, pp. 205-236.

Munden, J.M. (1967). *The relation between a driver's speed and his accident rate.* TRL, Crowthorne.

Neisser, U. (1976). *Cognition and Reality.* W.H. Freeman and Co.

Nilsson, L. & Berlin, M. (1992). *Driver attitudes and behavioural changes related to presentation of roadside information inside the car. A pilot of the CAROSI system.* Report 698A. Swedish Road and Traffic Institute (VII), Linköping.

Noy, Y.I. (1987). *Theoretical review of the secondary task methodology for evaluating intelligent automobile displays; Proceedings of the 31th Human Factors Society, Santa Monica, CA.* In: Human Factors Society, pp. 205-209.

Noy, Y.I. (1990). *Attention and performance while driving with auxiliary in-vehicle displays.* Publication TP 10727 E. Transport Canada, Ottawa, Ontario.

O'Hanlon, J.F., Haak, T.W., Blaauw, G.J. & Riemersma, J.B.J. (1982). *Diazepam impairs lateral position control in highway driving.* In: Science 217, pp. 79-81.

Palmquist, U. (1993). *Intelligent Cruise Control and roadside information.* IEEE Micro, February 1993, pp. 20-28.

Parasuraman, R. (1985). *Sustained attention: a multifactorial approach.* In: M. Posner & O.S.M. Marin (Eds.), Attention and Performance XI: Mechanisms of attention. Lawrence Erlbaum Ass., Hillsdale, pp. 492-511.

Parkes, A. (1989). *Changes in driver behavior due to two modes of route guidance information presentation: A multi-level approach.* HUSAT Research Centre, Loughborough.

Parkes, A.M. (1991). *The effects of driving and handsfree telephone use on conversation structure and style.* In: Proceedings of the 24th Conference of the Human Factors Association of Canada.

Parkes, A.M. & Ross, T. (1991). *The need for performance based standards in future vehicle man-machine intefaces.* Advanced ATT in road transport (Vol. II, pp. 1312-1321). In: Proceedings of the Drive conference, Brussels, February 1991. Elsevier, Amsterdam.

Pelz, D.C. & Schuman, S.H. (1971). *Are younger drivers really more dangerous after controlling for exposure and experience?* In: Journal of Safety Research 3, pp. 68-79.

Persson, H., Towliat, M., Almqvist, S., Risser, R. & Magdeburg, M. (1993). *Hastighetsbegrdnsare i bib. Fdltstudie av hastigheter, beteenden, konflikter och fo-rarkommentarer vid kdrning i tdtort.* Technical report. University of Lund, Lund institute of technology, Department of traffic planning and engineering, Lund.

Petit, C., Chaput, D., Tarriare, C., LeCoz, J.Y. & Planque, S. (1990). *Research to prevent the driver from falling asleep behind the wheel.* In: Proceedings of the 34th AAAM conference, pp. 505-523.

Quimby, A.R. (1988). *In-car observation of unsafe driving actions.* Research Report ARR 153. Australian Research Board.

Reid, G.B. & Nygren, T.E. (1988). *The Subjective Workload Assessment Technique; A scaling procedure for measuring mental workload.* In: P.A. Hancock & N. Meshkati (Eds.), Human mental workload. Elsevier, Amsterdam, pp. 185-218.

Riemersma, J.B.J., Sanders, A.F., Wildervanck. C., & Gaillard, A.W. (1977). *Performance decrement during prolonged night driving.* In: R.R. Mackie (Ed.), Vigilance. Theory, Operational Performance, and Physiological Correlates. Plenum Press, New York, pp. 41-58.

Risser, R. & Brandstdtter, Ch. (1985). *Die Wiener, Fahrprobe.* In: Kleine Fachbuchreihe, Band 21. Kuratorium für Verkehrssicherheit, Wien.

Risser, R., Antoniades, A. & Lehner, U. (in prep.). *Route Guidance Evaluation: In-car observation studies. Deliverable 28, HOPES V2002.*

Rothengatter, J.A. (1991) *Automatic policing and information systems for increasing traffic law compliance.* In: Journal of Applied Behavior Analysis 24, pp. 85-87.

Rothengatter, J.A. (1993) *De risico's van rijplezier (The risks of driving pleasure).* Inaugural lecture, University of Groningen, Haren.

Rumar, K. (1990). *The basic driver error; Late detection.* In: Ergonomics 33, pp. 1281-1290.

Salusjärvi, M. (1981) *Speed limits and traffic accidents in Finland.* In: Proceedings of the OECD Symposiwn on The Effects of Speed Limits on Traffic Accidents & Transport Energy Use. Dubliri, An Foms Forbartha.

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. Van Gorcum, Assen.

Steyvers, F.J.J.M., Winsum, W. van & Brookhuis, K. (1997). *Testing safety aspects of in-vehicle information systems.* [to be published]

Steyvers, F.J.J.M., Waard, D. de, Jessunm, M., Rooijers, A.J. & Brookhuis, K.A. (1992). *Een voorstudie naar de effecten van snelheittbeperkende maatregelen op het rijgedrag op 80-km wegen.* Rapport VK 92-05. Rijksuniversiteit Groningen, Verkeerskundig Studiecentrum, Haren.

Stove, A.G. (1992). Linear FMCW Radar Techniques. *Institution of Electrical Engineers Proceedings, 139,5*

Taylor, D.H. (1964). *Drivers' galvanic skin response and the risk of accident.* In: Ergonomics 7, pp. 439-451.

Theeuwes, J. & Godthelp, J. (1995). *Self-Explaining Roads: How people categorize roads outside the built-up area.* In: Proceedings of 'International Conference Road Safety in Europe and Strategic Highway Research Program' 26-28 September, Lille, France.

Thomas, D.B. et al. (1989). *Monitoring driver status: the state of the art. DRIVE report V1004.* T1DV Rheinland, Koln.

Treat, J.R. et al. (1977). *Tri-level study of accident causes of traffic accidents.* Report No. DOT-HS-034-3-535-7. Indiana University.

Treat, J.R., Jones, R.K. & Joscelyn, K.B. (1981). *Analysis of unsafe driving actions; Data requirements and methodological issues.* Society of Automotive Engineers. Warrendale, PA: SAE.

Verwey, W.B. (1991a). *Towards guidelines for in-car information management: Driverworkload in specific driving situations.* Report IZF 1991 C-13. TNO Institute for Perception, Soesterberg.

Verwey, W.B. (1991b). *Jaarverslag ongevalsanalyse voor de KL in 1990.* Memo IZF 1991 M-35. Instituut voor Zintuigfysiologie, Soesterberg.

Verwey, W.B. (1993a). *How can we prevent overload of the driver?* In: A.M. Parkes & S. Franzen (Eds.), Driving future vehicles. Taylor & Francis, London, pp. 235-244.

Verwey, W.B. (1993b). *Driver workload as a function of road situation, age, traffic density, and route familiarity.* Report IZF 1993 C-11. TNO Institute for Perception, Soesterberg.

Verwey, W.B. (1996a). *Evaluating safety effects of in-vehicle information systems. A detailed research proposal.* Report TM-96-C045. TNO Human Factors Research Institute, Soesterberg.

Verwey, W.B. (1996b). *Evaluating safety effects of in-vehicle information systems. Testing the method.* Report TM-96-C067. TNO Human Factors Research Institute, Soesterberg.

Verwey, W.B. (1996c). *Evaluating safety effects of in-vehicle information systems. A field experiment with traffic congestion information systems (RDS-TMC) and preliminary guidelines for IVIS.* Report TM-96-C068. TNO Human Factors Research Institute, Soesterberg.

Verwey, W.B. & Veenbaas, R. (1993). *Driving performance of paraplegic drivers using common hand controls for accelerating and braking.* Report IZF 1993 C-3 1. TNO Institute for Perception, Soesterberg.

Verwey, W.B., & Veltman, J.A. (1995). *Measuring workload peaks while driving. A comparison of nine common workload assessment techniques.* Report TM B-4. TNO Human Factors Research Institute, Soesterberg.

Verwey, W.B., Brookhuis, K.A. & Janssen, W.H. (1996d). *Safety effects of in-vehicle information systems.* Report TM-96-C002. TNO Human Factors Research Institute, Soesterberg.

Waard, D. de (1995). *It's a load of work! Measurement of driver mental workload.* PhDThesis. Traffic Research Centre, University of Groningen, Haren.

Waard, D. de & Brookhuis, K.A. (1991). *Assessing driver status; A demonstration experiment on the road.* In: Accident Analysis & Prevention 23, pp. 297-307.

Waard, D. de & Brookhuis, K.A. (1995a). *Experimental test of the integrated DETER system.* In: K.A. Brookhuis (Ed.), Integrated systems. Results of experimental test, recommendations for introduction. Report V20091 DETER/Delivemble 18 (331A, 440AB). University of Groningen, Traffic Research Centre, Haren.

Waard, D. de & Brookhuis, K.A. (1995b). *Een intelligent toezichtsysteem in de auto.* In: Verkeerskunde 11, pp. 16-20.

Waard, D. de & Steyvers, F.J.J.M. (1996). *Wegbelijning in rurale verbliffgebieden: een experiment met kantbelijning.* Rapport VK 95-07. Rijksuniversiteit Groningen, Verkeerskundig Studiecentrum, Haren.

Waard, D. de, Brookhuis, K.A., Hulst, M. van der & Laan, J.D. van der (1994). *Behaviour comparator prototype test in a driving simulator.* Report 2009/DETER/Deliverable 10 (321B). University of Groningen, Traffic Research Centre, Haren.

Waard, D. de, Jessurun, M., Steyvers, F.J.J.M., Raggatt, P.T.F. & Brookhuis, K.A. (1995). *Effect of road layout and road environment on driving performance, drivers' physiology and road appreciation.* In: Ergonomics 38, pp. 1395-1407.

Wegman, F. et al. (1991). *Voor alle veiligheid (For all safety I To be safe).* SDU, The Hague.

Wertheim A.H., (1978). *Explaining highway hypnosis; Experimental evidence for the role of eye movements.* In: Accident Analysis & Prevention 10, pp. 111-129.

Wickens, C.D. (1984). *Processing resources in attention.* In: R. Parasuraman & D.R. Davies (Eds.), Varieties of Attention. Academic Press, London, pp. 63-102.

Wiener, E.L. (1987). *Applications of vigilance research; Rare, medium or well done?* In: Human Factors 29, pp. 725-736.

Wiener, E.L., Curry, R.E. & Faustina, M.L. (1984). *Vigilance and task load: in search of the inverted-U.* In: Human Factors 26, pp. 215-222.

Wierwille, W.W. (1993). *An initial model of visual sampling of in-car displays and controls.* In: A.G. Gale, I.D. Brown, C.M. Haslegrave, H.W. Kruysse, & S.P. Taylor (Eds.), Vision in Vehicles IV. North-Holland, Amsterdam, pp. 271-280.

Wilbers, P.T. (1992). *Cruise control en grenssnelheid 90 km/uur hebben de voorkeur onder chauffeurs.* In: Verkeerskunde 3, pp. 27-29.

Winsum, W. van (1997). *A validation study of a PC-based test of safety aspects of in-vehicle information systems: a test of a map-display version of a RDS-TMC task.* [to be published]

Winsum, W. van & Wolffelaar, P.C. van (1993). *GIDS Small World simulation.* In: J.A. Michon (Ed.), Generic Intelligent Driver Support System. Taylor & Francis Ltd, London.

Wolffelaar, P.C. van & Winsum, W. van (1992). *A new driving simulator including an interactive intelligent traffic environment.* In: L. Olausen (Ed.), Vehicle navigation, information systems. Moberg & Helli, Oslo, pp. 499-506.

Yaouta, K, Iizuka, H., Yanagishima, T., Kataika, Y. & Seno, T. (1985). *The development of drowsiness warning devices.* Paper presented at the Tenth International Technical Conference on Experimental Safety Vehicles, Oxford.

Yeh, Y.Y. & Wickens, C.D. (1988). *Dissociation of performance and subjective measures of workload.* In: Human Factors 30, pp. 111-120.

Zaidel, D.M. (1991). *Specification of a methodology for investigating the human factors of advanced driver information systems.* Publication TP 11199 E. Transport Canada, Ottawa, Ontario.

Zaidel, D.M., Treat, J.R., & Hume, R.D. (1980). *Pre-crash factors methodology study.* Institute for Public Safety, Indiana University.

Zeitlin, L.R. (1993). *Subsidiary task measures of driver loading: A long term field study.* In: Transportation Research Record 1403, pp. 23-27.

Zijlstra, F.R.H. (1993). *Efficiency in work behavior.* Thesis. Technical University Delft.

Zijlstra, F.R.H. & Meijman, T.F. (1989). *Het meten van mentale inspanning met behulp van een subjectieve methode.* In: T.F. Meijman (Ed.), Mentale belasting en werkstress. Een arbeidspsychologische benadering. Van Gorcum, Assen, pp. 42- 61.

Zuckerman, M. (1979). *Sensation seeking; Beyond the optimal level of arousal.* Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Zwahlen, H.T., Adams, C.C., & DeBald, D.P. (1988). *Safety aspects of CRT touch panel controls in automobiles.* In: A.G. Gale et al. (eds.), Vision in Vehicles 11. Elsevier Amsterdam, pp. 335-344.

Glossary

AHS	Automated Highway System
alpha waves	Waves in the EEG with a frequency range of 8-12 Hz, associated with lack of visual stimulation (eyes closed) or relaxation
arousal	Physiological state of increased energetical activation
ATT	Advanced Transport Telematics
AICC	Autonomous Intelligent Cruise Control
BAC	Blood Alcohol Concentration
beta waves	Waves in the EEG with high frequency (> 12 Hz) associated with high mental activity and arousal
CAROSI	CAr ROadside Signalling
CMT	Continuous Memory Task
control level	The level of the driving task that is involved in vehicle handling
DETER	Research project sponsored by the european community, aimed at Detection, Enforcement & Tutoring for Error Reduction
driver paced	The driver determines the pacing of the interaction
ECG	Electrocardiogram. A measure of cardiac activity
EEG	Electroencephalogram. A measure of brain activity
EMG	Electromyogram. A measure of muscular activity
forced-paced	The driving task characteristics are determined by other forces (such as the experimenter) than the driver. This means that the driver is not able to adjust driving behaviour in some respects
glance duration	The time the subject looks at a display
glance frequency	The number of glances towards a display
headway	Distance between the vehicle and the leadvehicle
INRETS	Institut National de Recherches sur les Transports et leur Sécurité
ISA	Intelligent Speed Adaptor
ISO	International Standardisation Organisation
ITS	Intelligent Transport Systems
IVIS	In-Vehicle Information System
IVIS-LabTest	A computer program that runs on a PC and emulates the driving task that was developed to test traffic safety of in-vehicle information systems
lateral control	Control of position on the road with respect to lane boundary
lateral position	Position on the road with respect to lane boundary
longitudinal control	Control of speed and distance to other objects
manoeuvring level	The level of the driving task that is involved in reacting to the (traffic) environment
MANOVA	Multivariate Analysis of Variance
NASA-TLX	NASA Task Load Index. A subjective measure of workload
overload	The situation that the amount of sensory input is too high to be processed adequately
RDS-TMC	Radio Data System-Traffic Message Channel
RSME	Rating Scale Mental Effort

SCR	Skin Conductance Response. A physiological measure for workload.
SD	Standard Deviation
SDLP	Standard Deviation Lateral Position
SDSTW	Standard Deviation Steering Wheel
secondary task	A task with a lower priority that is added to a main task
SWAT	Subjective Workload Assessment Technique. A subjective measure of workload
strategic level	The level of the driving task that is related to choice of transport modality, route choice and navigation
self-paced	See driver paced
system paced	The system determines the pacing of the interaction
theta waves	Waves in the EEG with a low frequency, associated with relaxation, sleep and drowsiness
time-on-task	The effect of task duration on performance
time-to-collision	The time before the vehicle collides with another object if it is assumed that vehicle speed (or relative speed) is constant
time-to-intersection	The time before the intersection is reached assuming a constant vehicle speed
time-to-line crossing	The time before any part of the vehicle crosses the lane boundary, assuming that vehicle speed does not change
TLC	Time-to-Line Crossing
TTC	Time-To-Collision
TTI	Time-To-Intersection
TNO-HFRI	TNO Human Factors Research Institute
TRC	Transport Research Centre of the Dutch Ministry of Transport and Public Works
COV	Centrum voor Omgevings- en Verkeerspsychologie (previously called : TRC)
underload	Low state of vigilance because of understimulation