A safety checklist for ATT devices

A summary of the results so far of the project 'Automation of the Driving Task'

R-97-19
Ir. T. Heijer
Leidschendam, 1997
SWOV Institute for Road Safety Research, The Netherlands

Report documentation

Number:	R-97-19
Title:	A safety checklist for ATT devices
Subtitle:	A summary of the results so far of the project 'Automation of the Driving Task'
Author(s):	T. Heijer
Research manager:	S. Oppe
Project number SWOV:	54.511
Client:	This research was funded by the Netherlands Transport Research Centre (AVV), Directorate-General of Public Works and Water Management.
Keywords:	Driver information, safety, traffic, driving (veh), driver, attention, error, behaviour, method, cost benefit analysis.
Contents of the project:	This report is aimed at investigating the effects on road safety of various applications of Advanced Transport Telematics (ATT) intended to support the driver.
Number of pages:	13 p.
Price:	Dfl. 15,-
Published by:	SWOV, Leidschendam, 1997

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

Contents

1. Introduction

This report is part of a project instigated by the Transport Research Centre (AVV) of the Dutch Ministry of Transport and Public Works. The project is aimed at investigating the effects on road safety of various applications of Telematics intended to support the driver.
Apart from AVV, which has commissioned the project, participates in some of the work and also provides overall project management, the project is carried out by three institutes:

- TNO Human Factors Research Institute;
- Traffic Research Centre of the University of Groningen (TRC);
- SWOV Institute for Road Safety Research.

During the first stage of this project the safety effects of single Advanced Transport Telematics (ATT) systems have been investigated in both a number of theoretical studies and a series of experiments. The overall aim of this research is to provide policy makers with a well-based tool to assess the safety-effects of existing and new telematic systems in road vehicles. The project must result in a set of guidelines and methods to identify potential safety hazards that single or multiple applications of these ATT systems may produce.

In the current report an attempt is made to *summarise the results of the project so far* (where possible) in the form of a checklist. Also, certain results of other recent projects will be incorporated here, e.g. the IVIS project but also the general checklist for safety effects of telematic applications developed by AVV and SWOV. In the final stage of this project this intermediate result will have to be integrated with other existing checklist procedures in order to produce a comprehensive, state of the art, checklist.

The general set-up of this project contains three or four stages:

- First, a checklist is defined that summarises available knowledge on known safety effects and diagnoses which part or function of a given ATT device may prove unsafe or doubtful.
- The second step is the definition of standard procedures for laboratory testing to produce a verdict on the ATT device or parts of the device for which the checklist was inconclusive.
- The third step determines if and what modifications of the ATT application will be necessary.

Finally, a last step should contain a cost/benefit consideration, aimed at determining whether safety benefits (if any) of the application outweigh the possible problems and costs to user and society (the problem of how to conduct such a comparison has not been addressed by this project).

However, the overall results of the project so far show that existing knowledge still only provides fragmented knowledge and not a clear, comprehensive picture. Therefore, as long as this situation remains, the second step of the scheme (laboratory testing) should be complemented by another possible testing method: full field testing.
In the following paragraphs the procedure will be elaborated.

2. A checklist

2.1. Procedures for determination of safety effects

The checklist is intended for detailed analysis of aspects of ATT systems for which sufficient knowledge exists to conduct a safety check. However, in the course of this project is has become clear that knowledge in this field is still fragmented. Therefore the checklist proposed here contains checks that refer to known criteria that are mainly based upon task load considerations. Most of the research so far has yielded criteria related to overload situations; research into underload and counterproductive behavioural adaptation, for various reasons, has provided much less readily usable criteria.

Laboratory testing
If one or more checks of the checklist result in the determination of a problem, or even when an answer can not be conclusive (not sure), two possibilities arise.
The first is, to reject the ATT application when the shortcomings are deemed too serious.
The second is a recommendation to carry out a standardized laboratory test of the specific aspect or part. Such a test can be used to resolve uncertainty which can mean definitive acceptation or rejection or recommendations for redesign of the particular function or procedure.
These standardised test procedures have not been developed yet. What is possible to determine about these test in this stage has been described in § 2.4.

Field testing
As long as we have not obtained a sufficiently complete and reliable checklist, expert judgement and field testing will be necessary to fill in the gaps both in a general sense (the development of a complete checklist) and in the occasional testing of a submitted ATT device. These field tests should be conducted conform to a standardised set-up that is referred in § 2.5.

2.2. A checklist based upon current research

It has been attempted to summarise the insights obtained with the current project in a checklist of successive questions.
The questions of the list have been formulated in such a way that a confirmative answer signals a possible problem with a certain aspect or part of the device.
Missing out on a check means that an adverse safety effect is *possible*, not that it certainly will occur. This implies that the final decision on the necessity of one or more tests 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 requires expert knowledge.

In order to provide some simple handholds to the checklist, the following characteristics of ATT systems have been used:

- The level of autonomy of the application: information only, semi-autonomous, autonomous systems.
- The modality of the user interface (if any): visual, auditory, haptic, combinations.
- The levels of the driving task: strategical, tactical, operational level.
- The main purpose of the application for general safety analysis: general effect on traffic stream, special circumstances, etc.

2.2.1. *Information systems*

Check on traffic circumstances

The first checks of an information system concern the general circumstances in which it is intended to work:

- If the ATT system under scrutiny only functions when the vehicle is at a standstill no further safety checks are necessary.
- If the system is intended for use in specific traffic circumstances only e.g. on a motorway or in a build-up area the checklist should be applied. The possible resulting tests should depend on the check whether the system will (redundantly) function outside the specified area and can not be switched off manually:
  - if so, testing should represent all possible traffic circumstances;
  - if not, testing may be limited to the specified operational circumstances.
- In all other cases the checklist should be applied and doubtful aspects or functions should be tested in sufficiently varied traffic circumstances.

Visual displays

Overload checking:

- Do any messages require more than three glances of at most one second?
- Are any messages exclusively system-paced and short-lived?
- Can messages be seen well in extreme lighting conditions (at night, in heavy sun) → is there *no* automatic adaptation to external lighting conditions?
- Does the display fail to comply with any of the legibility conditions:
  - viewing distance 70-75 cm;
  - character size 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;
  - do not only use capitals on messages longer than three words.
- Do any messages require extended decision making?
- Are any messages confusing or ambiguous; is specialised terminology used?
- Is control of the system context dependent → are there multi-level menus?

- Is sometimes immediate manual control required (e.g. deactivating an alarm)?

See further: common criteria for underload and counterproductive adaptation and physical taskload.

Auditory displays

Overload checking:

- Is the loudness of the message outside the limits:
  - 15-25 dB over background noise;
  - never over 115dB;
- Are the alarms used outside the following specifications:
  - frequency range 500-2000Hz;
  - repetition rate 1-8 1/sec;
  - non-speech messages only.
- Do any speech messages not comply to:
  - the use of familiar words only;
  - a limited set of phrases.
- Are the messages only system-paced (they can not be repeated or switched off at drivers request); do any messages require extended decision making?
- Are any messages confusing or ambiguous; is specialised terminology used?
- Is sometimes immediate manual control required (e.g. deactivating an alarm)

See further: common criteria for underload and counterproductive adaptation and physical task load.

Haptic displays

- Do any messages require extended decision making?
- Are any messages confusing or ambiguous; is the sensory message always distinguishable from random environmental inputs?

See further: common criteria for underload and counterproductive adaptation and physical task load.

Combinations of displays

In this case all relevant checklists of the involved types of display should be used as well as the common criteria for underload and counterproductive adaptation and physical task load.

Common criteria for underload and counterproductive adaptation and physical task load

Underload criteria:

- Does the system take care of obstacle detection?
- Does the system take care of signal input?
- 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, drugs)?

Counterproductive adaptation checking:

- 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?

Physical task load in operation of the device:

- 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)?

2.2.2. *Autonomous systems*

So far those elements of the checklists that pertain to the effects of autonomous functions are limited to two types: underload and counterproductive adaptation.

Underload:

- Does the system take over pedal control?
- Does the system take over part of manual control?
- Does the system affect behaviour when the driver is in a unfavourable state (fatigue, drugs)?

Counterproductive adaptation checking:

- 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?

2.2.3. *Semi-autonomous systems*

These systems can be seen as hybrids the two previous types of systems. Therefore the relevant parts of the checklists of both types should be applied first to check the general characteristics of the systems.
Apart from this judgement on the interface characteristics and the characteristics of the autonomous functions as such, separate checklist should be applied to gauge the effects of the interaction of user-interface and autonomous part; such a list is not available at this time.

2.3. **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 'unacceptable' special measures, like prioritizing the messages of the separate systems, are recommended to avoid the interference.

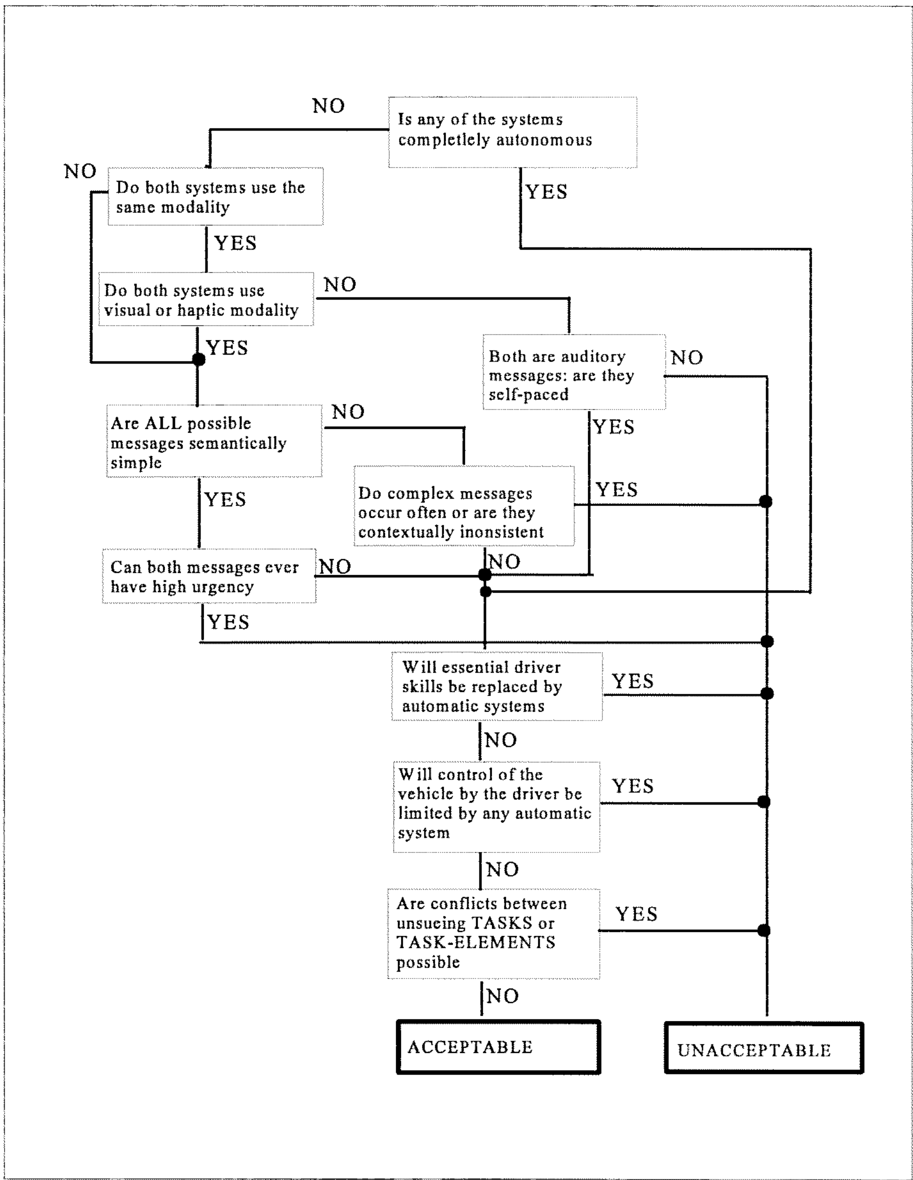

Figure 1. *Common criteria for underload and counterproductive adaptation and physical task load.*

2.4. Standard laboratory tests

If the checklist leads to the conclusion that a device contains functions or parts that do not qualify or are doubtful, a standardised testing method should provide a final verdict.
At this moment, this testing method is not yet available.
However, it is clear that these laboratory tests should eventually be able to replace the more elaborate field-tests that are described in the next chapter. the fundamental set-up of these standard field tests should also apply to the laboratory tests, specifically with respect to the composition of the group of test-subjects.

The eventual testing method should comprise a series of well defined test to represent critical aspects of the driving task and a standardised way to record and rate the performance on these tests.

2.5. Set-up and criteria for field testing

2.5.1. *Design for testing*

The design for a general field test of ATT applications (which can be a test with an instrumented test-vehicle in real traffic or a simulator test) is described in studies of Verwey et al. (1996), Verwey (1996a, 1996b) and Brookhuis & De Waard (1996). It may be summarised as follows:

- Appropriate time should be allocated to both the training of the subjects and for the tests themselves in order to ensure that 'normal driving behaviour' will be observed.
- The test requires a well defined variation of infrastructure and traffic circumstances.
- The test should be conducted in this environment with both an experimental group and a control group without the ATT application.
- The groups should be carefully matched on parameters like: gender, age, driving experience, risk-seeking/risk-avoidance.
- There should be a sufficient distribution over age groups, as ATT may affect different ages differently.

2.5.2. *Safety criteria*

Again referring to Verwey et al. (1996), Verwey (1996a, 1996b) and Brookhuis & De Waard (1996), applicable safety gauges involve both objectively measurable parameters as expert judgement.
In general, the measurements of objective parameters should at least enable to judge:

- characteristics of driving speeds;
- characteristics of momentary accelerations and decelerations;
- characteristics of headway;
- characteristics of lateral position.

Taking into account that critical values for the parameters may change with external conditions of the test, the following further specification of objectively measurable safety-related measures can be measured:

Overload indicators

- Subjective estimates (by the subject) of task load.
- Performance on secondary tasks.
- A range of physical indicators like:
    - heart rate variability;
    - steering frequency;
    - possible others like skin resistance.

Underload indicators

- Time To Collision with lead vehicle, at intersections and at traffic lights.
- Pedal use, especially the timing of pedal use.

Counterproductive adaptation
Testing for this, more long term, phenomenon requires the extension of the design in such a way that before- and after (ATT application) effects in the experimental group are compared with dual tests of the control group (tests in the same before- and after periods in the same conditions but always without ATT application). This in order to compensate for long term behavioural changes due to other effects.

In this case, the same safety-parameters are recorded and particular attention must be paid to distribution characteristics of those parameters.

Subjective safety criteria
In practice, the safety criteria defined so far have not shown to be conclusive measures and there is certainly need for extension and improvement of those measures. For that reason the use of subjective expert judgement, along with the objective measures, is a necessity.

This expert judgement can be recorded and structured according to a checklist that is derived from the one defined and used by De Gier (1980); this checklist has been developed to gauge the effects of drugs on driving behaviour. It is set up in such a way that behavioural changes due to drugs can be measured on all sorts of actions including preparation to driving, and driving itself.
This list contains a comprehensive series of checks to judge driving behaviour in all possible driving situations and this part is also applicable to this project.

2.6. Determining acceptability: cost/benefit analysis

Although not explicitly included in the set-up of this research project, a checklist for overall safety effects is not complete if it can not lead to a balanced insight in the possible gains and losses that may be related to an ATT application. So far, the checklist has concentrated upon the determination possible detrimental effects for road safety due to changes in individual behaviour, but it is of course very well possible that the general effect of a device on road safety must be judged positive. By Oppe & Bos (1996) a general method is proposed to assess these general safety effect on several levels: effects on mobility, effects on traffic streams, effects on general traffic behaviour, effects on the individual driving task (the subject of the current research) and effects on accidents in general.
This method in itself does not constitute a checklist but a way to combine several types of more long term observation techniques. Still, this method may also lead to the 'condensed' form of an off-line checklist for overall safety effects, necessary to obtain a final cost/benefit judgement of ATT devices in terms of general increase or decrease of specific types of accidents.
Both the completion of the checklist and the definition and validation of standardised laboratory tests require standardised field testing.

3. Discussion and conclusion

3.1. Results so far

The checklist constructed on the basis of results so far is predominantly based upon criteria related to *overload*. Criteria to gauge underload are significantly fewer and especially criteria to rate the effects of counterproductive behavioural adaptation are very incomplete.
To produce a more balanced result, the last phase of the current project should be aimed at obtaining more firm criteria for underload and possibly a separate checklist for counterproductive behavioural adaptation.
The problem with the latter aspect is, that the theoretical basis for criteria is not very well developed: the risk homeostasis theory is an obvious candidate for a basis, but this theory is not uncontested. Handholds may also be found in the concepts of 'situation awareness'; a theoretical framework that has so far mostly been applied to complex tasks of air traffic controllers or controllers of nuclear plants. Also theories concerning the human propensity to optimise task load against performance may be applicable, in which case we can e.g. analyse the nature of criteria functions that have been used in successful optimal control models for human control tasks.

As long as the safety checklist is incomplete and the standardised laboratory tests are not completely developed, field testing, either with instrumented vehicles in real traffic or with driving simulators, will be necessary. In order to provide a coordinated basis for the development of checklist and laboratory tests, the field tests should be conducted according to the standards that have been developed.

3.2. Project continuation

While keeping in mind that the development of a complete safety checklist *and* a comprehensive method for laboratory testing lies beyond the scope of the current project, the current project should at least provide a viable starting point.
To that end the last phase of this project should address the following items:
- completion of the provisional standard for field testing;
- enhancing the checklist criteria for underload;
- producing a set-up for a checklist for counterproductive behavioural adaptation;
- making a start with the definition of standardised laboratory tests.

Literature

Brookhuis, K. & Waard, D. de (1996). *Limiting speed through telematics.* Traffic Research Center, University of Groningen, Haaren.

Gier, J.J. de (1980). *Evaluatie van geneesmiddelen in het verkeer.* Proefschrift. Utrecht.

Kuiken, M.J. & Heijer, T. (1995). *Driver support systems and traffic safety.* R-95-68. Institute for Road Safety Research SWOV, Leidschendam.

Oppe, S. & Bos, M.J.M. (1996). *Systematiek voor het toetsen van telematica-systemen op verkeersveiligheidseffecten.* R-96-16. Institute for Road Safety Research SWOV, Leidschendam.

Verwey, W.B. (1996a). *Evaluating safety effects of in-vehicle information systems (IVIS).* TNO Human factors research institute, Delft.

Verwey, W.B. (1996b). *Evaluating safety effects of in-vehicle information systems; Testing the method.* TNO Human factors research institute, Delft.

Verwey, W.B., Brookhuis, K.A. & Jansen, Z.W.H. (1996). *Safety effects of in-vehicle information systems.* TNO Human factors research institute, Delft.