6th Framework Programme - Priority 2 “Information Society Technologies”
“HUMAN centred design for Information Society Technologies”
Proposal n°507 420
Contract n°507420

TRL (Transport Research Laboratory)
Transport Systems Unit, Crowthorne House, Nine Mile Ride, Wokingham, Berks, RG40 3GA, ENGLAND
e-mail: scotter@trl.co.uk
http://www.trl.co.uk

VTT Building and Transport
PO Box 1902
FIN 02044, VTT, FINLAND
http://www.vtt.fi

BAST (Federal Highway Research Institute)
Department of Automotive Engineering – Driver Assistance Systems, Brüderstraße 53, D-51427 Bergisch Gladbach, GERMANY
e-mail: gelau@bast.de
http://www.bast.de

Issues Related to Integrated Methodologies
Report of the Task Force E Santorini Workshop
June 13, 2005

Deliverable 3 of Task Force E

Reference: ETRL-050725-E1-C

CONFIDENTIALITY LEVEL: PUBLIC

Rev.	Issuing date	Pages	Written by	Visa	Verified by	Visa	Approved by	Visa
A	25th July 2005	198	Virpi Anttila (VTT), Christhard Gelau (BASt)	✓	Alan Stevens (TRL)	✓		
Modifications : Creation of Document								
B	31st July 2005	198	Sally Cotter (TRL)	✓	Christhard Gelau (BASt)	✓		
Modifications : Amendments following external review								
C	25th October 2006	201	Sally Cotter (TRL)	✓	C.Brusque (INRETS), L.Konstantinopoulou (ICCS)	✓		
Modifications : Amendments following internal review of external review comments								

Table of Contents

1. Introduction

Main objectives

One of the main objectives of the HUMANIST Task Force E is to exchange through the network, the knowledge and experience of projects which have applied or developed methodologies for the evaluation of IVIS both in terms of safety and usability. The task force will consider the integration of various methods into comprehensive methodologies. There is particularly close co-operation with Task Force D where studies are considered where these methods are applied.

The previous deliverable (Deliverable D.2/E.2) presented a matrix to summarise the main aspects of applicability of various available methods for assessing specific aspects of safety and usability. The deliverable also reported on the exchange of knowledge within the network from both Task Force D and E as a result of a workshop in Toulouse, September 2004. Having identified methods with different advantages and disadvantages, the next step of Task Force E was to consider how these methods could be integrated into appropriate methodologies.

In order to get a common understanding of what is meant by "Integrated Methodologies" the following preliminary definition was proposed as a result of the discussion during the Task Force E meeting on March 17 in Lyon:

"Structured human factors evaluation (of a driver interface) that combines evidence from multiple assessments of different aspects of driver-vehicle interaction within a conceptual framework."

The integration of methodologies will for example need to address the issue of how quantitative and qualitative output from the various methods within a methodology could be combined and applied into a meaningful result. In order to proceed towards integration of different methods, a workshop discussing the various issues related to integrated methodologies was held in Santorini, June 2005.

The main objective of this deliverable is to report on this first exchange of knowledge concerning the integration of individual methods into a more holistic approach to assessing safety and usability. The deliverable includes short summary of each presentation, minutes of the general discussion during the afternoon session and a brief list of actions. In the annexes the agenda (Annex 1), the list of Task Force E members (Annex 2) and the presentations (Annex 3a-3g) are included.

The Santorini workshop

The workshop was held in Santorini, Greece on 13th June 2005. The morning session was comprised of five presentations concerning different aspects and methods of safety and usability. The afternoon session started with the sixth and final presentation, which formed a transition to a more in-depth discussion about integrated methods and proceeding in Task Force E. The morning session was also open to ESRs (early stage researchers). The objective of this workshop was to provide introductions to the subjects in an interesting way as well as exploring more in-depth material.

2. Summary of Morning Presentations

ISO Standards and the Suitability process

Presentation made by *Alan Stevens, Transport Research Laboratory, UK*.

This presentation was in three parts:
The **first** was a short introduction to the first TF E deliverable that says a method is the combination of "Metric", "Technique", "Tool" and "Environment" and has other important properties. These key concepts where defined and explained in detail.

The **second** part was to introduce ESRs to the fascinating world of standardisation by addressing issues including:

- What standards are (and what they are not)
- The different types of standards
- Why standards are interesting and necessary
- How standards are developed

The **third** part was a summary of the "Suitability Standard" (ISO 17287) with a description of seven different stages of a evaluation process which could also become a part of an integrated methodology. The stages described were:

1. **Planning of resources and timing**
2. Detailed description of the system being studied
3. Context and any restrictions for system use
4. Setting success criteria based on specific metrics (variables)
5. Selecting other details of the method (techniques, tools, environments)
6. Data collection
7. **Interpretation of results**

Stages 1 and 7 are performed once only, while stages 2-6 can occur many times as part of the assessment.

HASTE Risk Assessment Procedure

Presentation made by *Wiel Jansen, TNO Human Factors, The Netherlands*.

The presentation covered the following steps for IVIS (not ADAS) evaluation developed by the recently finished European project HASTE ("*H*uman Machine Interface *A*nd the *S*afety of *T*raffic in *E*urope"):
(1) Estimating the effects an IVIS has on a selected (and validated) set of driving behaviour parameters, in a standardised environment (i.e., rural road, mid-range simulator).
(2) Doing this also for workload and distraction effects that are caused by IVIS.

The presentation began with a description of the criteria needed for good methods: Efficiency, effectiveness, relevance to safety, reliability over sites, comprehensive and safe for the participants.

The HASTE project validated selection of behavioural parameters (that differentiates between IVIS task difficulty levels) comprised of just four metrics and did not include any lateral positioning metrics:

- Subjective rating of own performance
- Average speed (compensating for increased workload)
- Amount of high-frequency steering
- Minimum headway

The useful distraction metrics included PDT and %road centre visual attention.

The HASTE proposal for experimental procedures is:

- At least a medium range simulator and standardised rural road
- 10-15 subjects age between 25 and 50 M & F with driving experience 10k annually for at least 5 years
- Task duration about 10 minutes
- A single baseline of 10 minutes

Combining the behavioural and distraction sets of effects into one overall estimate of accident/risk is a major challenge, where one possible approache might be micro-simulation. Also penetration rates of IVIS, number of times a display is used etc. needs to be taken into account when considering the overall risk. The research continues in AIDE project (ongoing research).

How to select and combine evidence from different sources: Problems to be solved

Presentation made by *Josef F. Krems, Chemnitz University of Technology (CUT), Germany*.

The presentation began with the positive aspects of IVIS and it was noted that a key point is to make dialogues “chunkable”, i.e. divisible into unit which enable an easy resumption after an interruption. Also fast and precise apprehension and learnability are key issues considering successful IVIS.

At the measurement level there are behavioural indicators and performance indicators and key criteria for choices between indicators were described. From the TF E deliverable matrix, five subsets were identified:

- driving performance,
- physiological measures,
- laboratory tasks,
- subjective ratings and
- incident analysis.

These can also be organised as objective, physiological and subjective measures.

However, the practical difficulties including validity and ethical considerations are difficult for behavioural measures and physiological measures are highly unspecific. Subjective measures appear valid but are not sufficient.

An alternative approach of laboratory measurements was then discussed based on the German industry project ADAM (“*A*dvanced *D*river *A*ttention *Me*trics”). These include

occlusion, PDT and the "Lane Change Test". The conclusion was that a combination of metrics could be proposed based on a test battery with a simple linear multiple regression:

Alpha = b1 X1 +b2X2 + ... + bnXn

where Alpha = distraction, workload, usability, safety … and Xn are subjective rating, occlusion …

Finally, the comment was made that we do not have a sufficient model of driver behaviour to set these parameters at the moment.

Socio-Economic Impact assessment

Presentation made by *Torsten Geißler, University of Cologne, Germany*.

In this presentation, it was argued that socio-economic impact assessment is strongly required for the evaluation of Intelligent Vehicle Safety Systems (IVSS) since the efficiency proof on individual and societal level represents an important force for acceptance and implementation of these applications. It was proposed that the balance between costs and benefits of those IVSS that can be expected to have a significant impact in the next 10-15 years will be the basis for the development of a roll-out strategy for IVSS.

The presentation demonstrated, based on the findings of the recently finished European project SeiSS ("Socio-Economic Impact Assessment of Intelligent Vehicle Safety Systems"), a methodological approach of how this socio-economic impact assessment can be carried out consistently. This involves Technology, Market, and Traffic considerations followed by impact assessment by modelling and monetarisation.

The process identifies safety-critical effects, possible accidents and their costs. It also identifies non safety-critical effects including travel time savings, vehicle costs and environmental effects. Subjective methods of monetarisation include willingness-to-pay. More objective methods use the cost-of-dame approach, the cost-of-avoidance approach or the market-data-divergence analysis (e.g. comparing real-estate prices caused by different noise levels). Finally, using standard methodology a benefit-cost ratio is calculated and if greater than one, the system provides societal benefits.

The workability of the proposed approach was then described based on exemplary case studies of eCall, Safe following (ACC) and Lane Change Assist/Lane Departure Warning. The main challenges are political and administrative issues of definitions as well as the need for better accident prediction models. A complimentary approach would involve multi-criteria as well as individual Stakeholder perspectives. More information about the SeiSS project and the report can be found at www.escope.info.

The EuroNCAP Rating Procedure

Presentation made by *Christhard Gelau, BASt, Germany*.

The presentation began with an overview of Euro NCAP (European New Car Assessment Programme) which is a consumer organisation formed in 1998. Whilst type approval provides minimum performance requirements to be fulfilled by every vehicle, NCAP

provides additional information to inform consumers specifically about safety aspects. This has recently been extended from passive safety (frontal, side and pedestrian crash performance using dummies) to active (or pre-crash) safety. The result is summarised as a “star rating” with 1-5 stars.

The scoring procedure was then described in detail. Points are calculated on a sliding scale for each separate parameter (e.g. for depression on the chest of a dummy). This is performed for all dummy body parts (head, chest, knee, leg etc.). There are also “modifiers” that can reduce the scoring based on vehicle displacements or contacts between the dummy and the vehicle. The worst points in each section are summed and the total points converted into the “star rating” for the respective vehicle.

Observe and Communicate

Presentation made by *Karin Ausserer, Factum, Austria*.

The presentation introduced system safety, HMI safety and traffic safety and described how traffic is both a social and a technical system. Traditionally, traffic safety is evaluated by means of accident criteria. Accident prognoses in connection with new equipment, however, are speculative and not very reliable. In addition they do not give any answers in which way a system has to be improved or further developed or if it should be implemented, at all.

Hypotheses concerning acceptance aspects and possible behaviour outcomes should be based on verbal data derived from target groups, if one does not want to rely on one’s own reasoning, solely. Verbal data have to be analysed and structured, among other things in the light of scientific tradition and results of earlier research.

Behaviour and interaction analyses have to be included in prognoses, as traffic is not only a technical, but a socio-technical system. Therefore changes in behaviour and interaction can be expected. Behavioural adaptation (risk compensation, delegation of responsibility, imitation, changes in communication patterns) is one phenomenon that has to be dealt with when evaluating new equipment.

It was stated that it is very important to assess possible consequences of adaptation processes (positive and negative) for the whole traffic system. This means that vulnerable road users also have to be involved in the evaluation process. Prognoses with respect to these questions have to be dealt with heuristically (expert discussion, workshops, etc.) as no empirical approaches are possible, yet.

It was suggested that social-scientific methods have to be made use of, and different method of analysing traffic safety were then described, including:

- Literature analysis
- Traffic safety checklist (developed within the Prometheus project)
- “Wiener Fahrprobe” – in-car observation method with two observers
- Focus group interviews with 5-8 people, either heterogeneous or homogeneous
- In-depth interviews with a single person taking 1-4 hours
- Standardised interviews/questionnaires
- Workshops as a heuristic approach to elaborate specific topics

3. Afternoon Presentation and Discussions

Evaluation, Environment and Questions

Presentation made by *Eva Gelova, CDV, Czech Republic*.

This presentation began with an overview of the environment in which evaluation is undertaken (including standards, regulations, and different actors including manufacturers, drivers – including drivers with special needs – and the public). It was suggested that the environment should be described completely by defining the roles, responsibilities and relations amongst the Stakeholders as well as the existing situation with standards, regulations etc. A possible protocol for the evaluation process was presented which included validation and certification.

The presentation finished with a range of challenging questions, including:

- Which measures (e.g. subjective and/or objective; how to ensure statistical minimum)?
- How to choose the methods (favourite, validity, etc.)?
- Should we prepare a specific set of measures and a protocol within Humanist or just general recommendations?
- Who should pay for any assessment?
- Should there be external authorised independent testing and certifying institutes?
- From which point of view should the results be presented (e.g. consumer rating, teachers rating)?
- Can we develop a suitable/unsuitable conclusion?
- How can we ensure acceptance from the Stakeholders (voluntary or enforced?)

Questions in order to proceed in Task Force E

Some question prepared by Christhard Gelau were presented in order to start discussion about how we should proceed in order to achieve structured human factors evaluation that combines evidence from multiple assessment of different aspects of driver vehicle interaction within conceptual framework.

Questions:

- Who are our "customers" in Humanist?
- Should operation be considered on individual or society level?
- Is there an agreement on the concept of risk?
- Do we have conceptual framework (do we know what it is to be measured)?
- According to what rule should we combine the evidence?

Summary of Afternoon Discussions

Alan Stevens proposed three different levels of analysis with different customers:

- Strategic issues discussed by Eva Gelova– Customer is “Society”
- Socio-economic level discussed by Torsten Geißler – Customer is “EC”

- Scientific level – Customer is “Humanist”

Torsten Geißler showed how the E-MERGE project has identified benefits for different stakeholders. This underlines the point that SeiSS is operating at a more strategic level.

At the scientific/Humanist level, four frameworks were discussed:
- Suitability – as an overall framework
- Linear regression model – to combine evidence
- HASTE method – has identified the important methods
- Euro NCAP expert experience approach

Q: Do we know what is to be measured and do we know the “rules” for combining evidence?
A: Linear model is most general approach and compatible with Suitability and HASTE and (possibly) Euro NCAP. However, we don’t have an overall answer at this point. To establish a valid linear regression model can be viewed as a long-term goal of the Task Force.

Q: Do we have an agreed conceptual framework of driver information processing and of risk that is sufficient to really make progress?
A: We need a conceptual framework, not a model. Several frameworks were discussed at the ISPRA workshop and these may be sufficiently detailed; they may not have to be too sophisticated. The framework made by AIDE appeared to be the most favoured.

4. Conclusions and Areas for Further Work

The main conclusions for the workshop were:

1. As a group, the Humanist partners understand their "customers" and their role: that is, scientific work on safety and usability of IVIS and ADAS. This work provides an output from the project into wider socio-economic and policy considerations of technology development

2. We have a general common understanding of safety and risk, but there is a need to identify and use a common conceptual framework of driver information processing and risk (it is not necessary to have a detailed model). Such conceptual models exist.

3. There remains a research gap concerning how to combine individual methods into an overall integrated methodology. The next deliverable (Deliverable E.4) aims to address this issue and will describe the status of research in developing integrated methods

Next actions for Humanist for development of an integrated methodology:

1. Identify a conceptual framework of driver information processing and risk

2. Examine the common set of parameters to measure driver behaviour being developed by Humanist TF2, and adopt or develop these as necessary.

3. Identify a list of specific detailed research questions for future integrated methodology development

List of Annexes

Annex	Contents	
1	Workshop Agenda	
2	List of TF E members	
3	**Partner**	**Title**
3a	Alan Stevens (TRL)	ISO Standards and the SUITABILITY process for assessment of Safety and Usability
3b	Wiel Jansen (TNO)	HASTE Risk Assessment Procedure
3c	Josef Krems (CUT)	How to select and combine evidences from different sources
3d	Torsten Geißler (U. Cologne)	Socio-Economic Impact Assessment of Intelligent Vehicle Safety Systems - Methodological Approach and Case Studies
3e	Christhard Gelau (BASt)	The EURONCAP/PNCAP rating procedure
3f	Karin Ausserer (Factum)	"Observe and Communicate"
3g	Eva Gelova (CDV)	Evaluation, Environment and Questions to answer

ANNEX 1: WORKSHOP AGENDA

09:00 - 09:30	Alan Stevens (TRL) ISO Standards and the SUITABILITY process for assessment of Safety and Usability
09:30 – 10:00	Wiel Janssen HASTE Risk Assessment Procedure
10:00 – 10:30	Josef Krems (CUT) How to select and combine evidences from different sources
10:30 – 11:00	Coffee break
11:00 – 11:30	Torsten Geißler (University of Cologne). Socio-Economic Impact Assessment of Intelligent Vehicle Safety Systems - Methodological Approach and Case Studies
11:30 – 12:00	Christhard Gelau (BASt) The EURONCAP/PNCAP rating procedure
12:00 – 12:30	Karin Ausserer (Factum) "Observe and Communicate"
12:30 – 14:00	Lunch break
14:00 – 14:20	Eva Gelova (CDV) Evaluation, Environment and Questions to answer
14:20 – 16:00	Discussion and conclusions from lectures, discussion of ‘key words’
16:00 – 16:30	Coffee break
16:30 – 17:00	How to proceed in TF E
17:00 – 17:20	Discussion of Indicators (Reference page 7 of Terms of Reference document 7ERT-050308-T1 DA)
17:20 – 17:45	Ideas for 2006 ITS World Congress in London Future research needs for FP 7

ANNEX 2: LIST OF HUMANIST TF E MEMBERS

Christhard Gelau	BASt	Gelau@bast.de
Yvonne Barnard	Eurisco	Yvonne.Barnard@onecert.fr
Anke Mogilka	CUT	Anke.mogilka@phil.tu-chemnitz.de
Annie Pauzine	INRETS	Annie.pauzie@inrets.fr
Christine Turetschek	Factum	Christine.turetschek@factum.at
Eva Gelova	CDV	Eva.Gelova@cdv.cz
Ioanna Spyropoulou	NTUA	iospyrop@central.ntua.gr
Iva Hanzlikova	CDV	iva.hanzlikova@cdv.cz
José Manuel Menéndez	UPM	jmm@etsit.upm.es
Karel Schmeidler	CDV	Schmeidler@cdv.cz
Lena Nilsson	VTI	Lena.nilsson@vti.se
Martin Baumann	CUT	Martin.baumann@phil.tu-chemnitz.de
Matthew Kalaftis	NTUA	mgk@central.ntua.gr
Merja Penttinen	VTT	Merja.Penttinen@vtt.fi
Virpi Antilla	VTT	virpi.anttila@vtt.fi
Nina Dragutinovic	SWOV	Nina.dragutinovic@SWOV.nl
Stevens, Alan	TRL	astevens@trl.co.uk
Thomas Troglauer	DtF	tt@dtf.dk
Truls Vaa	TOI	TVa@toi.no
Wiel Janssen	TNO	janssen@tm.tno.nl
Sally Cotter	TRL	scotter@trl.co.uk

ANNEX 3A-G : PRESENTATIONS