HASTE Workshop
Brussels
22 March 2005

What should criteria for in-vehicle HMI be like?

Oliver Carsten

Institute for Transport Studies

University of Leeds

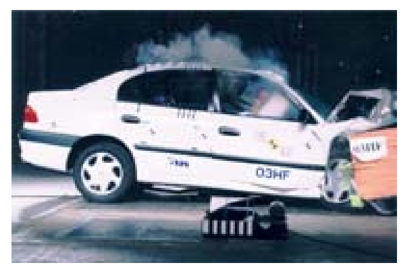
Types of Standard

- Product

- Procedural

ISO 9001 Certified

- Performance



Product standards

- Specify physical aspects of system
- Example: home rows of keyboard should be no more than 50mm above desk level
- Advantages:
  - Clear
  - Easy to comprehend
- Disadvantages:
  - Tend to apply only to parts of a system
  - Technology dependent
  - Tend to become outdated

Procedural standards

- Describe process of analysis and testing which a manufacturer must use
- Example: government requirements of suppliers in military purchasing
- Advantage:
  - Can ensure participation of relevant specialists in the design process
- Disadvantages:
  - Require strong certification body (audit)
  - Need large amount of documentation
  - Can slow deployment of new systems

Performance standards

- Specify performance levels of or with a system
- Example: drivers should be able to obtain the information from a display without taking their eyes off the road for more than a certain amount of time
- Advantages:
  - Technology independent
  - Can apply to whole systems
  - Encourage innovative design
- Disadvantage:
  - Require thorough testing, e.g. by test organisations

The HASTE approach

- Performance testing is the best option
- Focus should be on effect of IVIS on the driving task
- 2 major studies:

1. Does greater secondary task load from an In-Vehicle Information System (IVIS) lead to an identifiably worse performance in the primary task of driving?
2. How can the methods and indicators developed in (1) be applied to assessing tasks on real systems?

Criteria for a test regime, 1/6

- ***Efficiency***
  - Any unnecessary elaboration or duplication should be removed.

Criteria for a test regime, 2/6

- ***Effectiveness***
  - The sample size (number of tests) needs to be sufficient to reveal differences between good and poor designs.

Criteria for a test regime, 3/6

- ***Reliability***
  - The tests, when repeated at different test sites or with different drivers, should produce similar results.
  - This could argue in favour of using a driving simulator or laboratory environment, because in such an environment it is easier to control the conditions and situations encountered.

Criteria for a test regime, 4/6

- ***Relevance***
  - The criteria being used to assess the IVIS should be related to the safety of the driving task.
  - Poor functionality or usability of a system in aspects that cannot be used while driving, e.g. use of a menu that is locked while the engine is running, might affect the user’s impression of a system, but is not safety-relevant.

Criteria for a test regime, 5/6

- ***Comprehensiveness***
  - All important safety implications should be assessed.
  - This argues in favour of using driving in a naturalistic environment, i.e. on real roads, as part of the test regime, since such driving is more likely to reveal unanticipated problems which might not be revealed in the more constrained environment of a simulator.

Criteria for a test regime, 6/6

- ***Safety***
  - Neither the test subject (driver) nor the test administrator (e.g. an observer in a test vehicle) should be exposed to improper risk.

Main parts of project timetable

Jan to June 2002	WP1 Establish experimental protocol
July 2002 to March 2004	WP2 Examine distraction and driving performance with surrogate IVIS
March 2004 to March 2005	WP3 Refine test procedure and apply to “real” IVIS systems

HASTE Workshop
Brussels
22 March 2005

WP2: HMI and Safety-Related Driver Performance

Joakim Östlund, VTI

The objectives of WP2

- Investigate impact of IVIS on driving performance
- Identify **Indicators**
  - Sensitive
  - Reliable
  - Valid
- Develop **Test scenarios**
  - Realistic scenarios
  - Safe enough for attention to IVIS
  - Difficult enough to find effects on driving performance
- Identify advantages
  - simulators
  - laboratories
  - field experiments

Approach

- Surrogate IVIS (S-IVIS)
  - One cognitive, one visual
  - Three S-IVIS levels
- Assessment methods
  - Simulator, Laboratory and Field
- Road
  - Urban, rural and motorway
  - Road complexity level
- “Average” vs elderly drivers
- UK drivers vs Portuguese drivers

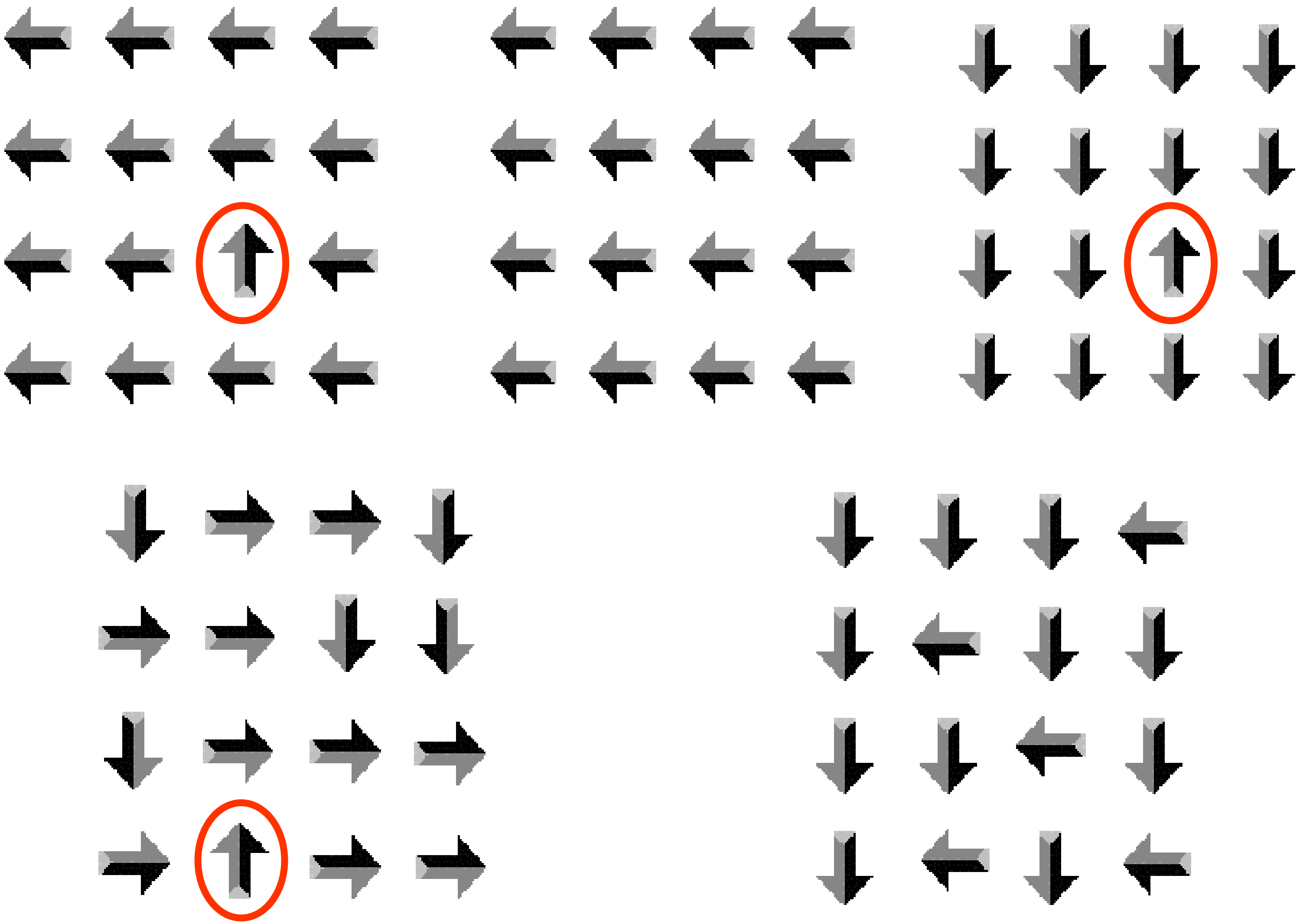
Scope

17 experiments, 527 participants

A standard rural road

- Was included in all simulator and laboratory experiments
- Identical scenario
- Average drivers

The Visual S-IVIS



The Visual S-IVIS in a car

The Cognitive S-IVIS

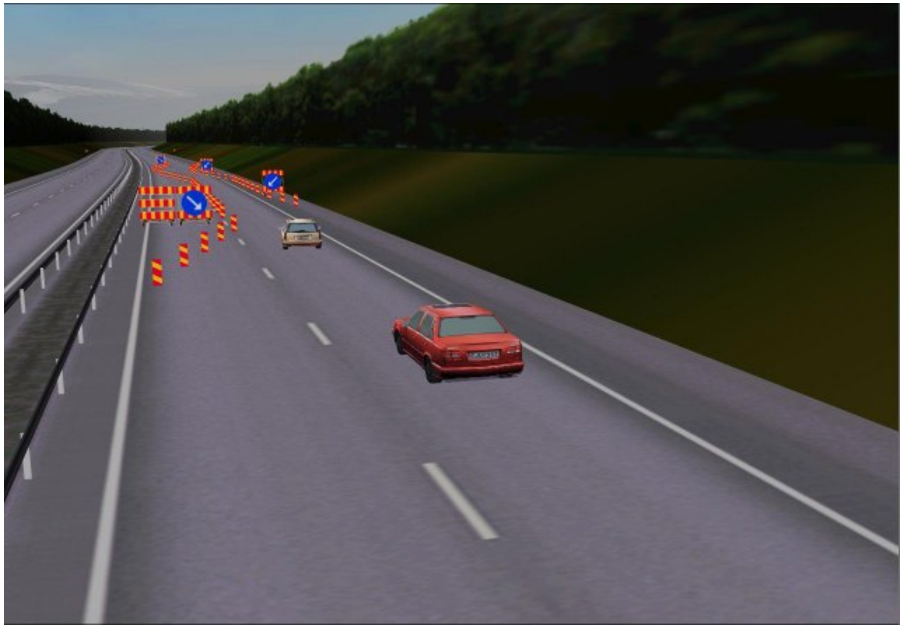
- Count target sounds – separate tally for each one
- 2-4 target sounds
- Max 2 seconds per sound
- 15 sounds – 45 seconds

Road Complexity Levels

- Factors determining road complexity
  - Road curvature
  - Intersections (urban only)
  - Interferring vehicles (sim/lab only)


R587 RM

Simulator motorway



Simulator urban road

Laboratory, rural road

Selected vehicle measures

- Speed measures
- Headway measures
- Steering control measures
- Lateral control measures
- Physiological measures
- Gaze angle measures
- Self report
- Observer ratings
- S-IVIS performance

Analysis

- Two factorial ANOVA
  - Road complexity
  - S-IVIS difficulty
- Effect of nationality (UK Vs Portuguese)
- Effect of age
- Analysis of each experiment
- Meta analysis including all experiments

Results – Effects of S-IVIS

- Most pronounced effects for the Visual task
- The two task types sometimes had different effects

The Visual Task

- Attention to visual task - Fewer glances straight ahead
- Increased lateral position variation
- (Compensatory) speed reduction …

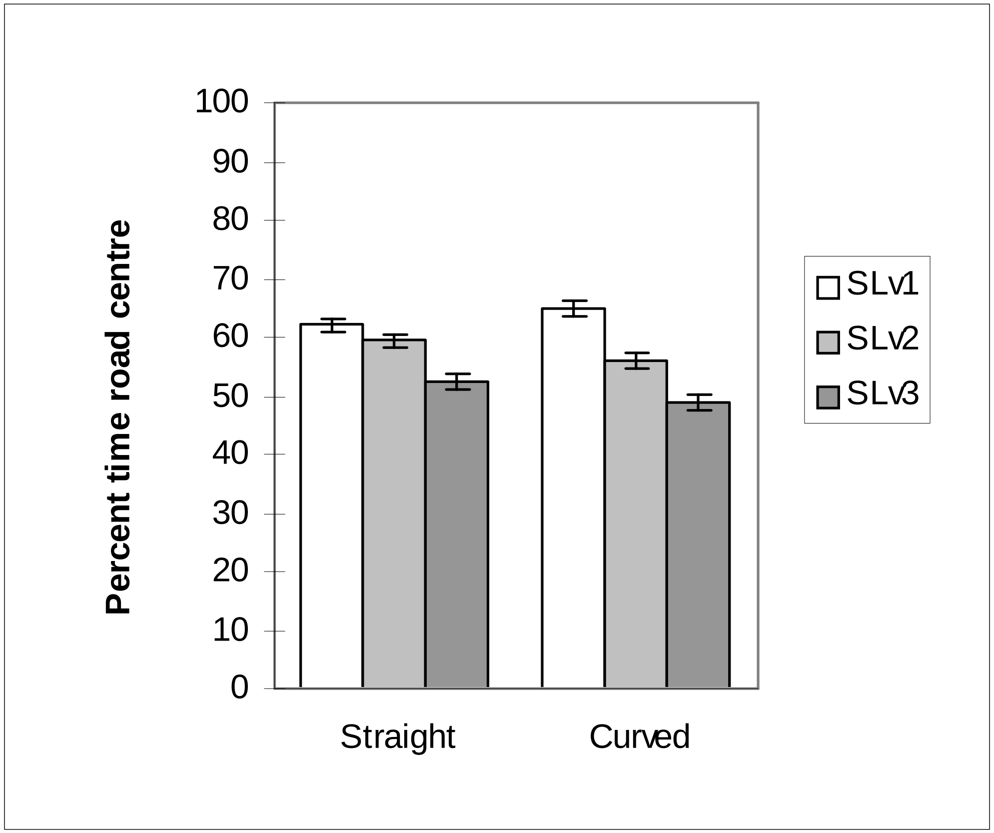
The Cognitive task

- More glances straight ahead
- Decreased lateral position variation
- Indications of decreased headway

Both tasks

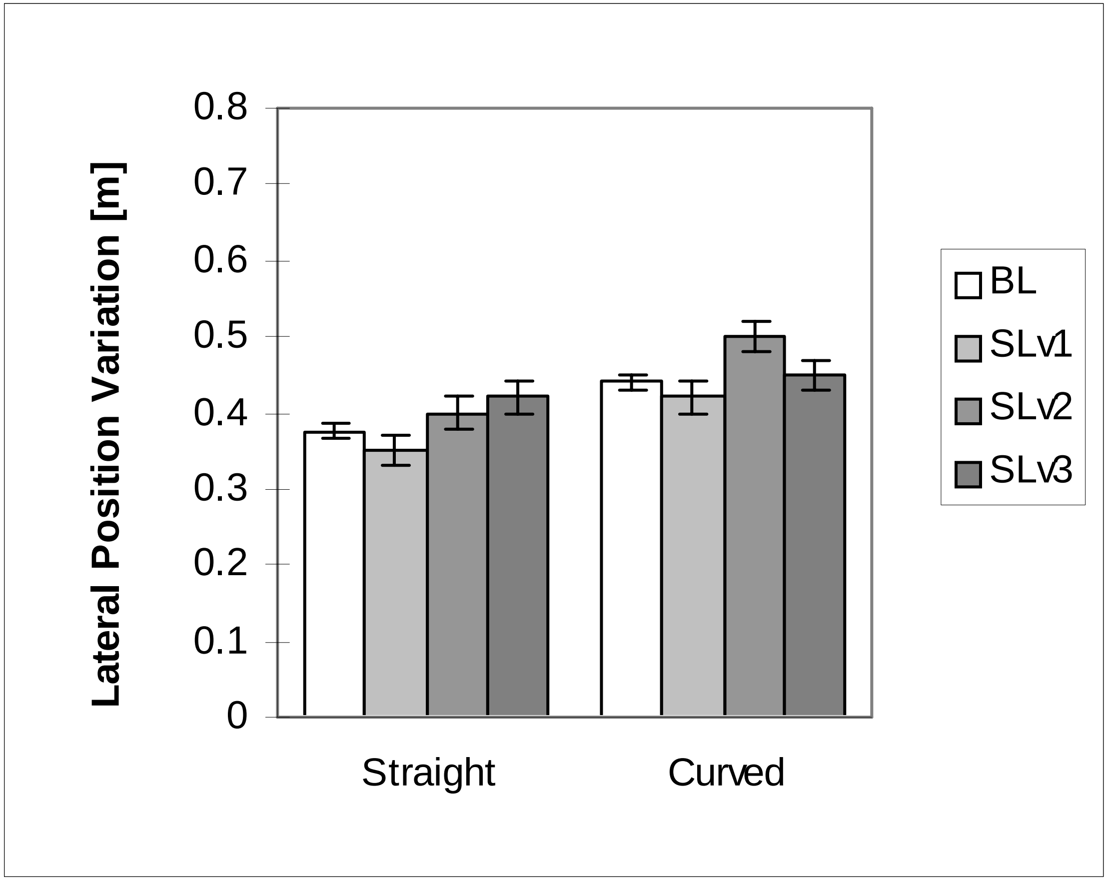
- Decreased self rated driving performance
- Increased steering activity
- Occasional speeding

The drivers looked less on the road centre while attending to the **visual** S-IVIS



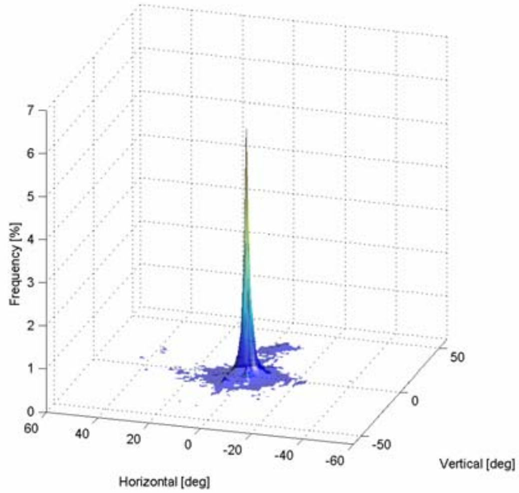
Time eyes on road centre

The lateral control was less stable when attending to the **visual** S-IVIS

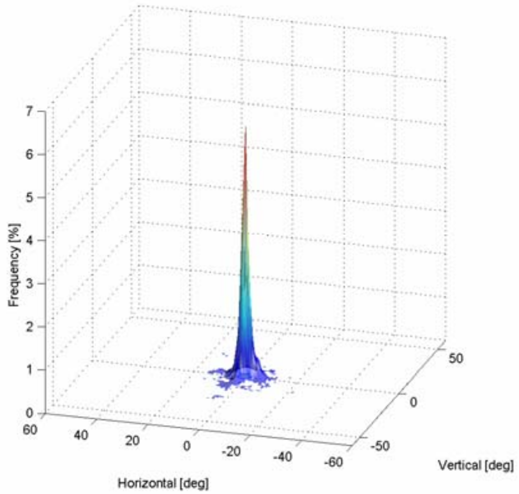


Lateral position variation

The drivers looked more on the road centre while attending to the **cognitive** S-IVIS



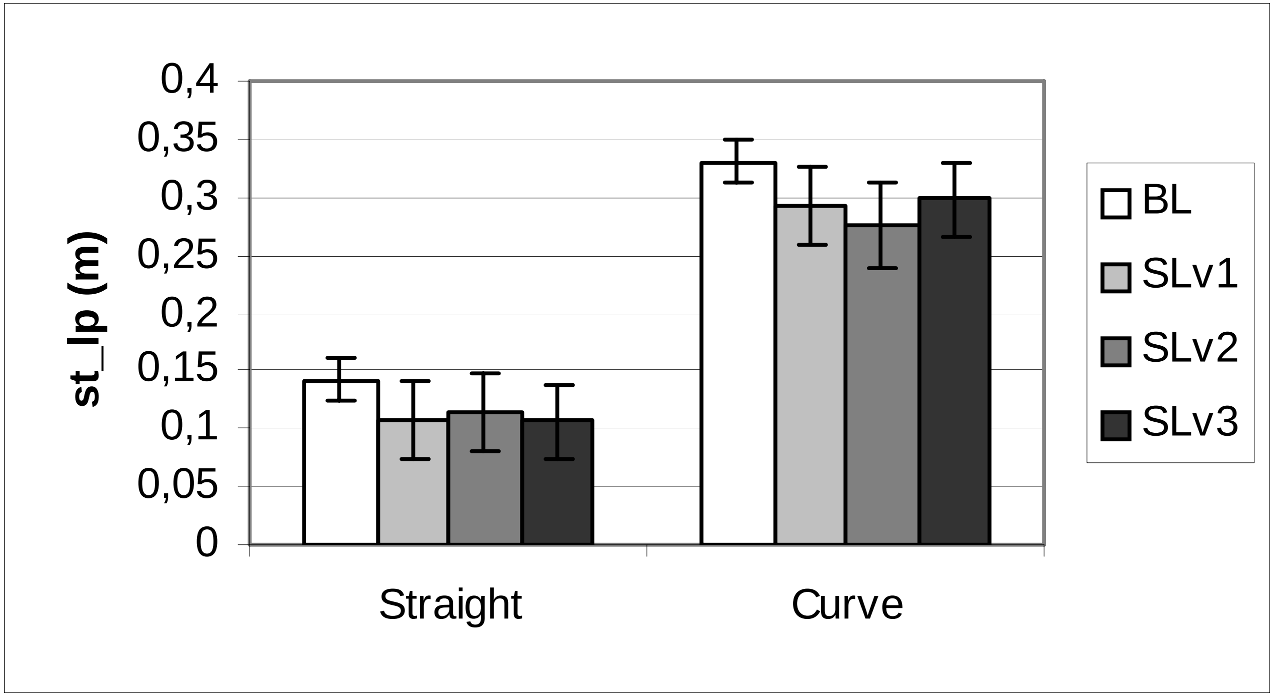
Baseline



Cognitive task

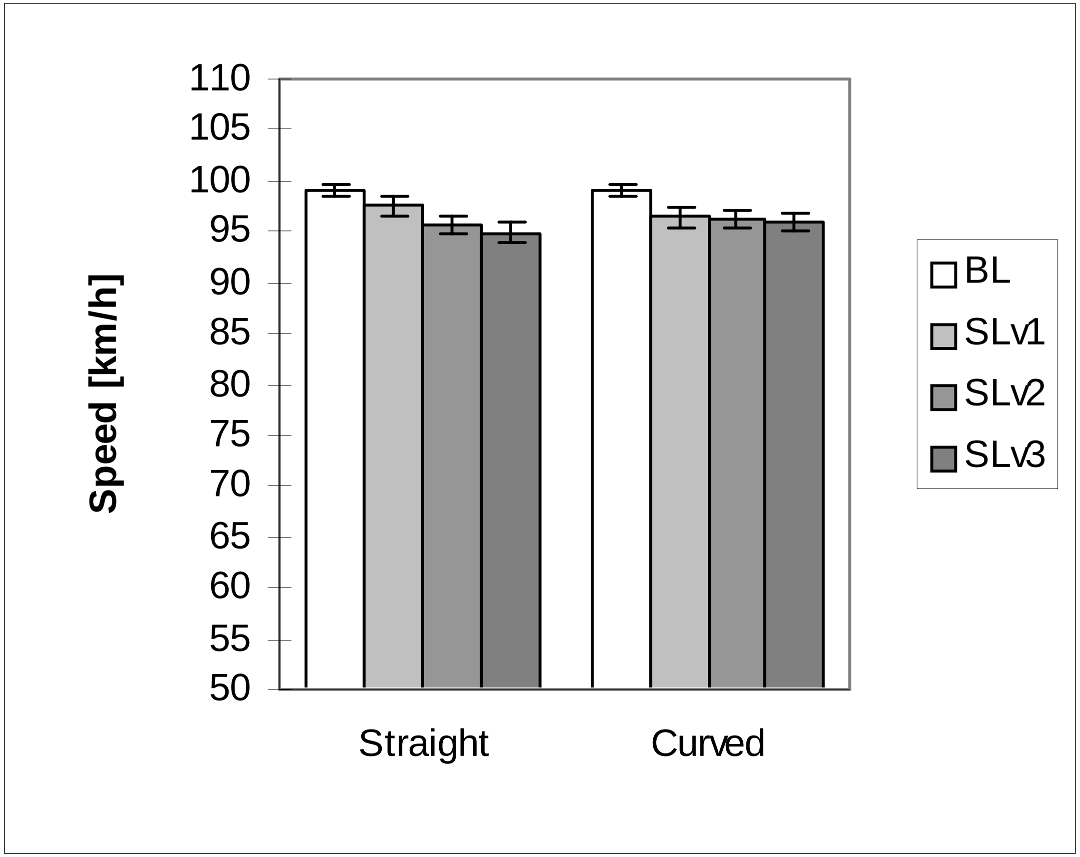
Gaze data distribution

The lateral control was more stable when attending to the **cognitive** S-IVIS



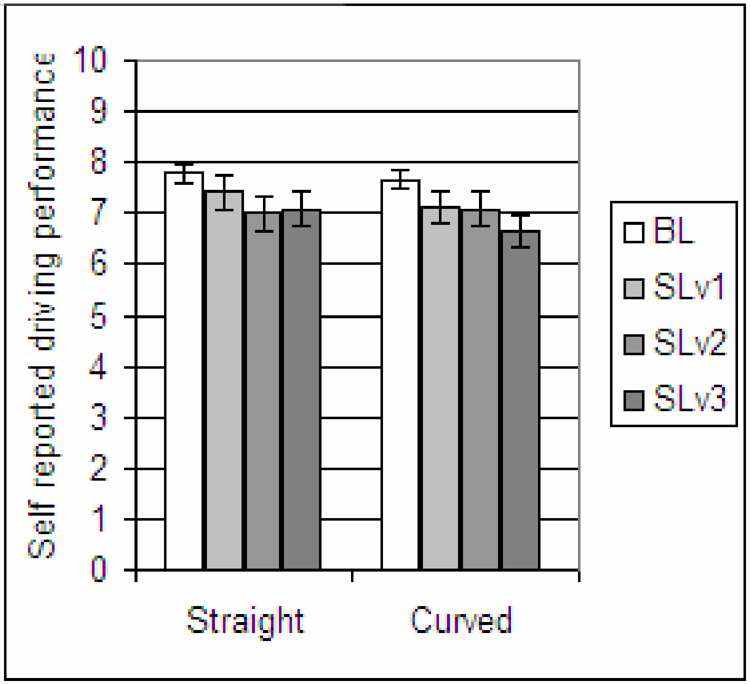
Lateral position variation

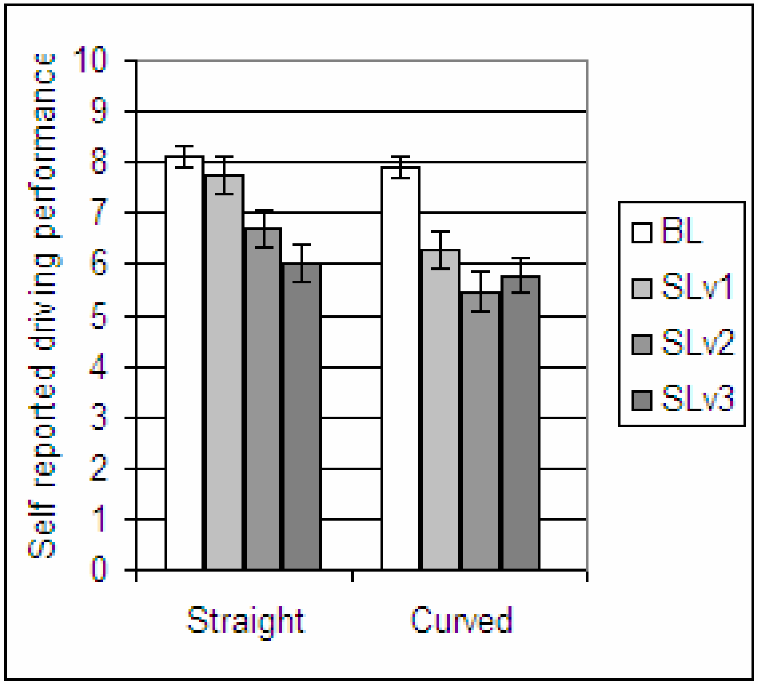
The speed decreased when attending to the visual S-IVIS, but not the cognitive



Mean travel speed

Self reported driving performance decreased for both S-IVIS

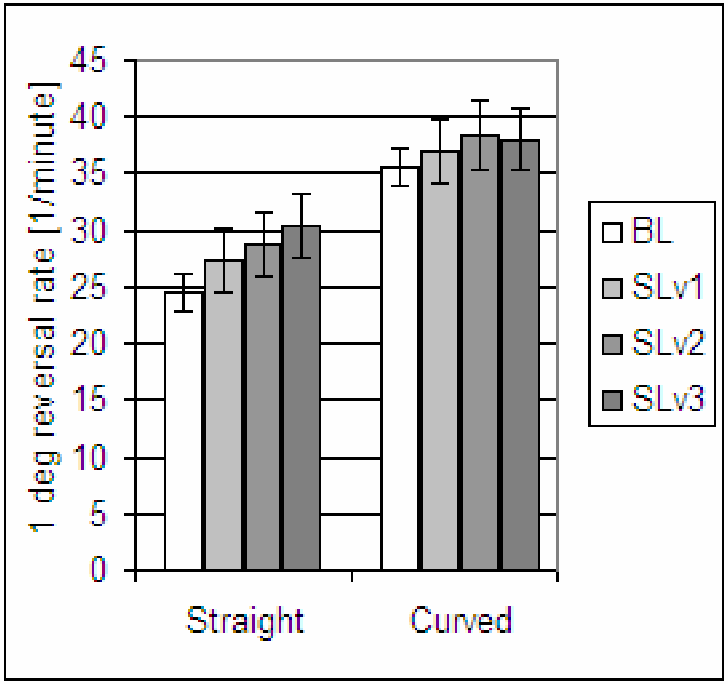




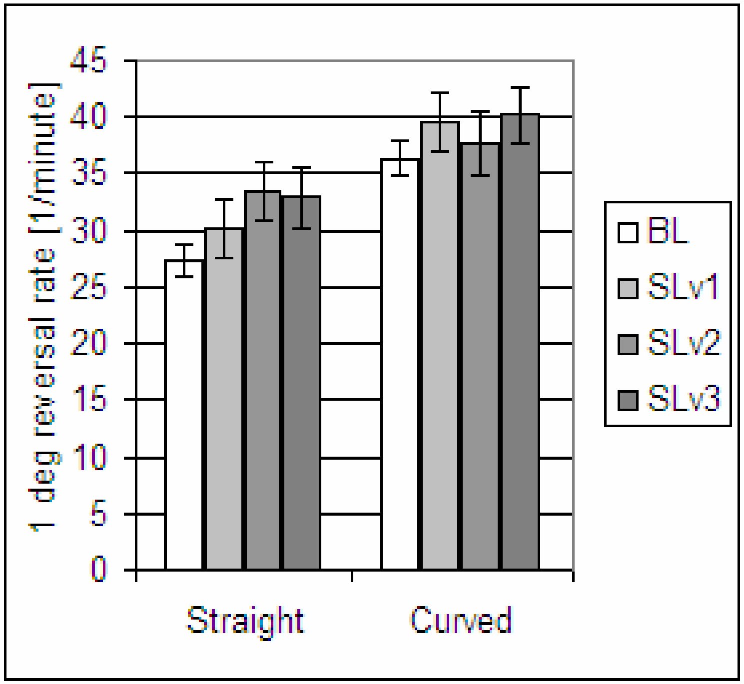
Cognitive

Visual

Steering reversal rate increased for both S-IVIS



Cognitive



Visual

Results - Elderly drivers

- Elderly drivers were worse than average drivers in…
  - managing the trade-off between driving and the S-IVIS
  - driving when attending to the S-IVIS
  - performing the S-IVIS tasks

- Also, elderly drivers were motion sick in the simulator when attending to the visual task

Results –
UK Vs Portuguese drivers

- Portuguese drivers exhibited riskier driving behaviours than UK drivers

Conclusions

- Visual distraction leads to problems in lateral control
- Cognitive load leads to gaze concentration and possibly also loss of peripheral information
- Rural road was most diagnostic in simulator and laboratory
- Motorway was most diagnostic in field
- Elderly drivers were exhibited very risky while performing S-IVIS tasks
- The field studies gave some information not found in the other experiments
- Some included measures were very sensitive and reliable. Some were not.

Thank you for your attention

HASTE Workshop
Brussels
22 March 2005

WP3, Validation of the HASTE protocol specification

Emma Johansson
Volvo Technology

The objectives of WP3

- Application of the methods devised to evaluating real systems

- Recommendation of a draft of a pre-deployment test regime that is both cost effective and possesses the validity to predict performance

The objectives of WP3

- The safety indicators should be
  - Sensitive
  - Reliable
  - Valid

The objectives of WP3

- Since one of the ambitions in HASTE has been to create a test regime as cost effective as possible the experimental set up in WP3 was greatly reduced, based on the results in WP2, with regard to:
  - No. of participants
  - Scenarios
  - Measures

Dimensions

- 4 real In-Vehicle Systems
  - PDA, navigation systems, simulation of a traffic information system
  - Systems assessed on task level
- Assessment methods
  - Simulator, Laboratory and Field
- Road
  - Rural: Laboratory and Simulators
    - Curved and Straight sections
  - Motorway: Field

Scope

- 7 test sites
- 13 experiments
- Approx. 15-20 participants in each experiment

	LABORATORY	SIMULATOR	FIELD
MINHO	C		
LEEDS	B	B	B
T. CANADA		A, D	
VTT			A, C
VOLVO			A, B
TNO		B	
VTI		A, B	

A-priori ranking

- For each system the tasks were a-priori divided into overall complexity level. This ranking was based on number of modalities, number of button presses and manual difficulty level.

Ex. System A

Description	Modality	Task
Route guidance message incl. arithmetic info.	Auditory	1
Route guidance message incl. arithmetic info. – more information than 1	Auditory	2
Route guidance message incl. spatial info. (turn by turn instructions)	Auditory	3
Route guidance message incl. spatial info. (turn by turn instructions) – more information than 3	Auditory	4
Alter volume	Visual-Manual	7
Change one item in map setting	Visual-Manual	8
Change several items in map settings	Visual-Manual	9
Destination entry – City*	Visual-Manual	5
Destination entry – City, Street*	Visual-Manual	6

* Two out of the nine tasks removed in the field experiments.

Measures

- The most promising measures from WP2
- The Peripheral Detection Task was added as a surrogate to the critical events in WP2

Analysis

- Same ANOVA as in WP2
  - Road complexity
  - Tasks
- Analysis of each experiment
- Meta analysis including all experiments

Results – Effects of Secondary Tasks

- Similar to our results in WP2, the effects from our WP3 experiments are more pronounced for the visual and visual-manual tasks
- Again, somewhat different effects for auditory vs. visual content in tasks. Lack in WP3 – not enough auditory/cognitive tasks

WP 3 results → Test Regime

- A recipe for the user (researcher, system engineer, human factors specialist) on how to conduct his/her safety assessment with regards to:
  - Test environment
  - Scenarios
  - Experimental design
  - Dependent measures
  - Safety criteria

HASTE Workshop
Brussels
22 March 2005

From results to regime: What have we learned?

Wiel Janssen

TNO Human Factors

Grasping what a set of some 30 experiments has told us

- Applying statistical meta-analysis
- So that we obtain robust results on sensitivity, reliability, and consistency of effects, and their links to safety
- Always in terms of IVIS difficulty level, relative to baseline; and of its modality (vis/vis-man/cogn)

Results of meta-analysis: effect sizes and task type

	Vis	Vis-man	Cogn	
Subj_R	-2.19	-2.49	-0.97	Of own perf.
MN_SP	-0.62	-0.84	-0.54	Mean speed
HI_ST	0.84	0.88	0.71	High-freq. steering
U_HWT	0.98	1.00	0.91	Min. time headway
PDT_HIT	-0.54	-0.84	-0.53	% Correct to PDT
PDT_RT	0.81	0.82	0.60	RT to PDT
PR_C	-2.74	-1.94	0.65	% in center

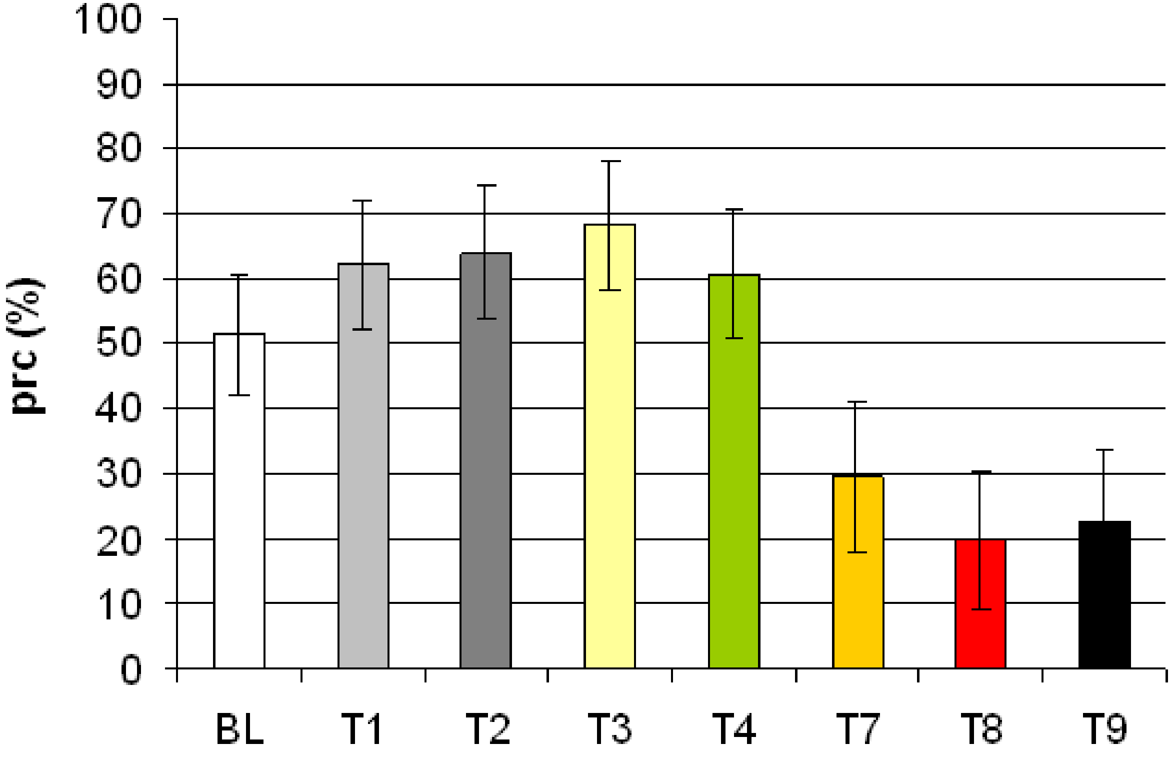
(NB: 0.20 = small, 0.50 = medium, 0.80 = large)

So which one(s) to select?

- That's what we want your opinion on
- (We have our own)
- Luckily enough, there is a lot to choose from
- Is pre-classification of system useful, so to tune selection?

Results: Percent Road Centre

- Example from Volvo's field results on System A
- Task 1-4: auditory, the others: visual-manual



Summary

	Strengths	Weaknesses	Opportunities	Threats
Subjective Rating (Subj_R)	Fast, cheap	Subjective, perceptions of driving performance may not be the same as actual performance	Different rating scales can be developed	Manipulation of data from instructions to participants
Mean Speed (MN_SP)	Easy signal to measure, on-road/sim	Safety interpretation of speed effects - speeding vs slowing down. Speed needs to stabilize to normal level again between tasks.		Slowing down may not be a relevant criterium for classification as unsafe.
Steering (HI_ST)	Easy signal to measure, on-road/sim, relevant			May reflect increased effort or sensitivity to steering error and not necessarily represent a threat to traffic safety.

Summary

	Strengths	Weaknesses	Opportunities	Threats
Minimum Headway (U_HMT)	Relevant	Needs lead vehicle, needs distance sensor	Different rating scales can be developed	Resource demanding in real traffic
Percent road centre (PRC)	Measures perceptual performance, relevant, high face validity, easy to calculate (much easier than glance measures)	Currently expensive hardware, Not calculated in all studies (Haste), needs eye tracker	Can be developed as inexpensive, easy to use tool. Can easily be used in product development.	
Peripheral Detection Task (PDT_RT; PDT_HIT)	Measures perceptual performance and reaction time, relevant, high face validity, easy to calculate	Somewhat intrusive, may effect other measures, Not calculated in all studies and not sufficient statistical reliability (Haste)	Can be augmented with other event detection stimuli	

Practicalities for a regime

- For design stage as well as final assessment
- Number of subjects: only 10-15
- Age between 25 and 50, M&F, sufficient driving experience (10 k annually, at least 5 yrs licence)
- Environment: at least medium-range simulator; rural road type
- Duration per task: about 10 min
- A single baseline ride is required (10 min)
- So full evaluation per system will take about 2 days, not including overall set-up

HASTE Workshop
Brussels
22 March 2005

What could happen to the HASTE regime

Oliver Carsten

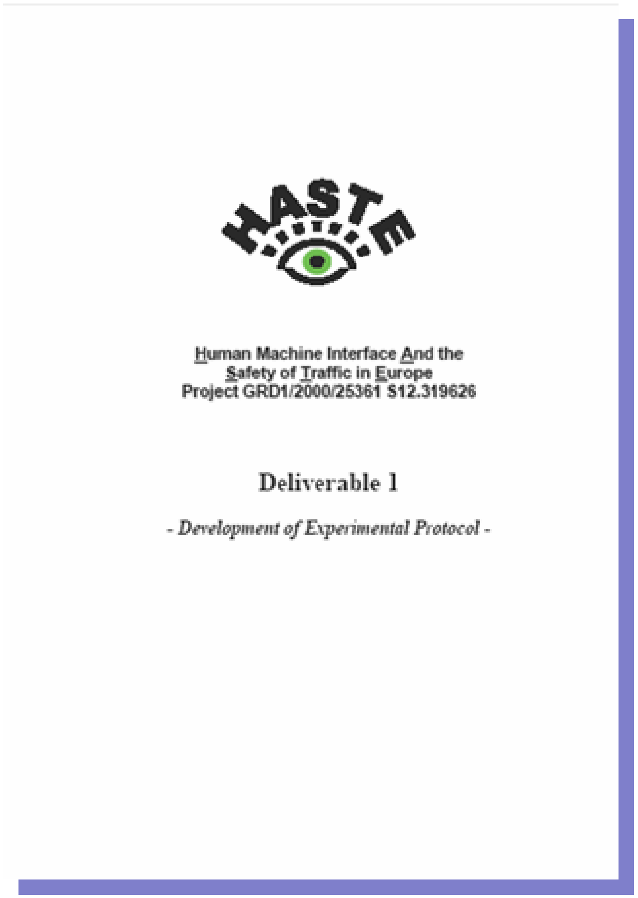
Institute for Transport Studies

University of Leeds

Choices

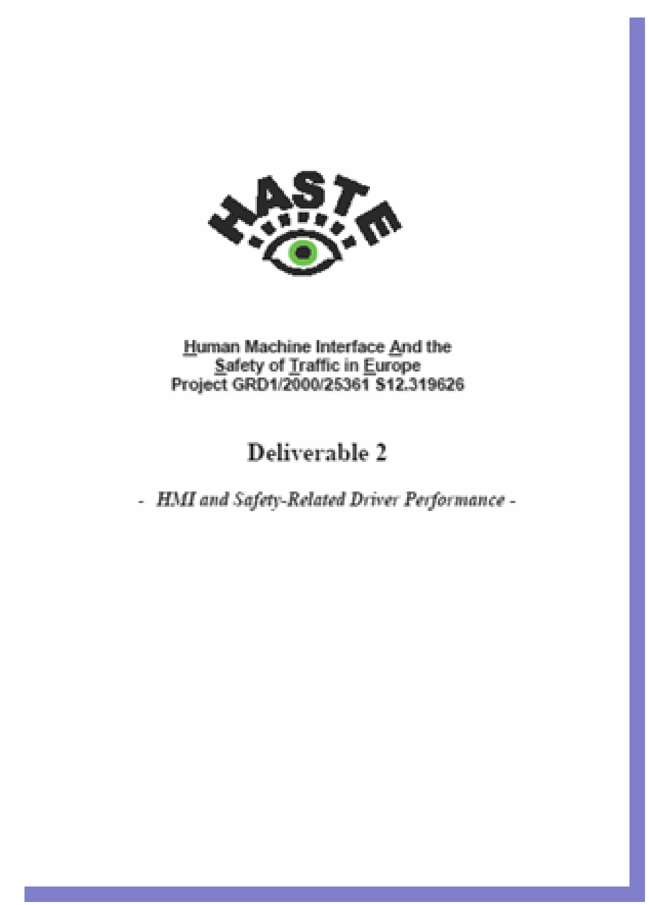
1. HASTE outputs remain as research
2. Enforced by legislation
   a. EU
   b. National
3. Issued as Commission Recommendation
4. Adopted voluntarily perhaps backed up by ISO
5. Used as consumer information (P-NCAP)

Research


Human Machine Interface And the
Safety of Traffic in Europe
Project GRD1/2000/25361 S12.319626

Deliverable 1

- Development of Experimental Protocol -


Human Machine Interface And the
Safety of Traffic in Europe
Project GRD1/2000/25361 S12.319626

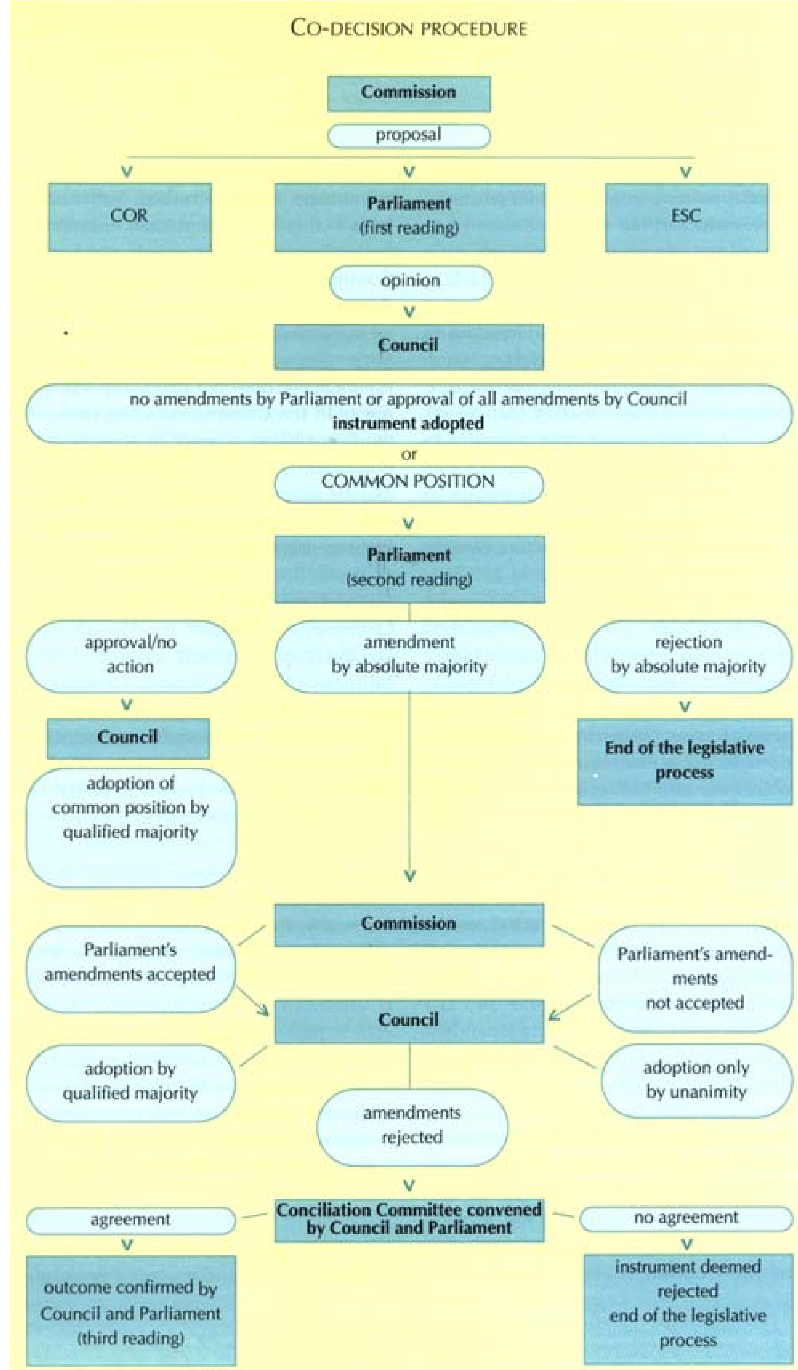
Deliverable 2

- HMI and Safety-Related Driver Performance -

Legislation

- An eventual EU directive was perhaps the original HASTE vision
- *But:*
  - Sets only a minimum threshold
  - Question of which systems have to be tested:
    - PDAs?
    - Mobile phones?

Legislation



Commission Recommendation

- Effectiveness?

Voluntary adoption

By whom:

- ACEA?
- CLEPA?

P-NCAP

- Currently vehicles are rated on secondary safety (protection of occupants and pedestrians)
- Plans to add rating on primary safety
- Subgroup on Ergonomics and Driver Assistance

Consumer information — issues

- Pros
  - Allows a range of scores, not just pass/fail
  - Any device (or feature of a device) could be assessed
- Cons
  - Not legally binding
  - Who pays the cost of testing?
  - Would consumers pay attention?

MINISTRY OF TRANSPORT
AND COMMUNICATIONS FINLAND

HASTE - Finnish experiences

Driver's HMI and Govmnt role

Anu Lamberg

Senior Adviser

MINISTRY OF TRANSPORT AND COMMUNICATIONS FINLAND

Background

- The aim of HASTE was to develop methodologies and guidelines for the assessment of In-Vehicle Information Systems (IVIS).
- There are many standards and guidelines available:
  - In Europe, European Statement of Principles on Human Machine Interface for In-Vehicle Information and Communication Systems as well as TRL safety Checklist for the Assessment of in vehicle information systems.
  - In the U.S., SAE standards, Battelle Guidelines and UMTRI Guidelines.
  - In Japan, JAMA Guidelines.

Background (2)

- **However, the guidelines are too general for the practical work of the industry and traffic safety authorities.**

- ***The safety relevance of the behavioural indices remains unclear as in almost all cases no proven explicit relationships exist between current HMI indices and accident risk.***

MINISTRY OF TRANSPORT
AND COMMUNICATIONS FINLAND

HASTE approach

- **The area of the project was significant and timely - the car has become a potential home to many different types of new systems.**

- ***There were clear societal needs associated with this project.***

MINISTRY OF TRANSPORT AND COMMUNICATIONS FINLAND

The objectives of the project were extremely challenging

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

Results

- HASTE produced a number of important and significant results.

- The results are theoretically and methodologically interesting and should be distributed to be dimlpemented and further developed.

- At the same time, they have practical relevance.
- However, further research and steps are needed to identify quantitative relationships between task load and road traffic risk.

MINISTRY OF TRANSPORT
AND COMMUNICATIONS FINLAND

Finland's role in HASTE

- VTT contributed to the project with extensive field studies.
- The field studies were conducted in three environments, i.e. in urban and rural areas as well as on motorways.
- Both average and elderly drivers were included in the studies.
- The studies produced valuable results, e.g. about the effects of IVIS in urban areas.
- Tool to improve traffic safety in MS Finland and ...

MINISTRY OF TRANSPORT
AND COMMUNICATIONS FINLAND

Government's role and the effects of IVIS

- **In-vehicle information and communication systems are designed to improve traffic safety and efficiency – ARE THEY?**
- **However, it has been recognised that there are potential negative safety effects,**
- **Knowledge is neede**
- **From the point of view of the road safety authorities, there is an urgent need for a research-based set of performance standards for in-vehicle human machine interfaces + other steps.**

Government's role and the effects of IVIS

- Awareness rising among all stakeholders: authorities, manufactures
- Public opinion
- Legislation on EU and MS level
- Competition between public transportations and private cars: travelling, working, entertainment
- etc
- THANK YOU!

INDUSTRY NEEDS & INTERESTS

- HASTE FINAL WORKSHOP -

KARIN SVENSSON
VOLVO TECHNOLOGY

BRUSSELS

AGENDA

- **Introduction**
  - **problem to solve**
  - **challenge**
  - **aim for in-vehicle information systems**
- **Background**
  - **user centered design approach**
  - **iterative and cost efficient development process**
  - **development process**
  - **test environment**
- **Industry needs & interests**
  - **aim of HASTE**
  - **key problems**
  - **requirements on test regime**
  - **requirements on measurements**
  - **what is needed to achive this?**
- **Conclusions**

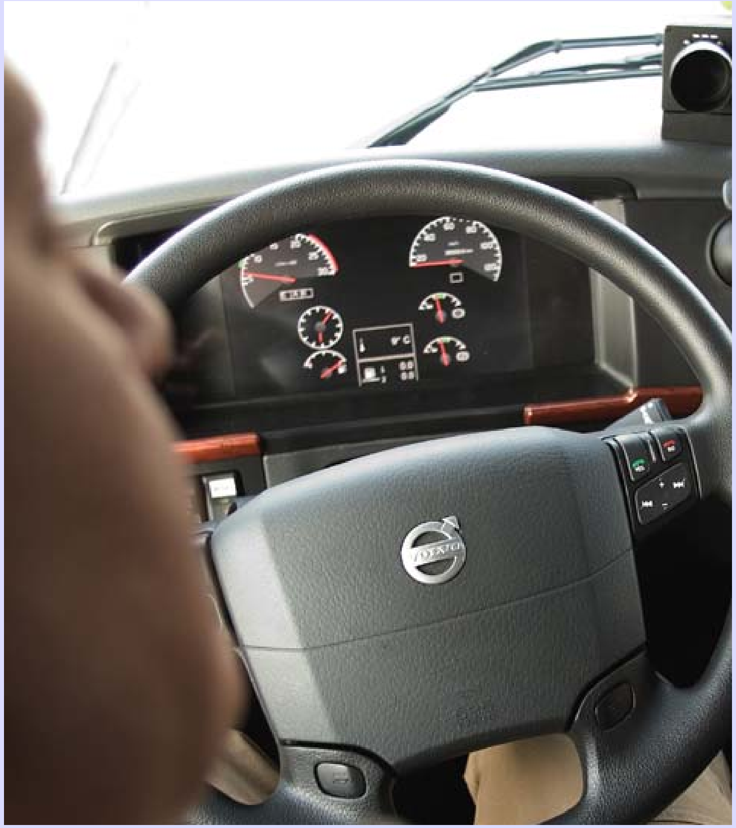
INTRODUCTION

PROBLEM TO SOLVE

CHALLENGE

- Increasing number of functions
  - **customer / market push**
  - **technology push**
  - **independent systems**
  - **increased number of functions have both positive and unwanted effects**
- Stressful working/driving situation
  - **traffic congestion**
  - **increasing demand on high productivity**
- Driver overload

AIM FOR IN-VEHICLE INFORMATION SYSTEMS

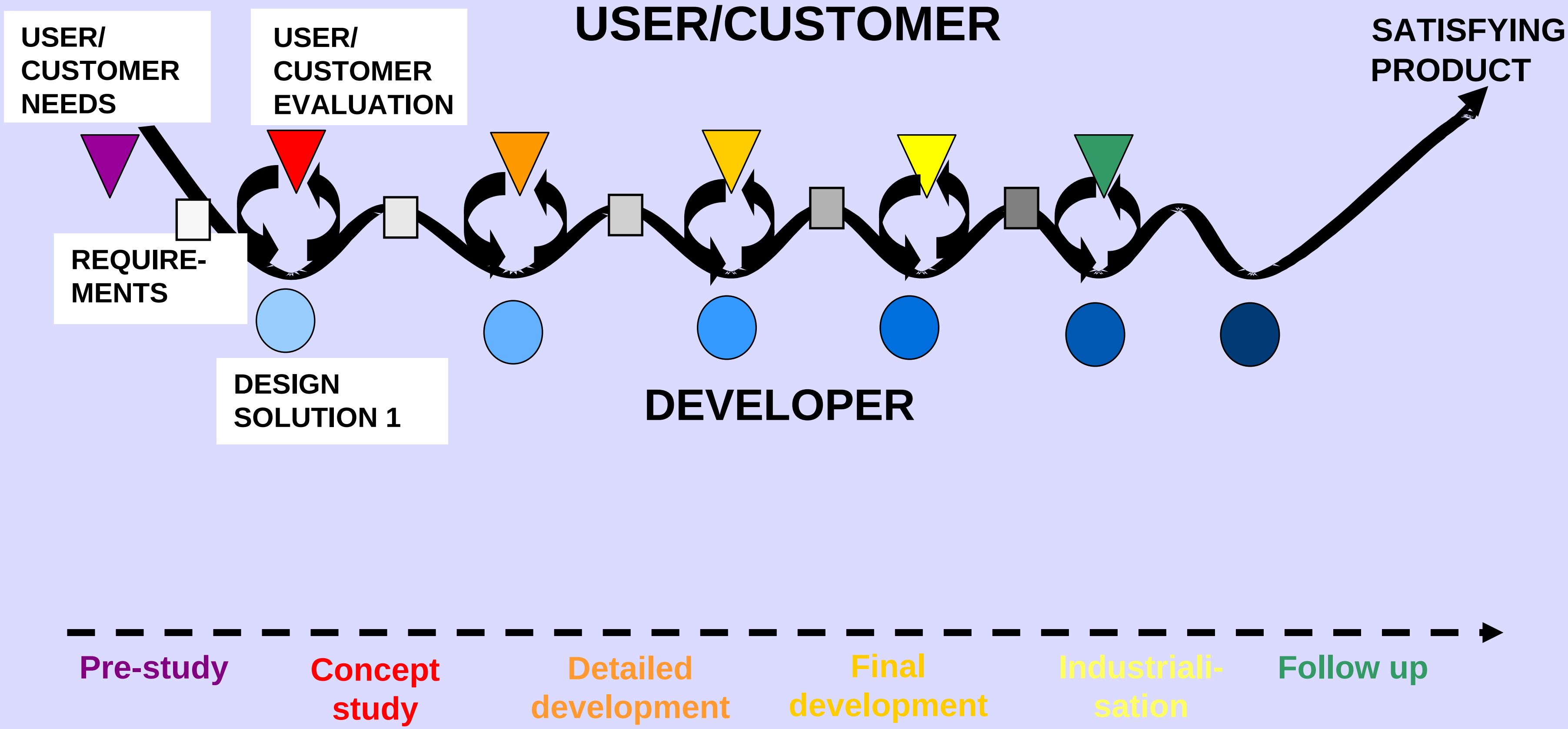


- ✓ Safe
- ✓ Reliable
- ✓ Efficient
- ✓ High usability
- ✓ High acceptance

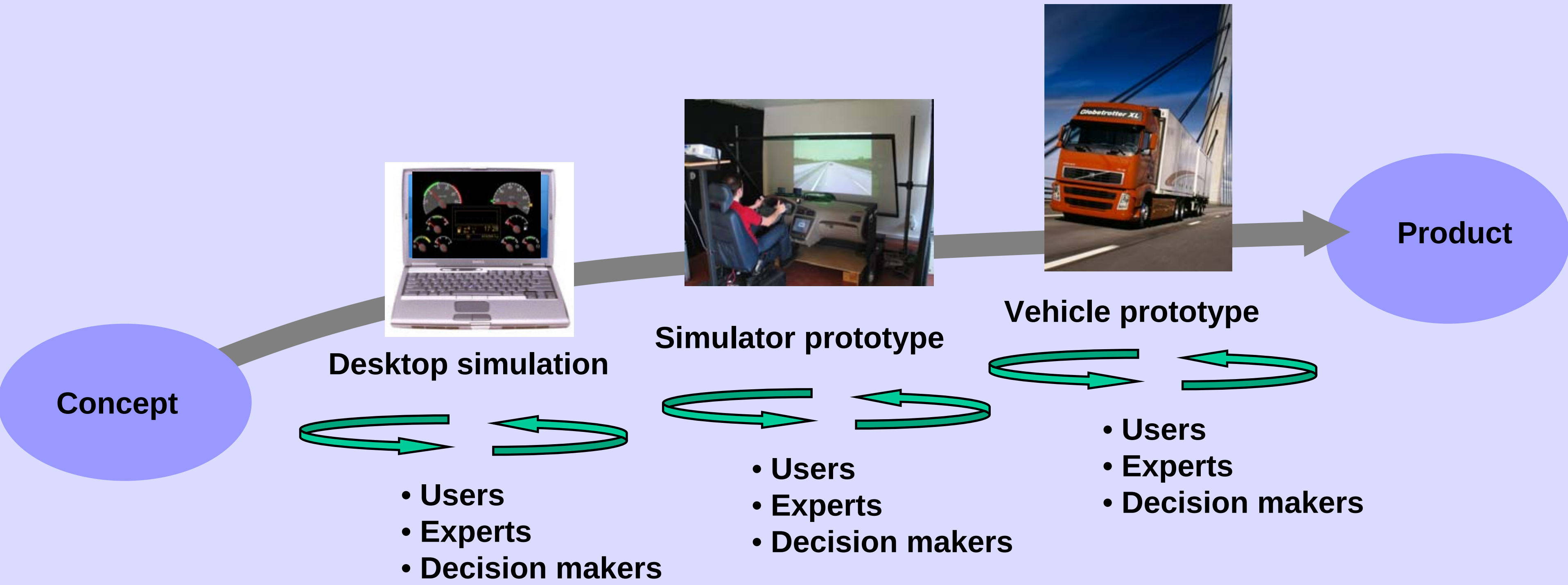
A development process which supports this!

BACKGROUND: DEVELOPMENT PROCESS

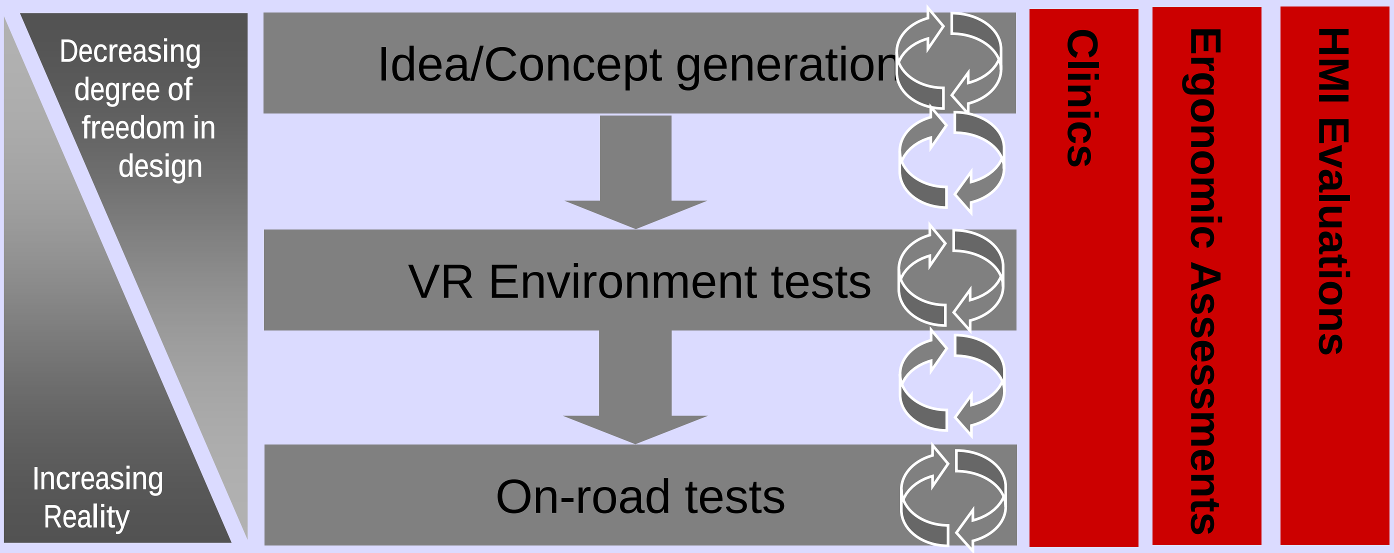
USER CENTRED DESIGN APPROACH



ITERATIVE AND COST EFFICIENT DEVELOPMENT PROCESS



DEVELOPMENT PROCESS



- **Inquiry** – surveys, questionnaires, interviews, FGDs
- **Inspection** – Heuristic evaluation, checklists, cognitive walkthroughs
- **Simulations** – ergonomics simulations, mock-up evaluations
- **Testing** – prototyping, performance measurements

TEST ENVIRONMENT

INDUSTRY NEEDS AND INTERESTS

AIM OF HASTE

- To develop ***methodologies and guidelines*** for the assessment of In-Vehicle Information Systems (IVIS).
- To present an outline for a ***test regime*** which could be used both throughout the ***design process*** at IVIS manufacturers as weel as in later stages for ***final verification and ceritification***.

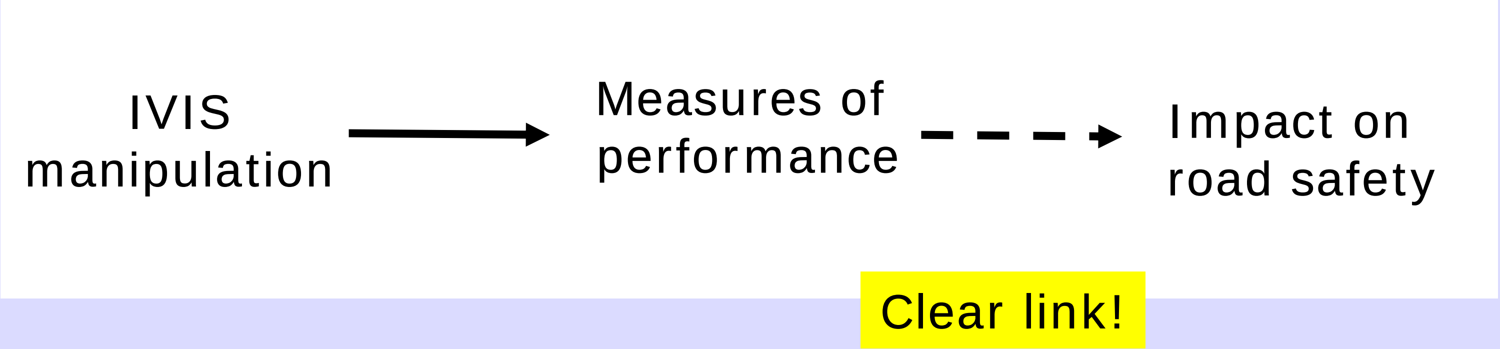
- Ideally, the test regime would ***specify methods and tools which would***:
  - ✓ Be technology independent
  - ✓ Have safety-related criteria
  - ✓ Be cost effective
  - ✓ Be appropriate for any system design
  - ✓ Have been validated through real-world testing

KEY PROBLEMS

1. Proliferation of methods, tools and performance metrics

- no consensus on when to use which method, tool and performance metric
- difficult to compare results from different studies
- HMI evaluation studies are costly and requires strong expertise

2. How to infer safety effects from measures of performance



REQUIREMENTS ON TEST REGIME 1(2)

Should support assessment at different stages of development

- Formative: Goal to improve design
- Summative: Verification, certification

Should allow for testing of different hypotheses

Should take into account system characteristics

REQUIREMENTS ON TEST REGIME 2 (2)

Should be associated with an agreed set of design guidelines (e.g. ESoP)

Should specify safety criteria

Should be cost-efficient and easy to use

No ”pass or fail” criteria!

REQUIREMENTS ON MEASUREMENTS

- ✓ high validity
- ✓ high reliability
- ✓ sensitive
- ✓ cost efficient
- ✓ simple to use

WHAT IS NEEDED TO ACHIVE THIS?

Better understanding of the effects of individual and combined in-vehicle systems on workload and performance

Better understanding of how driver errors cause accidents

- Not enough to say that 95% of accidents are caused by driver error...
- Difficult to infer the detailed causal chain from accident databases
- Promising approaches
  - In-depth on-site accident studie & incident and conflict analysis (e.g. Swedish national project ”Factors Influencing the Causation of Accidents and Incidents”)
  - Naturalistic field studies (e.g. US 100-car study)

CONCLUSIONS

CONCLUSIONS

Solid, structured and broad review of possible measurements

Positive to the development of cost-efficient and easy-to-use assessment methodologies

Reluctant to pass/fail criteria

Vehicle manufacturers are active in delivering easy to use and safe IVIS

FUTURE WORK AND NEEDS

Continue development of cost efficient and easy to use measurements

Continue development of “final” test regime

to use during development and for verification / certification

better understand the links between criteria and traffic safety

Continue work in i.e. AIDE and ISO groups

THANKS FOR YOUR ATTENTION!