Global evaluation methodology

SP 2.1 work description

Fiat Research Centre: Elisabetta Nodari

SP 2 General organization

SP 2

to develop a cost efficient and industrially applicable methodology for quantifying behavioural effects of IVIS and ADAS functions, and their relation to road safety

SP 2.1
Generic evaluation methodology

To develop a generic and cost efficient methodology for industrial human factors safety evaluation of integrated IVIS and ADAS.

Approach

The methodology developed should specify **methods**, **tools**, **scenarios** as well as guidelines and how and when apply them

SP 2.2
Workload and Distraction assessment method and tools

While WP2.1 will focus on the development of a general evaluation and assessment methodology, WP2.2 will focus on specific methods and tools for workload and distraction measurement, to be incorporated into the general methodology

SP 2.3
Estimating the risk reduction potential of Integrated Adaptive HMI

The objective of this WP is to develop methods for extrapolating from behavioural effects of IVIS and ADAS (e.g. the workload and distraction metrics developed in WP2.2), to actual road safety

2.1 Cookbook

To provide a detailed guide that allows readers to develop and apply an appropriate evaluation methodology of integrated as well as non - integrated ADA/ IVI Systems.

2.4 Prototype evaluation

To perform the final evaluation of the three AIDE prototype vehicles developed in SP3. The main objective is to assess the AIDE concept with respect to its potential safety-enhancing effects (compared to non-integrated, non-adaptive IVIS and ADAS and driving with no IVIS/ADAS at all).

SP 2.1 Preparatory tasks

Which evaluation methods and tools are present in literature?

2.1.1
Review of existing tools and methods

The aim is to collect methodologies for the evaluation of **usability** and **acceptability** of IVIS/ADAS in-vehicle applications during **different steps of system development**

Which IVI and ADA systems were developed until now?

2.1.2
Review and taxonomy of IVIS/ADA applications

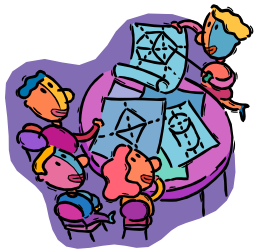
The main objective is to carry out a reasoned analysis of the IVIS (In Vehicle Information System) and ADAS (Advanced Driver Assistance System) applications in order to organize them in several categories

Which typology of scenario could be applied to different type of IVI/ ADA systems

2.1.3
Scenario development

The main objective of this document is to present a review of issues relevant to the field of testing environments (scenarios, simulators and/or driving environments, cohorts, use cases, etc.).

Output 2.1.1: Review of existing tools and methods



The state of the art of the methodologies used to evaluate IVIS and ADAS applications was collected. In particular, this activity focused mainly on techniques **used to estimate user mental model, usability and acceptability.** This work is strictly linked with 2.2.1 (focused on Workload) work and they are the starting point to define a common framework of tools and methods in order to conduct behavioural and usability evaluations of IVI / ADA systems and integrated ones.

The work provided an exhaustive overview about:

- The design cycle based on the **User-Centred Design**, the starting point to develop usable systems, prone to be positively accepted by final user;

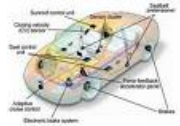
- Specific dimensions, important to the Design Cycle, as User Mental Model, Usability, Acceptance, Subjective Workload, Situation Awareness.

- Relative methods and techniques to evaluate these dimensions, beyond experimental description and main results;

At the end, it was derived :
- A list of dimension to take into consideration during a behavioural test. The dimensions are grouped on the basis of different modalities of interactions;
- Step-By-Step procedure for Test Design considering different phases of system development.

Output 2.1.2: Review and taxonomy of IVIS/ADA applications

A reasoned analysis of the IVIS (In Vehicle Information System) and ADAS (Advanced Driver Assistance System) applications was carried out and organized on the basis of different dimensions. The output of this work is the identification of a group of applications or prototypes to be used in later tasks of SP 2 (2.2 and 2.1.4)



The work provided :

- an extended database of existing and coming IVIS/ADAS applications;

- a classification of these systems on the basis of:

  - the availability of the system on the market or its existence in the form of prototype;
  - whether the system is realized or realizable (future application);
  - which system requirements have to be changed to allow its use in existing conditions;
  - what are the criticality and the advantages of the system in terms of safety, usability and acceptability;
  - which are the requirements to the evaluation procedures/tools derived both from the specific features and the various steps of the design process (requirements with regard to validity, reliability and practicability)

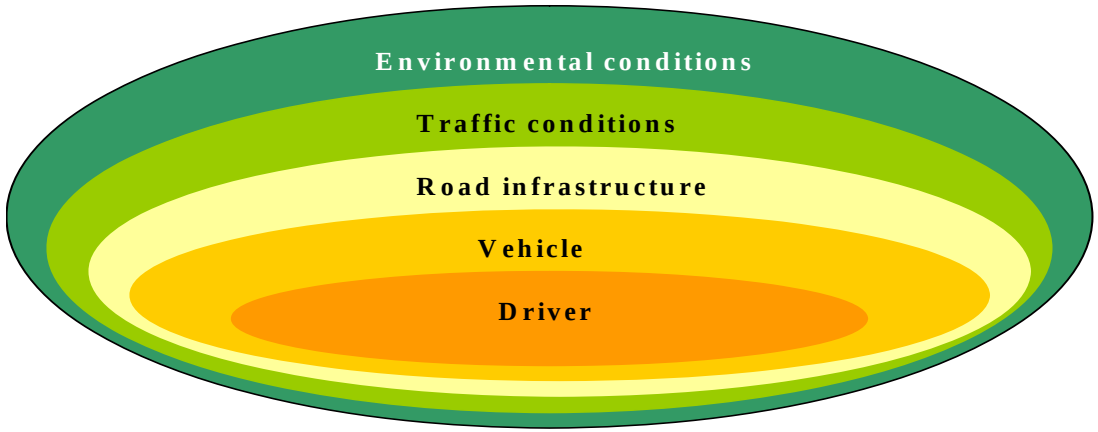
Output 2.1.3: Scenario development

The work focused on the “building blocks” of scenario development including considerations on simulator and cohort choices.

The components identified as necessary to consider in scenario building are:

- the vehicle
- the road infrastructure
- the traffic conditions
- the environmental condition

The general scope is support the choice, implementation and execution of experimental trials to be carried out in WP 2.1.4 for the assessment and specification of the best evaluation methodologies for different IVIS/ADAS, their adaptivity and integration.



Output 2.1.3: Scenario development

The work provided :

- an exhaustive classification of driving simulators
  - scenario for specific drivers cohorts

- relevance Table of Scenario Building Blocks

General means in the ratings

Possible Scenarios	Road type & conditions, visibility						Traffic type and actors						Tasks and goals					
	Type of road: City, Interu.	Type of road: Motorway	Type of road: rural	Wet conditions	Visibility conditions	Weather conditions	Traffic density and driving style	On coming traffic	Crossing traffic	Pedestrians	Relaxed driving	Car following	Lane Change Task (LCT)	Occlusion	Distraction task	Blind spot detection and lane change task	City driving
IVIS:																		
Navigation Systems	3.5	1	2.5	2.5	1	1	1	1	1	1	3	3	1	1	1	1	3	3
Travelling/Traffic Related Information Systems	3	1	2	2	1	1.5	1.5	1	1	1	2.5	2	1	1	1	1	2	2
Vehicle Communication Systems	3	1	3	3	1	1	1	1.5	1	1	3	3	1	1	1	1	3	3
Driver Convenience Systems	3	1	3	3	1	1	1	1	1.5	1	3	3	1	1	1	1	3	3
ADAS:																		
Lateral Control	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Lane Keeping and warning	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Blind spot monitoring	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Lane change and merge collision avoidance	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Longitudinal Control	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Intelligent Speed Adaptation	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Road Low Friction Warning Systems	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Reversing/Parking Aid	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Vision Enhancement	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Driver Monitoring	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Pre-Crash Systems	3		3	3	3	3	3		3		3	3	3	3	3	3	3	3
Vulnerable Road Users Protection Systems	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

- suggested simulator category and AIDE design scenario for each ADAS/IVIS

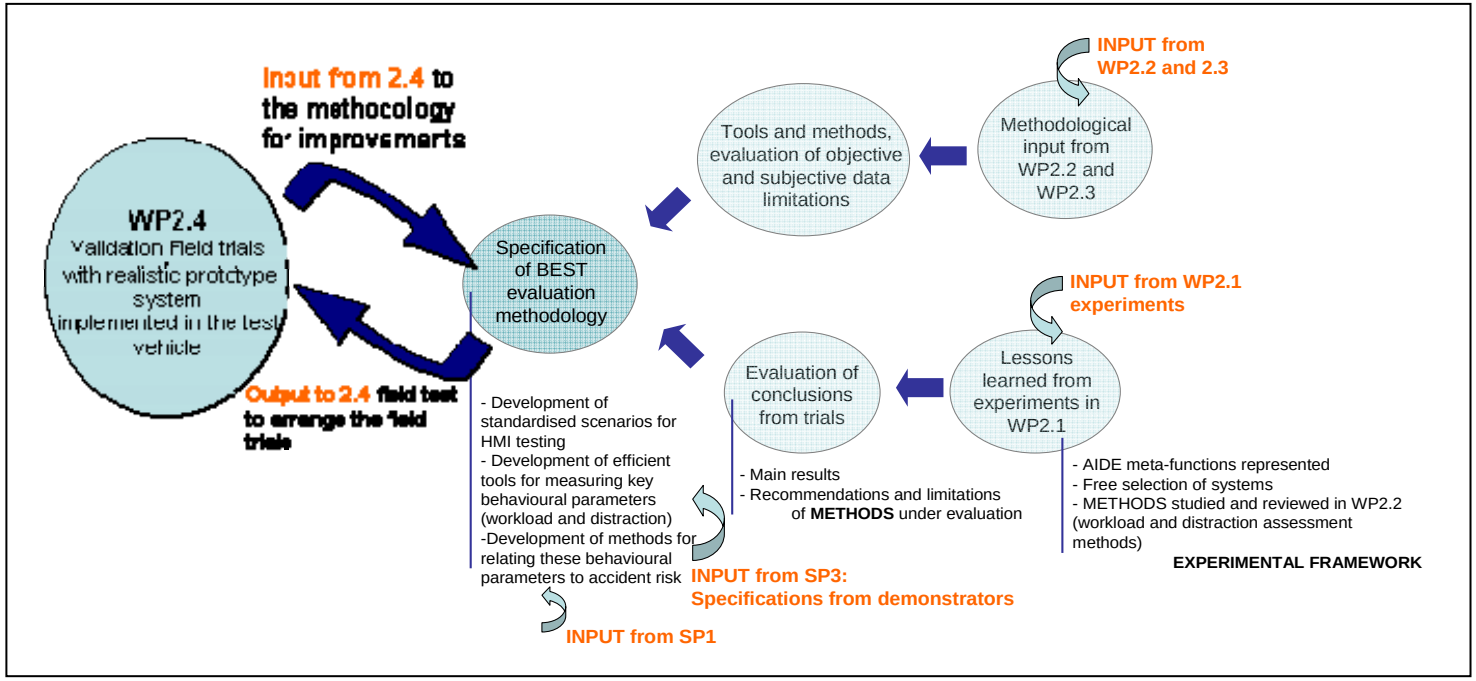
	Scalability over types of simulators							Vehicle on the road	Examples of AIDE driving scenarios
	A	B	C	D	E	F	G		
IVIS:									
Navigation Systems	LS	S	P	S	OD			S	1.5 or 2.1
Travelling/Traffic Related Information Systems	LS	S	P	S	OD			S	1.7 or 2.2
Vehicle Communication Systems	LS	S	P	S	OD			E	1.3 or 1.4
Driver Convenience Systems	S		P	S	OD			S (evtl. D)	1.2 or 1.6
ADAS:									
Lateral Control								S (evtl. D)	3.3
Lane Keeping and warning	LS	S	P	S	OD			S (evtl. D)	3.3
Blind spot monitoring	NS	NS	S	P	S			S (evtl. D)	3.3
Lane change and merge collision avoidance	NS	NS	LS	P	S			D	3.3
Longitudinal Control	S	S	S	P	S			S	3.3
Intelligent Speed Adaptation	NS	S	P	OD	OD			S	3.3
Road Low Friction Warning Systems	NS	NS	LS	P	S			T (S)	3.3
Reversing/Parking Aid	LS	S	S	S	S		P	S	3.3
Vision Enhancement	NS	NS	LS	LS	LS			S, T	3.1
Driver Monitoring	NS	NS	P	S	OD			D, T	2.3 or 2.4
Pre-Crash Systems	NS	NS	NS	LS	S		S	D, T	1.1
Vulnerable Road Users Protection Systems	LS	LS	LS	LS	LS	S	S	D, T	3.3

- scenario examples for AIDE system with conflicts between concurrent actions

AIDE design scenario 1.1: Conflict between concurrent warnings **(ADAS)**
Action/s = {W + W}, where W = Warning indicating high traffic risk.
DVE Condition/s: DVE_1= LOW/NO (otherwise see 3.5), $DVE_{2,3,4}$=irrelevant Traffic/environment Risk (DVE_5)=HIGH.
Flow of events 1. W_1 is initiated 2. W_2 is initiated while W_1 is executing
Example *Conflict scenario:* The driver swerves out of the lane. The lane departure warning system gives a warning (**Lat**). However, at the same time, the vehicle ahead brakes suddenly, triggering the forward collision warning (**PCS**).

Sp 2.1.4 the operative task

The previous revision tasks and experimental studies carried out in 2.1, 2.2 and 2.3 tasks were used to select the best methodologies and tools for the evaluation of INTEGRATED and ADAPTIVE system



Sp 2.1.4 the operative task

Four outputs

- a description of **several experimental studies** to test systems, scenario, methods and tools provided by previous review tasks (partners involved were REGIENOV, HIT/ CERTH study; USTUTT, BMW, BOSCH, TNO)

- An **exhaustive questionnaire** with the aim to measure different dimensions of usability and acceptability (partner involved was CRF)

- A **cookbook**: a chapter that suggests specific steps that should be followed to define the evaluation procedure of the AIDE final demonstrators in WP2.4 (Prototype evaluation). However, the AIDE cookbook is general enough to be successfully applied to other evaluation demands that are not directly linked to the AIDE project (partners involved were TNO, BMW, CIDAUT, CRF)

- **Transformation of behavioural/driver state data to risk**

Sp 2.1.4: CRF questionnaire

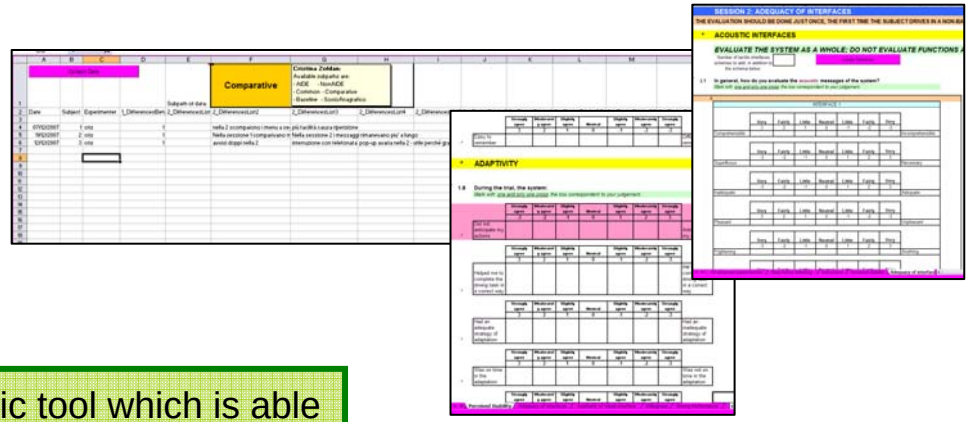
CRF **developed a subjective tool**:
- as general as possible, in order to apply it to different systems both adaptive and not adaptive;

- focused on **the aspects** that seem to be **important to understand whether a system could be acceptable by potential final users** and therefore successful on the market;

- with a modular structure, divided in different sub-dimensions which can be applied totally or partially, accordingly to the specific analysed system.

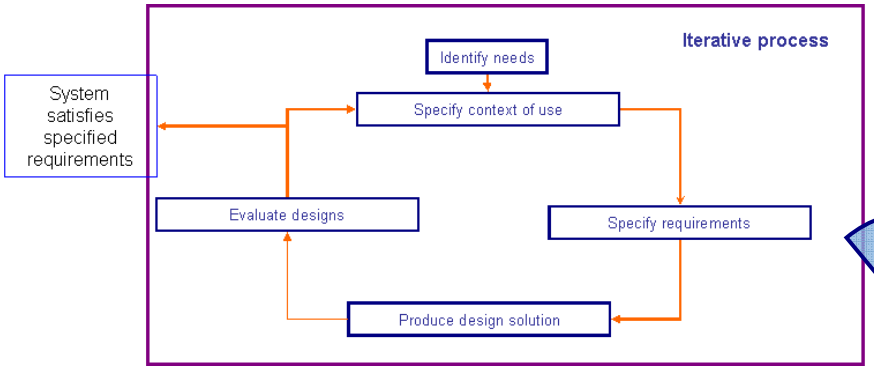
The sub-dimensions explored are

- **Perceived usability**
- **Adequacy** of different modalities
- (visual, haptic, auditory)
- **Perceived Utility**
- **Perceived Safety**
- **Reflective Dimensions**

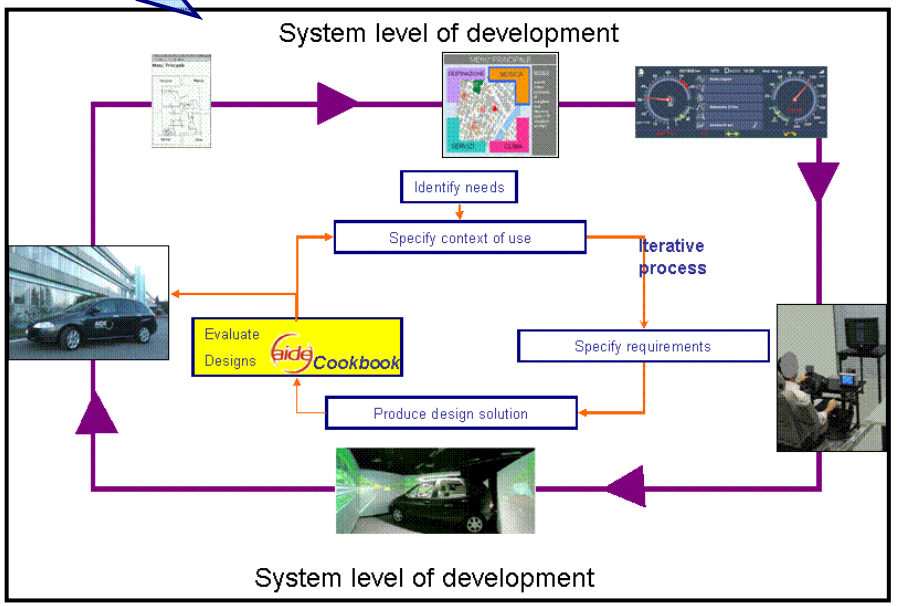


The questionnaire was implemented as an electronic tool which is able to automatically collect and organize data.
The organized data can be easily imported into several statistical tools (i.e. SPSS)

Sp 2.1.4: the cookbook



The UCD process



AIDE Cookbook in the UCD process

Sp 2.1.4 The Cookbook

The aim is to provide a detailed guide that allows to develop and apply an **evaluation methodology** with end-users of integrated as well as non- integrated ADA/ IVI Systems.

All the results of the research carried out within AIDE SP2 are used here to **create a practical instrument for Human Factor experts to develop an evaluation methodology** able to maximize the different types of validity and the reliability

The cookbook describes the following steps to carry out the evaluation:

- To define the aims of the evaluation
- To describe the system to be evaluated
- To define a scenario
- To define the subject sample
- To define subjective and objective parameters and instruments to collect them
- To define the experimental design
- To develop experimental instruction both for participants and for experimenters
- To finalize the experimental set-up
- To carry out the experiment
- To analyse the data
- To produce summary indications about the system

Sp 2.1.4 The Cookbook: an example the first step

- **To define the aims of the evaluation**

System level of development	Paper mock-up (only for initial development phases)		Bench Prototype (simulated an a PC computer or with its own display and junction box);		Simulated prototype		On-Board prototype		Final on-board system already on the market;	
	IMS	ADAS	IMS	ADAS	IMS	ADAS	IMS	ADAS	IMS	ADAS
Usability evaluation first impact	**Suitable for**		**Quite Suitable**	**Quite Suitable**	**Preferable**	**Suitable/ Preferable**	**Preferable**	**Suitable/ Preferable**	**Suitable**	**Suitable**
	Effective	**Not suitable**	Effective	Effective	Effective	Effective	Effective	Effective	Effective	Effective
	Efficient		Efficient	Efficient	Efficient	Efficient	Efficient	Efficient	Efficient	Efficient
	Error Tolerant		Error Tolerant	Error Tolerant	Engaging	Engaging	Engaging	Engaging	Engaging	Engaging
	Easy to Learn		Easy to Learn	Easy to Learn	Error Tolerant	Error Tolerant	Error Tolerant	Error Tolerant	Error Tolerant	Error Tolerant
					Easy to Learn (first impact)	Easy to Learn (first impact)	Easy to Learn (first impact)	Easy to Learn (first impact)	Easy to Learn (first impact)	Easy to Learn (first impact)
	Not suitable for		**Not suitable**	**Not suitable**	Adaptivity (first impact)	Adaptivity (first impact)	Adaptivity (first impact)	Adaptivity (first impact)	Adaptivity (first impact)	Adaptivity (first impact)
	Engaging		Engaging	Engaging		It depends on the scalability of simulator				
	Adaptivity		Adaptivity	Adaptivity			it could be dangerous	it could be dangerous	it is expensive to obtain objective logging data. it could be dangerous	it is expensive to obtain objective logging data. it could be dangerous
Usability evaluation long term	**Not suitable**	**Not suitable**	**Not suitable**	**Not suitable**	**Suitable**	**Suitable**	**Preferable**	**Suitable/Preferable**	**Suitable**	**Suitable/less suitable**
					Effective	Effective	Effective	Effective	Effective	Effective
					Efficient	Efficient	Efficient	Efficient	Efficient	Efficient
					Engaging	Engaging	Engaging	Engaging	Engaging	Engaging
					Error Tolerant	Error Tolerant	Error Tolerant	Error Tolerant	Error Tolerant	Error Tolerant
					Easy to Learn (deeper learning)	Easy to Learn (deeper learning)	Easy to Learn (deeper learning)	Easy to Learn (deeper learning)	Easy to Learn (deeper learning)	Easy to Learn (deeper learning)
					Adaptivity (deeper learning)	Adaptivity (deeper learning)	Adaptivity (deeper learning)	Adaptivity (deeper learning)	Adaptivity (deeper learning)	Adaptivity (deeper learning)
							it could be dangerous	It could be dangerous	it is expensive to obtain objective logging data, it could be dangerous	it is expensive to obtain objective logging data; it could be dangerous
Workload and safety evaluation (first impact)	**Not suitable**	**Not suitable**	**Not suitable** as no primary task exists	**Not suitable** as no primary task exists	**Preferable**	**Suitable/ Preferable**	**Preferable**	**Suitable**	**Suitable**	**Suitable/Less suitable**
					Both subjective and objective parameters	Both subjective and objective parameters	Both subjective and objective parameters	Both subjective and objective parameters	Good for subjective parameters and visual behaviour.	Good for subjective parameters and visual behaviour.
						It depends on the simulator scalability	it could be dangerous	It could be dangerous	Difficult for objective parameters	Difficult for objective parameters. It could be dangerous
Workload and safety evaluation (long term)	**Not suitable**	**Not suitable**	**Not suitable** as no primary task exists	**Not suitable** as no primary task exists	**Less Suitable**	**Less Suitable**	**Preferable**	**Suitable/ Preferable**	**Suitable**	**Suitable**
					Both subjective and objective parameters	Both subjective and objective parameters	Both subjective and objective parameters	Both subjective and objective parameters	Good for subjective parameters and visual behaviour.	Good for subjective parameters and visual behaviour.
					It depends on the simulator scalability	It depends on the simulator scalability	It could be dangerous	It could be dangerous	Difficult for objective parameters	Difficult for objective parameters. It could be dangerous
Acceptability evaluation	**Not suitable**		**Not suitable**	**Not suitable**	**Less suitable**	**Less suitable**	**Preferable**	**Preferable**	**Preferable**	**Preferable**
					It depends on the scalability of simulator	It depends on the scalability of simulator				

use the following table organized for level of system development vs. aim of evaluation

Sp 2.1.4 The Cookbook: its characteristics

1. it is ***GENERAL*** as it can be applied to different systems, with different level of development, both IVI and ADA;
2. it is a **phase of the UCD**;
3. it is focused on **EVALUATION phase** and not on the initial requirement phase**;**
4. it is an evaluation that requires **END-USERS** involvement.;
5. it allows to evaluate **USABILIY, ACCEPTABILITY, WORKLOAD and SAFETY** dimensions;
6. it provides a **number of subjective and objective measurements** to evaluate the impact of these dimensions;
7. it provides a method to collect data that can be used by technicians and designers to **improve prototypes and/ or products**;
8. It provides **a modular questionnaire** able to evaluate different perceived aspects listed in point 5.

CONCLUSIONS

- The Global Evaluation Methodology allowed to developed a number of steps necessary to be made for setting up, conducting and analysing an experiment on HMI evaluation.
- This test protocol can decrease the costs of the final validations of this kind of systems since the time to market can be reduced. Moreover, this protocol could help to focus the attention in the important aspects when testing the prototypes with real drivers
- This method was applied finally in the final exploitation of the three AIDE demonstrators.
- Further improvements of cookbook could come from three AIDE demonstrators studies