Evaluation Methodology

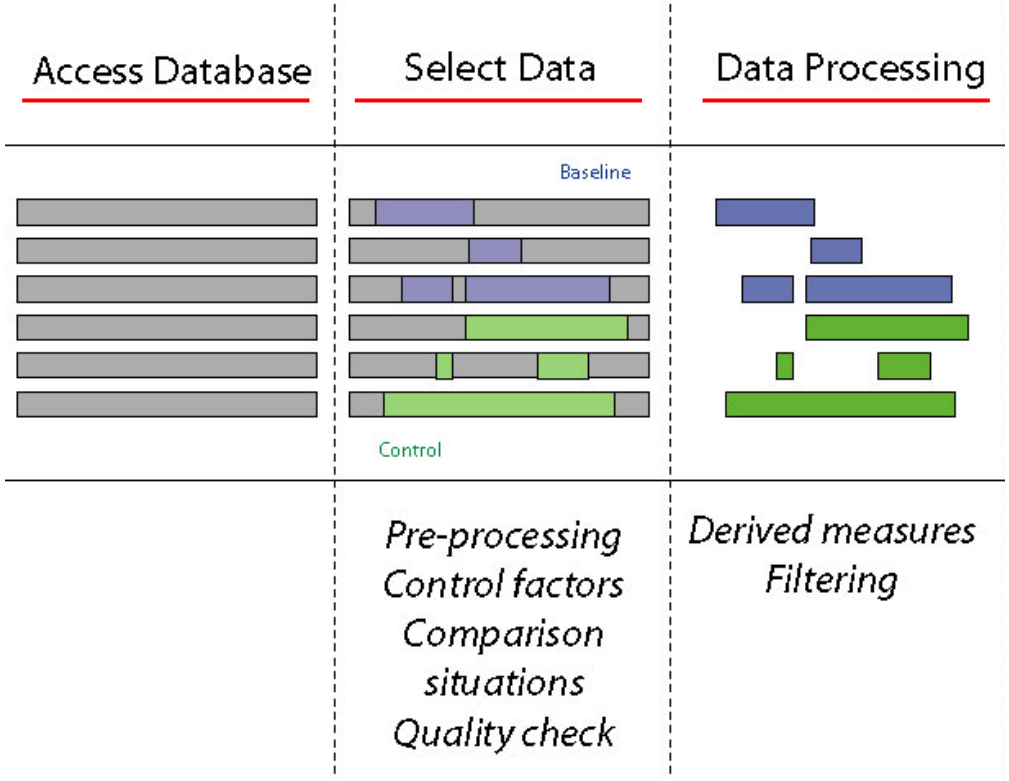
Methodology for Hypothesis Testing and Questionnaire Analysis

Samantha Jamson
University of Leeds
Marco Dozza
Chalmers University of Technology

Final Event
26-27 June 2012
Autoworld, Brussels

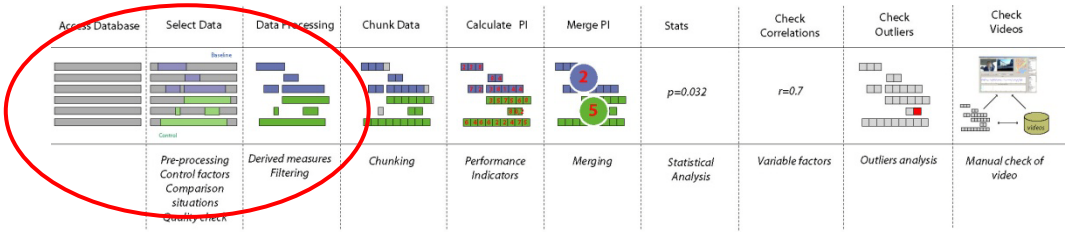
www.eurofot-ip.eu

Data selection and pre-processing

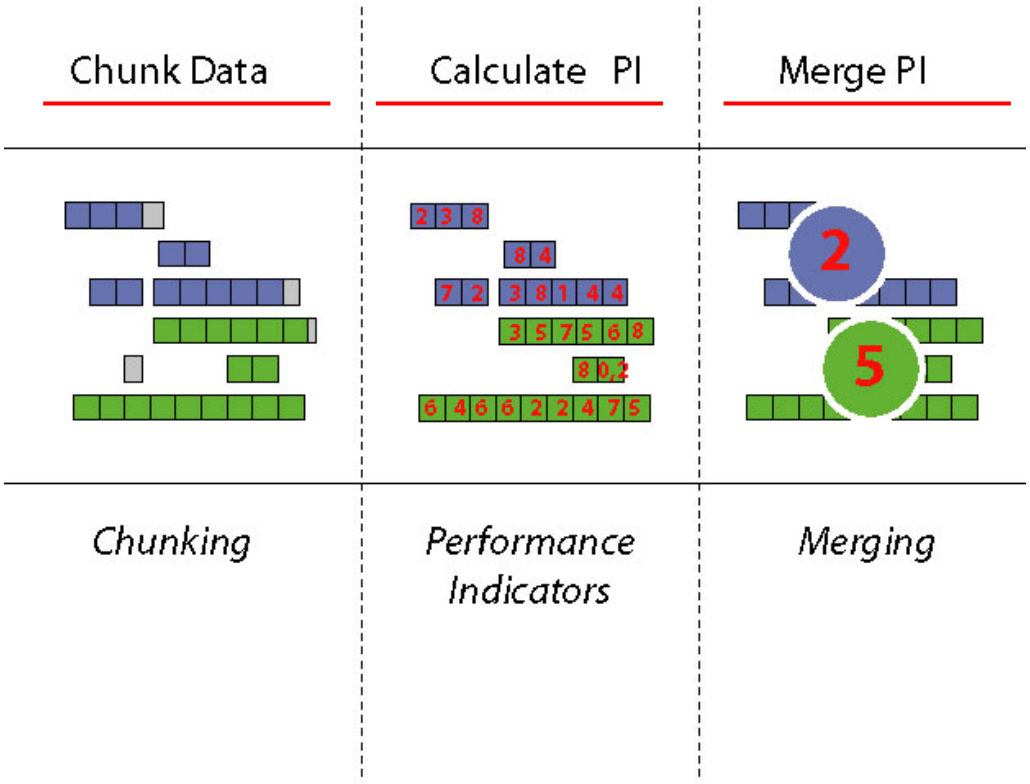


SELECT * FROM ….

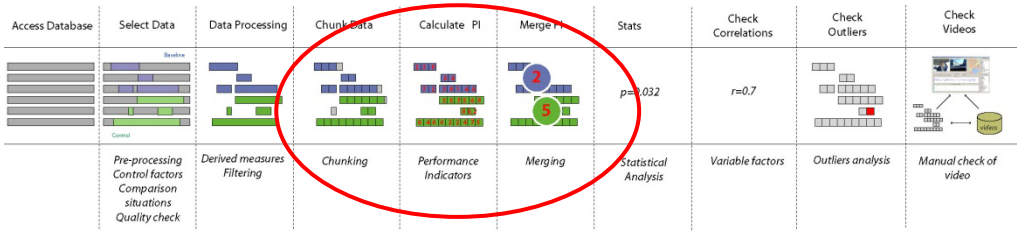
$$k = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2} - 3$$



Performance indicators calculation

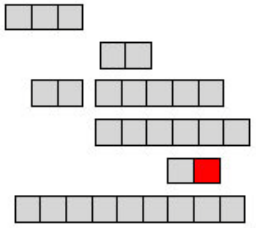


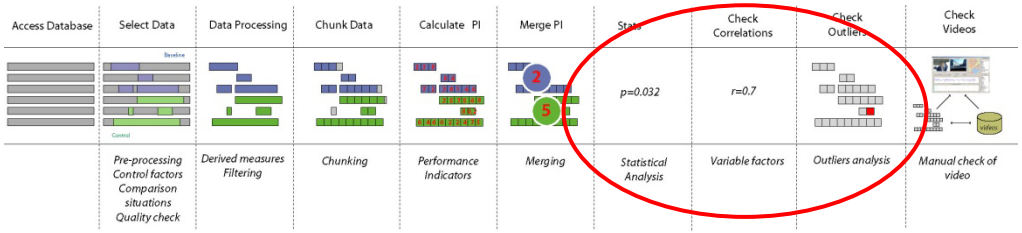
Dozza et al. "**Chunking: a procedure to improve naturalistic data analysis**" Accident Analysis & Prevention, 2012.



$$k = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2} - 3$$

Statistics and verification

Stats	Check Correlations	Check Outliers
p=0.032	*r=0.7*	
Statistical Analysis	*Variable factors*	*Outliers analysis*



Methodology for the questionnaire analysis

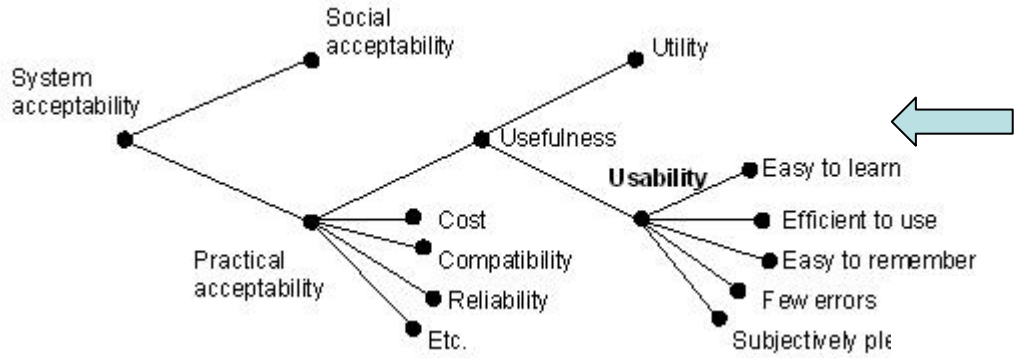
Aims

i. To develop a set of questionnaires to characterise and measure drivers ***acceptance*** of and ***trust*** in the in-vehicle systems under evaluation and the ***workload*** experienced when driving with these systems.

ii. Harmonisation of the questionnaires between the VMCs was envisaged

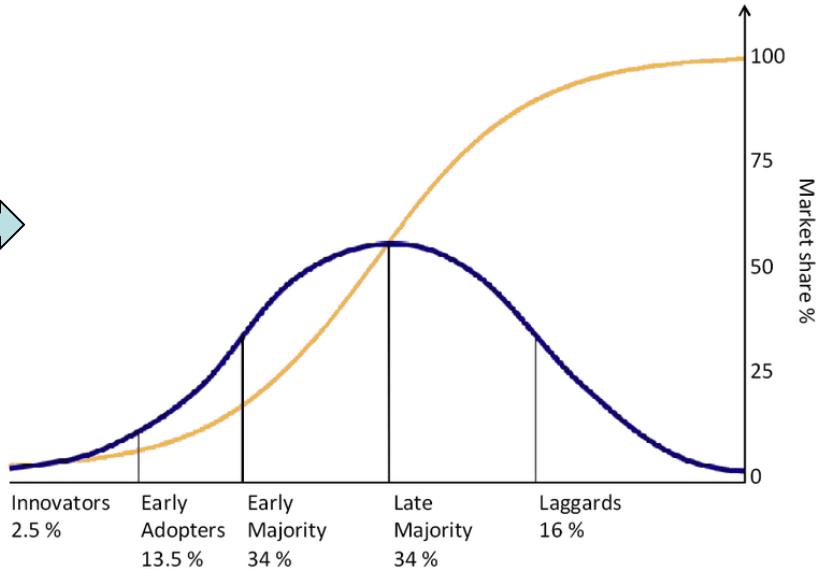
iii. The common questions allowed and overall evaluation and a comparison of results between the VMC related tests.

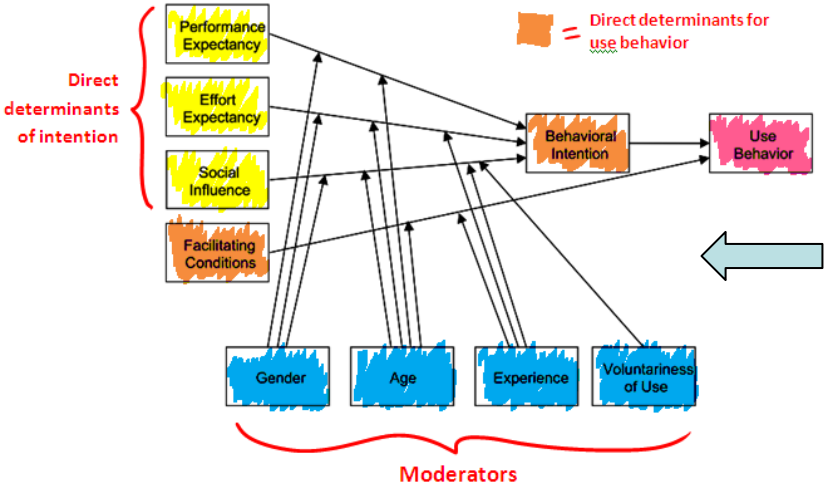
Acceptability/acceptance



Neilsen's (1993) framework of acceptability is rooted in usability engineering – "can an individual use the system?"

Whereas Rogers (1995) models patterns of adoption "who will use the system?"





UTAUT tries to synthesise various models of acceptance

www.eurofot-ip.eu

Trust

Trust is often used interchangeably with acceptance, but users:

- sometimes use automation that they are suspicious of (perhaps because they do not have the time or workload capacity to do otherwise) and
- sometimes do not use automation they believe is competent (perhaps because they enjoy doing the job themselves).
- Trust changes over time, increasing as users experience a reliable system, diminishing when they experience system failures

Questionnaire items

1. Workload
2. Usefulness
3. Satisfaction
4. Social acceptability
5. Affordability
6. Trust
7. Effectiveness
8. Ease of use
9. Misuse/abuse
10. Social influence
11. Behavioural intention
12. Exp. with vehicle technology
13. Exp. with other technology
14. Attitude to target behaviours
15. Demographic data
16. Personality
17. Travel patterns
18. Driving behaviour
19. User practice

Questionnaire procedure

Screening questionnaire:

- focused on demographics, personal data, travel patterns

Questionnaire 1:

- administered prior to system exposure
- measured driver attitudes, accident records

Questionnaire 2:

- determined pre-trial acceptability, workload and behaviour

Questionnaire 3:

- acceptability and workload measures
- administered during the Treatment phase and at repeated time points

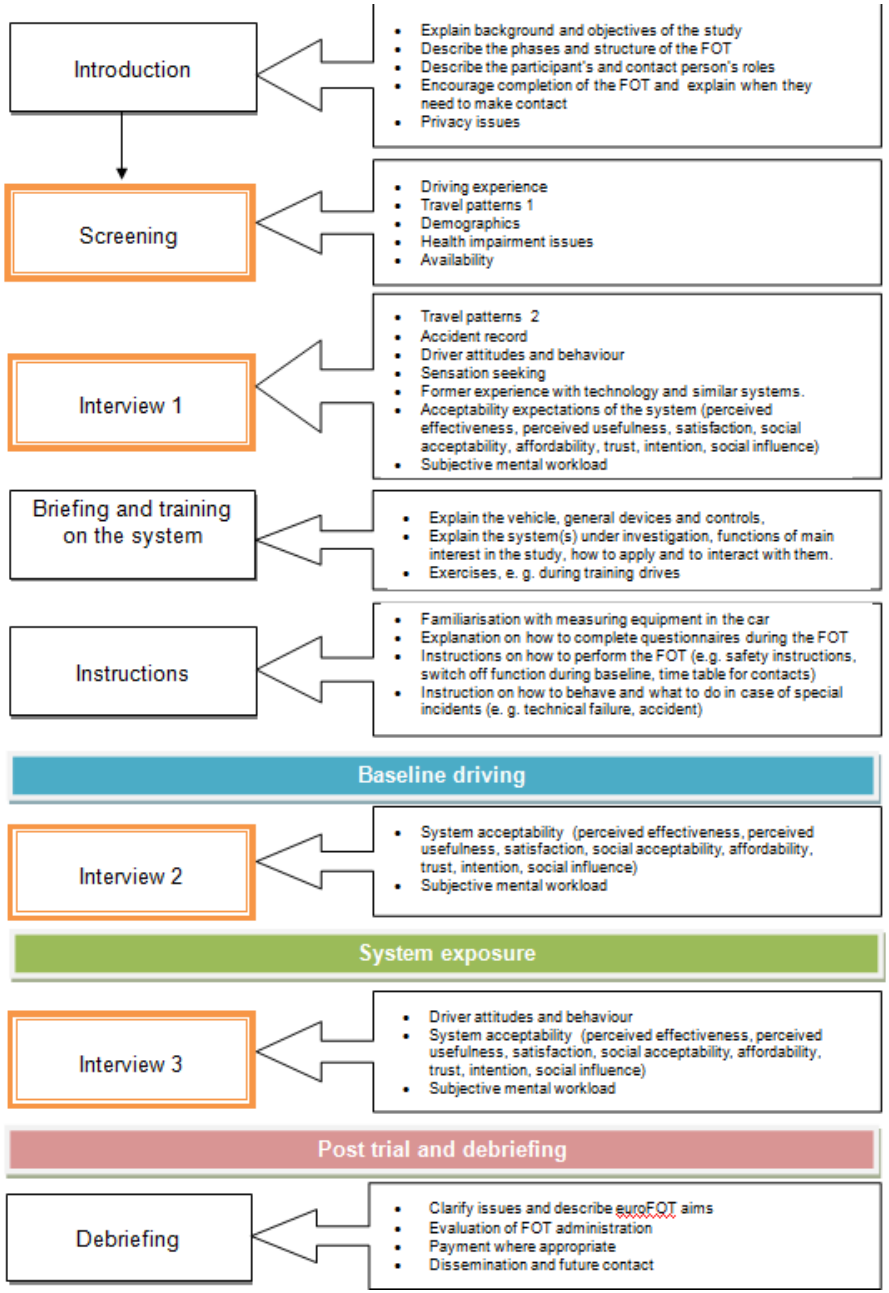
Questionnaire 4:

- acceptability and workload in more depth

Debriefing questionnaire :

- reflected on safety-critical scenarios

Harmonised questionnaires



All test sites used the same core questionnaires (same scales, items etc) with flexibility for additional items

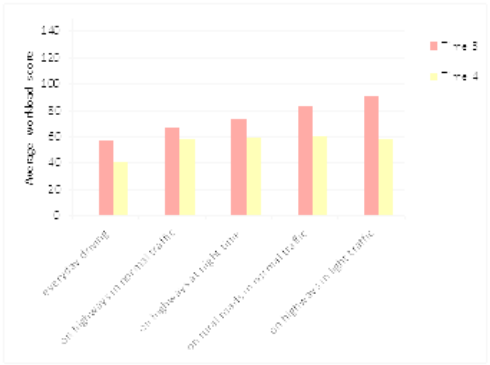
Analysis procedure

i. A template was developed allowing VMCs to “drop” their data straight in (a separate worksheet for each hypothesis)

ii. It automatically recoded and reversed scores where necessary

iii. Standardised graphs were then populated automatically when the data was fed in

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	participant ID												
2		Time 3 - mental workload						Time 4 - mental workload					
3		everyday driv ays in nornways at nicads in nor ways in light traffic						everyday driv ays in nornways at nicads in nor ways in light traffic					
4		xxx_mw_1t	xxx_mw_1c	xxx_mw_1c	xxx_mw_1c	xxx_mw_1f_3		xxx_mw_1t	xxx_mw_1c	xxx_mw_1c	xxx_mw_1c	xxx_mw_1f_4	
5	1	22	32	42	52	58		17	75	71	81	91	
6	2	35	45	55	65	71		30	10	22	32	42	
7	3	33	43	53	63	69		28	75	141	85	78	
8	4	66	76	86	96	102		61	10	55	65	75	
9	5	55	65	75	85	91		50	86	85	95	14	
10	6	17	27	37	47	53		12	53	65	75	85	
11	7	99	109	119	129	135		41	55	44	54	53	
12	8	97	107	117	127	133		41	77	71	81	55	
13	9	95	105	115	125	131		25	88	22	32	77	
14	10	31	41	51	61	67		26	99	14	76	88	
15	11	34	44	54	64	70		29	71	81	91	99	
16	12	61	71	81	91	97		56	22	32	42	4	
17	13	12	22	32	42	48		7	14	78	45	85	
18	14	131	141	51	64	89		12	55	65	75	12	
19	15	45	55	65	75	81		40	85	95	11	85	
20	16	75	85	95	105	111		70	65	75	85	14	
21	17	55	65	75	85	91		50	22	22	32	96	
22	18	88	98	108	118	124		83	14	12	45	85	
23	19	55	65	75	85	91		50	55	55	65	47	
24	20	88	98	108	118	124		24	85	54	64	12	
25	21	66	76	86	96	102		61	71	81	91	12	
26	22	33	43	53	63	69		28	22	32	42	12	
27	23	35	45	55	65	71		30	14	89	78	61	
28	24	57	67	77	87	93		52	55	65	75	64	
29	25	68	78	88	98	104		63	85	95	10	91	
30	26	79	89	99	109	115		74	65	75	85	65	
31	27	64	74	84	94	100		59	98	10	11	95	
32	28	45	55	65	75	81		40	65	75	85	75	
33	29	14	24	34	44	50		9	98	10	11	10	
34	30	55	65	75	85	91		50	76	86	96	56	
35													
36		everyday driv ays in nornways at nicads in nor ways in light traffic											
37	Time 3	57.00	67.00	73.67	83.77	90.40							
38	Time 4	40.60	58.83	59.23	60.50	57.93							
39													
40													
41													
42													
43													
44													
45													
46													
47													
48													
49													
50													
51													
52													
53													
54													
55													
56													
57													
58													
59													
60													
61													
62		Conduct t -tests between Time 3 and Time 4											
63		Also check there are no differences in scores in the first item in the workload section (											



8 Functionalities, 28 Partners, 1000 Vehicles

1 Field Operational Test, **8 Functionalities**

28 Partners, 1000 Vehicles, 1 Field Operational Test

8 Functionalities, **28 Partners**, 1000 Vehicles

1 Field Operational Test, 8 Functionalities

28 Partners, **1000 Vehicles**, 1 Field Operational Test

8 Functionalities, 28 Partners, 1000 Vehicles

1 Field Operational Test