Pergamon

Transpn Res.-C, Vol. 5, No. 1, pp. 1–10, 1997

0968-090X/97 $17.00 + 0.00

PII: S0968-090X(96)00025-3

A SIMPLE PROCEDURE FOR THE ASSESSMENT OF ACCEPTANCE OF ADVANCED TRANSPORT TELEMATICS

JINKE D. VAN DER LAAN, ADRIAAN HEINO and DICK DE WAARD*
Traffic Research Centre, University of Groningen, P.O. Box 69, 9750 AB, Haren, The Netherlands

(Received 1 July 1996; in revised form 2 December 1996)

Abstract—There is no standard way of measuring driver acceptance of new technology. A review of the literature shows that there are almost as many methods of assessment of acceptance as there are acceptance studies. The tool for studying acceptance of new technological equipment that is presented here has a major advantage compared with many other studies in that esoteric knowledge of scaling techniques is not required. The technique is simple and consists of nine 5-point rating-scale items. These items load on two scales, a scale denoting the usefulness of the system, and a scale designating satisfaction. The technique has been applied in six different studies in different test environments and analyses performed over these studies show that it is a reliable instrument for the assessment of acceptance of new technology. The technique was sensitive to differences in opinion to specific aspects of in-vehicle systems, as well as to differences in opinion between driver groups. In a concluding section explicit recommendations for use of the scale are given. © 1997 Elsevier Science Ltd

INTRODUCTION

A variety of in-vehicle telematics systems is already available, while many new systems are being designed. The range of these systems varies from those that provide the driver with information, for example about congestion, to more complex systems that may even take over driver tasks in hazardous situations (Hancock and Parasuraman, 1992). Whereas excellent system performance may be sufficient for the technician, it is as important that the equipment is appealing for and accepted by the driver. For tutoring and advisory systems the main issues determining the feasibility are not of a technical nature, but concern the social context for introduction (Rothengatter *et al.*, 1991). A prerequisite for the introduction of new in-vehicle technology is acceptance by the public. It is unproductive to invest effort in designing and building an intelligent co-driver if the system is never switched on, or even disabled.

In-vehicle systems and acceptance

The different in-car systems can be positioned along a continuum from relatively simple systems that provide the driver with information only, to relatively complex systems that take over parts of the driving task, or force the driver to perform the demanded behaviour (see Fig. 1). An example of a purely informative system is RDS (Radio Data System; Chevreuil *et al.*, 1991), in which messages with respect to congestion, road works and weather information are conveyed to the driver. The driver is free to use the information or to neglect it. At the other end of the continuum interventionist devices can be positioned, such as policing systems or an automatic braking system that is triggered in case of very short headways. Most systems, however, will be positioned somewhere in between these extremes. Examples are feedback and support systems such as warning systems, monitoring and co-operative driving systems. These systems give information that may not necessarily be heeded by the driver, though the information conveyed is more compelling than RDS messages. Systems that restrict the driver's behaviour or policing systems that force a behavioural change are likely to be less accepted than nonrestricting, informative systems that are situated more towards the left-hand side of the continuum.

The difference between the two is not always clear, but in general user-acceptance is more directed towards evaluation of the ergonomics of the system, while social acceptance is a more indirect evaluation of (longer-term) consequences of the system. For example, Bonsall *et al.* (1991)

*Author for correspondence.

Provision of information	Provision of feedback & support	Intervention/ force
Free choice of behaviour		Completely restricted behaviour

Fig. 1. Above the line: a continuum of advanced transport telematics. Below the line: freedom of choice in behaviour.

used a questionnaire as well as individual and group interviews to measure user-acceptance. They mainly focused on the information such a system provided. Rothengatter *et al.* (1991) on the other hand, focused on social acceptance, and they measured indirect attitudes towards a system and used an extensive questionnaire to accomplish this.

Even if a study is aimed at the determination of the level of user-acceptance, the focus differs. Frequently the system as a whole is evaluated. Systems are also evaluated on their pleasantness and usefulness (Kuiken and Groeger, 1993), the comfort and benefit (Becker *et al.*, 1995) or the ease or degree of use (Crosby *et al.*, 1993). Highlighted aspects in the evaluation of the provided information are the content or the format, the reliability, the relevance (Schofer *et al.*, 1993), the accuracy (Crosby *et al.*, 1993; Schofer *et al.*, 1993) or the effectiveness of the information (Michon and McLoughlin, 1991). Acceptance sometimes includes the *intention to purchase* the system (e.g. Becker *et al.*, 1994, 1995) and assessment of the price people are willing to pay (Barham *et al.*, 1993; Becker *et al.*, 1995).

In conclusion, there seem to be as many questionnaires and methods to measure acceptance as there are system-evaluation studies. This is not a desirable situation. There is a strong need for a simple, standard tool for the assessment of acceptance that can be used by the majority of researchers and that allows a comparison of impact of new devices with other systems. In the present paper a simple inexpensive technique is presented that has been applied in different studies (reported in this paper), and can be applied by most people active in the field of advanced telematics.

A simple scale

This paper reports on a simple scale for which acceptance of a system has been operationalized as direct attitudes towards that system. Attitudes are here defined as predispositions to respond, or tendencies in terms of 'approach/avoidance' or 'favourable/unfavourable'. Osgood *et al.* (1957) appoint attitudes to a basic bipolar continuum with a neutral or zero reference point, implying that they have both direction and intensity. The type of evaluation, i.e. the terms on which a concept is evaluated, depends on the concept that has to be judged. Few dimensions, each more than one item that can form a scale, are to be preferred, because an important drawback of single-item assessments of attitude is a potentially weak reliability (O'Keefe, 1990). A set of nine scales that could be applied to telematics systems was selected from Osgood *et al.* (1957), who had repeatedly shown that these scales were the most useful for further factor analyses in the assessment of dimensions of opinion.

The present paper gives an overview of the technique itself, and experience with the scale in different studies. In addition, analyses over studies were performed to assess the technique's sensitivity.

METHOD

Results from application of the technique in six studies are used as input for statistical analysis. The studies are briefly summarized below. The questionnaire itself consists of nine items, and is shown in Table 1. The instruction given to subjects was: "Could you please indicate below what your opinion was about the equipment that (...) "

Individual item scores run from -2 to $+2$. Item numbers 3, 6 and 8 are mirrored, compared to the other items.

Studies

The studies in which the questionnaire was applied are listed in Table 2. Also indicated are the different groups that were included in the individual experiments. In all studies the different

Table 1. The nine items in the questionnaire

My judgements of the (...) system are... (please tick a box on every line)			
1	useful	_ _ _ _ _	useless
2	pleasant	_ _ _ _ _	unpleasant
3	bad	_ _ _ _ _	good
4	nice	_ _ _ _ _	annoying
5	effective	_ _ _ _ _	superfluous
6	irritating	_ _ _ _ _	likeable
7	assisting	_ _ _ _ _	worthless
8	undesirable	_ _ _ _ _	desirable
9	raising alertness	_ _ _ _ _	sleep-inducing

Table 2. Studies in which the questionnaire was used

Experiment	Reference	System evaluated	Specific Groups	Environment	*N* (before)	*N* (after)
a. Tutor-1	De Waard *et al.* (1994)	Tutoring and enforcement system	Young and elderly	Simulator	37	29
b. Tutor-2	De Waard and Brookhuis (1997)	Tutoring and enforcement system	Experimental and control	On-the-road	29	29
c. ICC-1	Rothengatter and Heino (1994)	Intelligent Cruise Control	Experimental (3) and control	Simulator	80	80
d. ICC-2	Hogema *et al.* (1994)	Intelligent Cruise Control	Experimental (2) and control	Simulator	—	54
e. CAS-1	Janssen *et al.* (1993)	Collision Avoidance System	Experimental (2)	Simulator	—	52
f. CAS-2	Janssen *et al.* (1993)	Collision Avoidance System	Experimental (2)	On-the-road	—	39

Under specific groups between-subjects factors are indicated with the number of conditions in brackets. Under '*N* (before)' the number of subjects that completed the before-trial questionnaire is indicated, while under '*N* (after)' the number of subjects that completed the system-evaluation questionnaire after the experimental trials is indicated.

systems were judged after subjects had been exposed to them. In the tutoring and ICC-1 studies subjects were also asked to appreciate a hypothetical system explained to them before they had any experience with it. The Tutor-1, ICC-1 and CAS-1 studies were performed in the high-fidelity driving simulator of the Traffic Research Centre (TRC), while the ICC-2 study was realized in the high-fidelity driving simulator of TNO Human Factors Research Institute. Two on-the-road studies were included, one by TNO Human Factors Research Institute (CAS-2) and one by the Traffic Research Centre (Tutor-2). A short outline of the studies is provided below.

Tutoring and enforcement system (Tutor-1 and Tutor-2). A system that monitored the drivers' law-abiding behaviour has been tested in two environments; in a driving simulator (De Waard *et al.*, 1994) and in a real car (De Waard and Brookhuis, 1997). In both cases, detected traffic-law violations such as speeding and not coming to a complete stop at a stop sign, led to warning messages of the tutoring system. In the simulator study (Tutor-1) two groups of drivers were included, relatively young drivers (between 30 and 45 yr of age) and elderly drivers (aged between 60 and 75). Drivers received feedback about detected violations both in the auditory (e.g. a voice digitized message "You drove through red") and visual modality (the same messages projected as text on the simulator screen). In the on-the-road study (Tutor-2) an experimental and a control group of young drivers were included. The control group did not receive tutoring messages if violations were detected, opposite to the experimental group who received auditory feedback when they violated traffic-rules. Drivers did not receive any penalty in case of violations in either study. The tutoring and enforcement systems can be situated in the right-hand section of the continuum displayed in Fig. 1.

Intelligent cruise control -1 (ICC-1). An Intelligent Cruise Control (ICC) not only takes care that a car maintains a pre-set speed, it also detects and responds to slower lead vehicles. In the TRC driving simulator, the behavioural effects of three ICCs were tested in a within-subject design (Rothengatter and Heino, 1994). Each subject performed two trials (i.e. the first trials served as subject's own control). For the second trial, the subjects were randomly assigned to one of four experimental conditions. Subjects in the ICC condition received information about the distance to a car in front by means of symbols, displayed on a screen mounted on the dashboard ('warning signal'). Subjects in the active gas pedal condition received a counter force on the gas pedal

(Godthelp and Schumann, 1991), a force that depended upon the distance to a car-in-front ('haptic feedback'). In the AICC (Autonomous Intelligent Cruise Control) condition, speed and distance to a car-in-front were controlled autonomously, subjects did not have to use the gas and brake pedal at all. Finally, subjects in the control condition did not receive any information about speed and distance to cars-in-front. Regarding the continuum in Fig. 1, the ICC can be assumed to be situated in the centre, the AICC at the right end, and the active gas pedal somewhere in between these two.

Intelligent cruise control -2 (ICC-2). Two different types of ICCs were tested in a between-subjects design in the TNO driving simulator (Hogema *et al.*, 1994). The two types of systems made use of roadside information (i.e., made use of local speed limits in settling the car's speed). A control condition had no such information. One system only notified ('warning signal'). The other also intervened and adapted the vehicle's speed ('warning signal plus intervention on speed'). Feedback to the driver was given by means of a flashing LED on the speedometer, indicating the local limit. Additional notifying feedback was given visually, aurally or haptically (a short vibration on the gas pedal). For the present report, these sub-conditions were taken together because these warning signals only differed in modality. The warning-signal-only system can be situated at the centre of the continuum (Fig. 1), whereas the warning plus speed-intervention system should be set to the right of the centre.

Collision avoidance system (CAS-1 and CAS-2). Two actions of a collision avoidance system (CAS) were evaluated in this study (Janssen *et al.*, 1993). The system actions that were tested were: an auditory alarm and an accelerator pedal counter force ('haptic feedback'). The sound feedback was the emission of a tone at the moment headway distance was considered unsafe according to a specific criterion. In the haptic feedback condition, the gas pedal counter force (0–100 N) was activated linearly with increasing counter force as headway distance decreased. The system was tested both in a driving simulator (CAS-1) and in a real vehicle (CAS-2). On the continuum (Fig. 1), gas pedal feedback is probably more intervening for the driver than auditory feedback.

Simultaneous components analysis (SCA)

To summarize the nine original items from Table 1 in a smaller number of components, a principal component analysis (PCA) was performed. By a rotation of these principal components, new components can be found that account for the same amount of variance as the original principal components. For example, the Kaiser's varimax applied to the component weights aims at finding either large or small component weights that define components more clearly in terms of subsets of variables. In the multiple group method (see Nunnaly, 1978) this idea of mainly large and small component weights is carried to the extreme in that all component weights are either 0, 1 or -1. As a result, the components are computed as sums of variables that belong to particular subsets only. Because these 'multiple group' components cannot be seen as rotations of the principal components, they usually do not explain the maximum possible amount of variance. However, if they account for a great part of the variance, the set of components can be interpreted in terms of subscales.

In ordinary PCA, components are determined for sets of variables that have been observed in a single sample. Here, however, the same items have been used in six samples (Table 2). Analysing these observations by PCA's for each of these samples separately will generally yield different results (Kiers and Ten Berge, 1989; Kiers, 1990). In order to find components across samples that have the same definition, components can be constructed that are based on the same set of weights for the variables in all samples. In this so called 'simultaneous component analysis' (SCA; Millsap and Meredith, 1988) component weights are found that define (identical) components which optimally account for the variance in all samples simultaneously. The amount of variance accounted for by the SCA components in each sample has to be computed to check whether these simultaneous components adequately summarize the information in all samples. As in PCA, transformations of the SCA components can be performed, including the multiple group method.

RESULTS

The before-measurement item-scores for the Tutor-1, Tutor-2 and ICC-1 experiments (Table 2) were analysed by simultaneous components analysis with the computer program SCA (Kiers,

Table 3. PCA and SCA of the nine items in Table 1 for the Tutor-1 ($N = 37$), Tutor-2 ($N = 29$) and ICC-1 ($N = 80$) samples

	Weights PCA Tutor-1		Weights PCA Tutor-2		Weights PCA ICC-1		Weights SCA Varimax		Simple weights	
Item	Component 1	Component 2	Component 1	Component 2	Component 1	Component 2	Component 1	Component 2	Component 1	Component 2
1. useful	0.258	0.254	0.144	0.361	0.345	0.169	0.312	0.190	1	0
2. pleasant	0.439	−0.043	0.492	0.004	0.018	0.531	−0.019	0.541	0	1
3. good	0.329	0.098	−0.145	0.497	0.463	−0.072	0.495	−0.083	1	0
4. nice	0.403	−0.059	0.452	0.068	−0.032	0.546	−0.027	0.523	0	1
5. effective	0.084	0.488	0.073	0.416	0.526	−0.198	0.507	−0.084	1	0
6. likeable	0.517	−0.258	0.469	0.039	−0.089	0.529	−0.066	0.522	0	1
7. assisting	0.283	0.187	−0.013	0.479	0.432	0.051	0.409	0.072	1	0
8. desirable	0.328	0.137	0.535	−0.114	0.319	0.197	0.163	0.325	0	1
9. raising alertness	−0.081	0.750	0.012	0.448	0.304	0.163	0.448	0.001	1	0
Variance accounted for										
PCA	80.7		64.8		67.6		71.0			
SCA	80.1		64.3		67.2		70.5			
Simple weights	79.9		64.5		66.3		70.2			

1990). A varimax rotation was performed in order to define components in terms of subsets of variables. For these analyses the Tutor-1, Tutor-2 and ICC-1 studies were selected because only in these experiments item-scores were taken before individuals had any experience with the systems under evaluation.

The explained percentages of variance as listed in Table 3 indicate that the separate PCA's were hardly any better than the simultaneous component analysis. This implies that in each data set, the single set of weights defined components that were almost as good as the best possible (PCA solutions).

Based on the (transformed) SCA weights, the simple weights in the last two columns of Table 3 were used to construct two subscales. The results show (Table 3) that the components (subscales) based on these simple weights accounted for 70.2% of variance, which is only 0.3% less than the SCA weights. If the simple weights are used, the first subscale (or 'component') contains evaluations in terms of useful, good, effective, assisting and raising alertness, and could be interpreted as denoting the usefulness of a system. The second subscale contains evaluations in terms of pleasant, nice, likeable and desirable, and could be interpreted as reflecting satisfaction associated with a system. In Table 4 the reliability coefficients (Cronbach's alpha) for the usefulness and statisfaction subscales are given for each trial of each of the six experiments. As can be seen in Table 4, the homogeneity values are high for all experimental trials. This means that scale reliability is high, items loading on the same factor are in all experiments highly correlated with each other.

Since a component (or subscale) is a new variable that is defined as the weighted sum of the original variables, scores on each component can be computed for each subject separately. Here, 'usefulness scores' were computed as the average of the items 1, 3, 5, 7 and 9, whereas 'satisfying scores' were computed as the average of the items 2, 4, 6 and 8. In Table 5, the usefulness and satisfying scores are given for each trial of each of the six experiments.

Analyses of variance were performed on the usefulness and satisfying scores to test whether there were effects on these scales of the various in-vehicle systems. A difference was made with respect to comparisons between subjects (Table 6) and comparisons within subjects (Table 7; repeated measurement analyses of variance).

In the after-measurements of the Tutor-1 and ICC-1 experiments, statistically significant effects of group (between-subjects) were found (Table 6). In the Tutor-1 experiment, the elderly, compared to the younger drivers, evaluated the tutoring and enforcement system as more acceptable, i.e. both more useful and more satisfying (see Table 5). In the ICC-1 experiment, all systems (including the control condition) were evaluated positively in terms of usefulness, with the AICC system receiving the lowest usefulness score. The haptic feedback condition has the lowest score on evaluated satisfaction (see Table 5).

The statistical results of the studies that included a before- and after-measurement are shown in Table 7. The 'trial' and 'trial by group' effects of the Tutor-1 and ICC-1 experiments were found to be significant. In the group elderly (Tutor-1), experience with the tutoring and enforcement system caused a statistically significant increase in both the usefulness score (t-value$=2.75$, df$=8$, $p=0.03$) and the satisfying score (t-value$=1.96$, df$=8$, $p=0.09$). For the younger drivers, however, experience with the system resulted in a significant *decrease* in the satisfying score (t-value $=2.55$, df$=18$, $p=0.02$) whereas the usefulness score did not change significantly (t-value $=1.00$, df$=18$, $p=0.33$).

Table 4. Reliability coefficients (Cronbach's α) for the usefulness and satisfying subscales for each trial of each of the six experiments

Study	Before measurement		After measurement	
	Usefulness	Satisfying	Usefulness	Satisfying
Tutor-1	0.91	0.94	0.87	0.85
Tutor-2	0.76	0.90	0.81	0.90
ICC-1	0.85	0.83	0.80	0.85
ICC-2			0.73	0.88
CAS-1			0.82	0.81
CAS-2			0.85	0.84

Table 5. Usefulness and satisfying scores for the different experimental and control conditions in the experiments listed in Table 2

Study	Specific groups	Symbol	Before measurement		After measurement	
			Usefulness	Satisfying	Usefulness	Satisfying
Tutor-1	Tutor (young)	□	0.88	0.24	0.66	−0.38
	Tutor (elderly)	■	0.56	−0.22	1.73	0.83
Tutor-2	Tutor (young)	□	1.30	0.30	1.43	−0.06
	Control	▼	1.34	0.27	1.19	0.31
ICC-1	Haptic feedback	◆	1.11	0.25	0.89	−0.35
	Warning signal	○	1.21	0.31	1.00	0.45
	AICC	▽	1.06	0.10	0.34	−0.18
	Control	▼	0.85	0.10	0.92	0.30
ICC-2	Warning signal	○			0.72	0.88
	Signal + intervention	●			0.68	0.45
	Control	▼			0.90	1.08
CAS-1	Haptic feedback	◆			0.57	−0.14
	Auditory feedback	◇			0.83	−0.10
CAS-2	Haptic feedback	◆			0.39	−0.40
	Auditory feedback	◇			0.45	−0.66

Symbols correspond with the symbols that are used in Fig. 2 and denote the type of in-vehicle system ▼: control (no system); ▽: AICC; □ and ■: tutoring and enforcement system; ○: warning signal; ●: warning signal plus intervention; ◆: haptic feedback; ◇: auditory feedback.

In the ICC-1 experiment, both the control condition and experience with the warning signal had no significant effects on the usefulness scores (t-value$_{control}$ = −0.46, df = 19, p = 0.65; t-value$_{warning}$ = 1.27, df = 19, p = 0.22) and satisfying scores (t-value$_{control}$ = −1.21, df = 19, p = 0.24; t-value$_{warning}$ = −0.72, df = 19, p = 0.48). Experience with the haptic feedback resulted in a significant decrease in the satisfying score (t-value = 3.32, df = 19, p = 0.00) and no effect on the usefulness score (t-value = 1.56, df = 19, p = 0.14), whereas the AICC resulted in a decrease in the usefulness score (t-value = 4.66, df = 19, p = 0.00) without an effect on the satisfying score (t-value = 1.08, df = 19, p = 0.30).

Based on the usefulness scores and satisfying scores in Table 5, and the characteristics of the different in-vehicle systems evaluated, the after-measurement scores were averaged across individual experiments, which resulted in eight (pairs of) scores. These scores are displayed in Fig. 2 in a two-dimensional space.

All systems depicted in Fig. 2 are in the right-hand section or in the mid of the continuum as displayed in Fig. 1. All systems are considered neutral to positive on the usefulness dimension. Whether a system is satisfying, however, seems to depend upon the degree to which a person is forced to act in a specific way. The more a system restrains free choice, the more unsatisfying the system was evaluated. Tutoring and AICC systems are situated in the left-hand section, and to the left of the control condition, in Fig. 2. The odd one out is the satisfying score on the tutoring

Table 6. Multivariate analyses of variance of average usefulness and satisfying scores: between-subjects effects (all six experiments)

Between-subjects effects (after-measurements)	F-value	df	p-value
a. Tutor-1			
Group (young vs elderly)	8.98	2,25	0.00
usefulness	15.67	1,26	0.00
satisfying	12.39	1,26	0.00
b. Tutor-2			
Group (experimental vs control)	2.31	2,26	0.12
c. ICC-1			
Group (experimental [3] vs control)	3.95	6,152	0.00
usefulness	3.41	3,76	0.02
satisfying	4.37	3,76	0.01
d. ICC-2			
Group (experimental [2] and control)	1.18	4,102	0.32
e. CAS-1			
Group (auditory vs haptic feedback)	0.77	2,49	0.47
f. CAS-2			
Group (auditory vs haptic feedback)	0.76	2,36	0.47

Table 7. Multivariate analyses of variance of average usefulness and satisfying scores: Within-subjects effects (Tutor-1, Tutor-2 and ICC-1)

Within-subjects effects (before vs after-measurements)	*F*-value	df	*p*-value
a. Tutor-1			
Trial	2.62	2,25	0.09
usefulness	4.83	1,26	0.04
satisfying	0.74	1,26	0.40
Group x trial (young vs elderly)	6.11	2,25	0.01
usefulness	10.32	1,26	0.00
satisfying	10.88	1,26	0.00
b. Tutor-2			
Trial	0.50	2,26	0.61
Group x trial (experimental vs control)	2.16	2,26	0.14
c. ICC-1			
Trial	6.19	2,75	0.00
usefulness	12.31	1,76	0.00
satisfying	1.79	1,76	0.19
Group x trial (experimental [3] vs control)	4.12	6,152	0.00
usefulness	4.56	3,76	0.01
satisfying	4.48	3,76	0.02

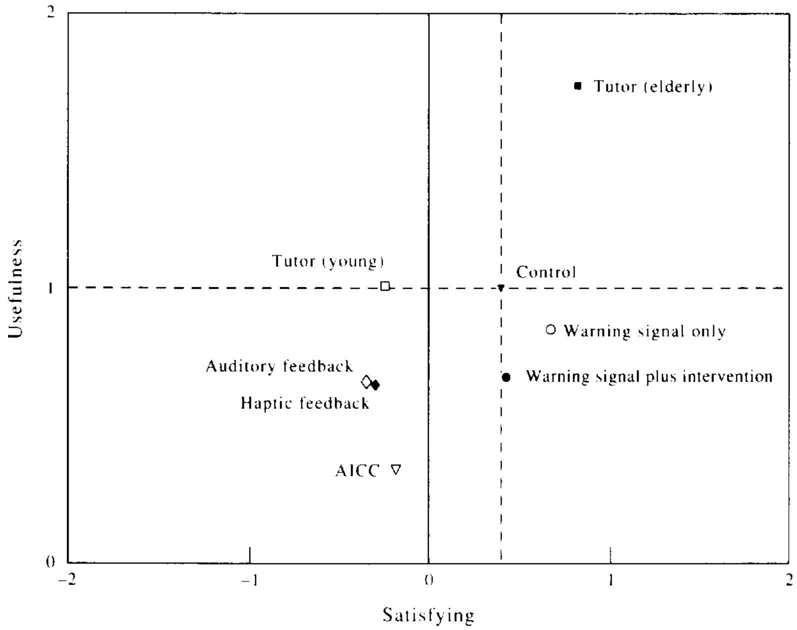

Fig. 2. After-measurement usefulness scores and satisfying scores averaged across individual experiments (see also Table 5). Both dimensions have a range between −2 and +2 (while the usefulness is only displayed between 0 and +2). The evaluation of the control group (no experience with ATT systems) is indicated with dashed lines.

system of elderly drivers. Elderly drivers however, sometimes miss information and were very pleased with the system's warnings that included information about local limits (see De Waard *et al.*, 1994). These drivers made use of the warnings as driver support, which is reflected in the high usefulness score (Fig. 2). The usefulness score of the AICC is low, AICC takes over control, it was not considered as supporting the driver. Relatively, the most satisfying are the warning signals, which only provide information. If the signal is coupled with intervention on longitudinal control, both the satisfying and usefulness scores decrease, though not very dramatically. Relatively, the least satisfying are the auditory and haptic feedback systems that gave information to the driver in a more compelling way than the warning signals did.

CONCLUSIONS

At present, the evaluation of acceptance of new technology is far from being standardized. The evaluation of systems as found in the literature is diverse and directed to different aspects of the system. The procedure presented here is relatively simple, and has been used in different studies including different test-environments. Here, acceptance is measured by direct attitudes towards a system (Osgood *et al.*, 1957) and provides the investigator with a system evaluation in two dimensions. Practical aspects of the system are reflected in the usefulness score, while the pleasantness is mirrored in the satisfying score.

The absolute values on these two scales give an impression of overall judgement (Fig. 2). However, besides looking at the absolute after-measurement scores, three experiments allowed for comparison of before and after-measurement score (Tutor-1, Tutor-2 and ICC-1). In these comparisons, deviations from expectations are emphasized. So, for example in the Tutor-1 experiment, the elderly changed their *a priori* evaluation of a tutoring and enforcement system, both in terms of usefulness and satisfaction, as a result of the actual experience with such a system. As such, the so called within-subject evaluations (before and after-measurement) provide additional information about system acceptance to the between-subject evaluations (after-measurement only).

Positioning of the evaluated systems on the continuum from informative to interventionist as presented in Fig. 1 matches the satisfying scores. As expected, more take-over or the monitoring of behaviour on law-compliance is rated as less satisfying by the (young) driver. Haptic feedback is also rated as less satisfying compared to a warning signal. Purely informative systems were not included in the evaluation, all conditions of the six studies either provide the driver with feedback about his or her behaviour, or go beyond that point and take over control (AICC). The systems that were evaluated can therefore be positioned in the right-hand section of the continuum of Fig. 1. It is likely that informative systems such as RDS will score high on satisfaction, though no supporting evidence for this is gathered yet.

Since the nine items were translated from Dutch, reliability analyses are still required. Based on the rather consistent results found in the six studies described, major differences are not expected, also because the English expressions match the Dutch terminology* very closely. Advice on how to use the scale is given below.

Guide for scale users

1. Describe the system to be evaluated in terms of 'what is your judgement about a system that would...', and present the nine items (before-measurement).
2. Present the nine items again after experience with the system under evaluation with the description: 'what is your judgement about the system... you just finished driving with'.
3. Individual items should be coded from +2 to −2 from left to right, scores on items 3, 6, and 8 should be coded ranging from −2 to +2 (the items are mirrored).
4. Perform reliability analyses on the before-measurement. Use item 1, 3, 5, 7, and 9 for the usefulness scale, and item 2, 4, 6, and 8 for the satisfying scale.
5. If reliability (Cronbach's α) is sufficiently high (above 0.65), compute per subject the end-score for the two scales by averaging the scores on items 1, 3, 5, 7, and 9 for the usefulness score, and averaging scores on items 2, 4, 6, and 8 for the satisfying score.
6. The usefulness scale can now be averaged over subjects to obtain an overall system practical evaluation. The same can be done with the satisfying scores.
7. Compute difference-scores per subject by subtracting the before-measurement score from the after-measurement score per scale. The difference scores show whether and in which direction subjects' opinion was altered as a result of experience with the system.

Steps 1 and 7 can be left out if evaluation of acceptance is limited to after-measurement. Step 4 is still required because the scale was translated from Dutch. In this case in Step 4 the after-measurement scores should be used.

Acknowledgements—We wish to acknowledge Dr Karel Brookhuis for his effort in developing the scale and his help during the preparation of this manuscript. Dr Henk Kiers of the Department Psychology, University of Groningen, has made very useful remarks on SCA methodology, and Dr Wiel Janssen of the TNO Human Factors Research Institute is thanked for being so kind as to provide us with the raw data set of the CAS-2 and ICC-2 studies.

*The Dutch items were: nuttig/zinloos; plezierig/onplezierig; slecht/goed; leuk/vervelend; effectief/onnodig; irritant/aangenaam; behulpzaam/waardeloos; ongewenst/gewenst; waakzaamheidverhogend/slaapverwekkend.

REFERENCES

Barham, P. A. J., Mitchell, K., Oxley, P. R. and Ayala, B. (1993) *In-Vehicle Driving Tests—Research Protocols for EDDIT (Part One). EDDIT, Elderly and Disabled Drivers Information Telematics.* Report V2031/EDDIT/Deliverable 16 (WP 7) to the commission of the European Communities. Cranfield University, Bedford, U.K.

Becker, S., Bork, M., Dorissen, H. T., Geduld, G., Hofmann, O., Naab, K., Nöcker, G., Rieth, P. and Sonntag, J. (1994) Summary of experiences with autonomous intelligent cruise control (AICC). Part 2: results and conclusions. In *Towards an Intelligent Transport System. Proceedings of the First World Congress on Applications of Transport Telematics and Intelligent Vehicle-highway Systems*, ed. ERTICO, pp. 1836–1843. Artech House, Boston.

Becker, S., Brockmann, M., Mertens, A., Nui, R. and Sonntag, A. (1995) User acceptance and willingness to pay for advanced driver support systems. Paper presented at TRAFFIC Technology Europe '95, Berlin, 6–7 April 1995.

Bonsall, P. W., Pickup, L. and Stathopoulos, A. (1991) Measuring behavioural responses to road transport informatics. In *Advanced Telematics in Road Transport; Proceedings of the DRIVE Conference*, Vol. 2, pp. 1457–1487. Elsevier, Amsterdam.

Chevreuil, M., Carrara, M., Ruenhaufer, P., Schuessler, R. and Karlsson, T. (1991) The implementation draft plan and the particular need for European traffic information exchange. In *Advanced Telematics in Road Transport; Proceedings of the DRIVE Conference*,Vol.1, pp. 434–453. Elsevier Amsterdam.

Crosby, P., Spyridakis, J., Ramey, J., Haselkorn, M. and Barfield, W. (1993) A primer on usability testing for developers of travel information systems. *Transportation Research C* **2**, 143–157.

De Waard, D., Brookhuis, K. A., Van der Hulst, M., Van der Laan, J. D. (1994) *Behaviour Comparator Prototype Test in a Driving Simulator.* Report V2009/DETER/Deliverable 10 (321B) to the commission of the European Union. University of Groningen, Traffic Research Centre, The Netherlands.

De Waard, D. and Brookhuis, K. A. (1997) Behavioural adaptation to warning and tutoring messages. Results from an on-the-road and simulator test. Special issue 'driver vehicle interactions'. *International Journal of Vehicle Design.*

Godthelp, H. and Schumann, J. (1991) The use of an intelligent accelerator as an element of a driver support system. In *Proceedings of the 24th ISATA International Symposium on Automotive Technology and Automation*, pp. 615-622. Automotive Automation Limited, Croydon, U.K.

Hancock, P. A. and Parasuraman, R. (1992) Human factors and safety in the design of intelligent vehicle-highway systems (IVHS). *Journal of Safety Research* **23**, 181–198.

Hogema, J. H., Van der Horst, A. R. A. and Janssen, W. H. (1994) *A Simulator Evaluation of Different Forms of Intelligent Cruise Control.* Report 1994 C-30. Soesterberg, TNO Human Factors Research Institute, the Netherlands.

Janssen, W. H., Brookhuis, K. A. and Kuiken, M. (1993) *Simulator and Field Evaluation of In-vehicle Collision Avoidance Systems: Report of First Year of Research for Nissan.* Report 1993 C-46. TNO Human Factors Research Institute, The Netherlands.

Kiers, H. A. L. (1990) *SCA: A Program for Simultaneous Component Analysis.* IEC PROGAMMA, University of Groningen.

Kiers, H. A. L. and Ten Berge, J. M. F. (1989) Alternating least squares algorithms for simultaneous components analysis with equal component weight matrices in two or more populations. *Psychometrika* **54**, 467–473.

Kuiken, M. J. and Groeger, J. A. (1993) *Effects of feedback on driving performance at crossroads and on curves.* Report VK 93-12. Haren, University of Groningen, Traffic Research Centre, The Netherlands.

Michon, J. A. and McLoughlin, H. (1991) The intelligence of GIDS. *Advanced Telematics in Road Transport; Proceedings of the DRIVE Conference*, Vol. 1., pp. 371–376. Elsevier, Amsterdam.

Millsap, R. E. and Meredith, W. (1988) Component analysis in cross-sectional and longitudinal data. *Psychometrika* **53**, 123–134.

Nunnaly, J. C. (1978) *Psychometric Theory*, 2nd edn. McGraw-Hill, New York.

O'Keefe, D. J. (1990) *Persuasion. Theory and Research.* Sage Publications, London.

Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. (1957) *The Measurement of Meaning.* University of Illinois Press, Urbana.

Rothengatter, J. A., De Waard, D., Slotegraaf, G., Carbonell Vaya, E. and Muskaug, R. (1991) *Social Acceptance of Automatic Policing and Information Systems.* Report 1033/AUTOPOLIS/Deliverable 7. University of Groningen, Traffic Research Centre, The Netherlands.

Rothengatter, J. A. and Heino, A. (1994) Safety evaluation of collision avoidance systems. In *Towards an Intelligent Transport System. Proceedings of the First World Congress on Applications of Transport Telematics and Intelligent Vehicle-highway Systems*, ed. ERTICO, pp. 2047–2054. Artech House, Boston.

Schofer, J. L., Khattak, A. and Koppelman, F. S. (1993) Behavioural issues in the design and evaluation of advanced traveller information systems. *Transportation Research C* **2**, 107–117.