Towards an On-line Assessment of Subjective Driver Workload

J.M. Girard, M. Wilczyk, Y. Barloy, P. Simon, J.C. Popieul

Laboratory of Industrial and Human Automation, Mechanics and Computer Science (LAMIH), *Université de Valenciennes and Hainaut Cambrésis*, Le Mont Houy, 59313 Valenciennes Cedex 9, FRANCE

Abstract

The number of Advanced Driving Assistance Systems (ADAS) integrated in modern cars has grown over the last few years (e.g., BAS, ACC, navigation systems). Because car manufacturers must ensure that all these systems are compatible with the driving task, evaluation using an experimental protocol is often necessary. When analyzing the data from such evaluations, the driving context must be accurately taken into account since it has an important influence on the driving task. Thus, methods for assessing "mental workload" are often considered useful tools, as they can take both the driving context (e.g., kind of road, traffic, meteorological conditions) and the use of ADAS into account. Though subjective workload assessment methods have proved quite sensitive in many studies, they are not well suited to on-line assessment. To solve this problem, we have adapted the Instantaneous Self Assessment method (ISA), initially developed for air traffic control, for use during the driving task. An experiment was conducted on the SHERPA full-size driving simulator. Fourteen subjects were asked to complete the same 30-kilometer run twice, once with an Adaptive Cruise Control (ACC; a kind of ADAS) and once without. The route comprised stretches on motorways, major roads and mountain roads, and included various traffic densities. The ISA scale was quite senstive to traffic conditions. Moreover, ISA highlighted the prevailing influence of traffic density over road profile. However, the impact of ADAS use could not be ascertained. This study constitutes a first step towards the development of a new on-line workload assessment technique based on an automatic classification system able to take several indicators into account, including vehicle indicators (e.g., speed, brake and clutch pedals) and driver indicators (e.g., eye-gaze position).

INTRODUCTION

The number of systems integrated into cars has grown over the last few years, due to the ease of integration of new technologies and the appeal of these technologies for the consumer. When a new system is implemented in a vehicle, the manufacturer has to ensure that this system is compatible with the driving task. Despite the existence of guidelines for the usability, acceptability [1], and interface design [2] [3] of such systems, Advanced Driving Assistance Systems (ADAS) often need to be evaluated using experimental protocols, involving either real cars or driving simulators.

When evaluating an ADAS, the compatibility of the system with the driving task must be verified with respect to various driving conditions. The purpose of experimental protocols is to gather data related to both vehicle and driver behavior under these varied conditions. During data analysis, the road profile, traffic and, if appropriate, weather conditions must be taken into account, as they have an important effect on the gathered data. Thus, incorporating the idea of driver workload assessment into ADAS evaluations would appear to be a positive addition, since workload is not only influenced by ADAS use, but also by driving conditions. Furthermore, if available on line, workload assessments could be quite a valuable tool for adapting ADAS actions to driver actions in real time, which would improve driving safety [4].

To accomplish both the evaluation and the adaptation of ADAS to the driver and driving conditions, our long-term goal is to develop a non-invasive workload assessment technique based on multivariate classification techniques using objective measurements of elements in the driver-vehicle-environment (DVE) system. However, prior to developing such a technique, we need a reliable on-line subjective workload indicator that is as non-intrusive as possible. This paper presents the method that we have developed and validated through simulation.

Mental workload can be defined as “the degree to which the operator's cognitive resources are absorbed by the task at hand” [5]. The main problem is that most of the workload assessment methods are not adapted for non-intrusive, on-line measurement. For instance, with physiological methods, the driver usually needs to be equipped with sensors, which is hardly non-intrusive. The Modified Cooper Harper Method [6] has already shown its usability for measuring driver workload; however, it is not well adapted for real-time evaluation. This, unfortunately, is true of most existing subjective workload measurement methods.

One method, called Instantaneous Self Assessment (ISA) [7], was developed and initially used in the air traffic control domain for the purpose real-time assessment. Based on a 5-level absolute scale, the original implementation interfaced with a 5-key keyboard. For our study, this method was adapted for a car; the driver-car interface included a small display screen and the +/- buttons of a car radio control unit. A 5-level “relative” scale was added to the original 5-level ISA scale. The basic idea was that it would be easier for drivers to assess variations in their workload than to assess their actual workload level.

The use of the +/- buttons was then easily understandable by the driver: + for an increase in the workload and – for a decrease.

The experiment described this paper examines the variations in driver workload in terms of the driving task demands. These demands are defined according to three factors: road profile, traffic density and use of an ADAS. Three road profiles (i.e., kinds of road) were chosen: motorways, major roads and mountain roads. In order to exaggerate the demands on the driver, opposing conditions were defined for the traffic density factor: high traffic density was defined as roughly 200 vehicles per hour and low traffic density was defined as less than 20 vehicles. The ADAS used during the experimentation was a simulated Adaptive Cruise Control (ACC) that had previously been developed in our laboratory [8]. The ACC was implemented during only one of the two experimental runs. We assumed that if the ACC was available to the driver, then it was used. The subjective driver workload was measured during each of the two runs on both the absolute and relative scales of our adapted ISA method and on the Cooper Harper method's scale. The results obtained with these 3 scales are compared and discussed later on in this paper.

EXPERIMENTAL APPARATUS

The driving simulator

Fig. 1. The Citröen Xantia and the visualization screens

The experiment was conducted on a full-size SHERPA simulator (developed by PSA Peugeot Citroën), reproducing a fully equipped Citroën Xantia This simulator employs four video projectors to provide a 180° front view and a 45° rear view. The environment is computed on a silicon Graphics Onyx 2 workstation. To make the simulation more realistic, the simulator is equipped with a surround sound generator, which uses a 5.1 sound system to emit engine and wind sound inside of the mock-up and the sound of other vehicles outside of the mock-up. The traffic is fully configurable and programmable with a set of scripts. In order to improve the accuracy of the simulation, the experimenter can also control one or more of the vehicles in traffic manually, which makes the interactions between the driver and the traffic more realistic without using complex programming.

The ACC was previously developed on the simulator for another experiment. Using a set cruise speed and a set time headway, the ACC regulates the vehicle speed in terms of the speed of preceding vehicles. The ACC interface is composed of an LCD screen display installed on the dashboard and a car radio remote control unit placed on the side of the steering wheel.

Implementation of the Instantaneous Self Assessment method

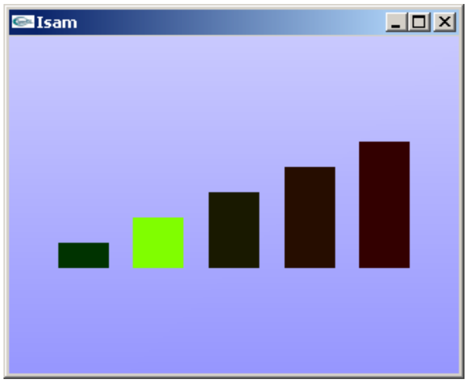

Fig. 2. The absolute scale as displayed on the interface in the simulator mock-up

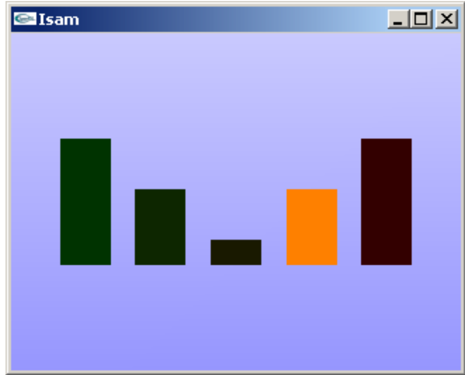

Fig. 3. The relative scale

The ISA scale is shown on the same LCD display interface as the ACC. This scale appears when drivers are asked to assess their workload. When no request is made, the display remains black or shows the ACC interface. When a workload assessment is requested, the ISA emits a sound, and the interface shows a first screen representing a 5-level scale, ranging from low load to high load (fig. 2). Drivers must assess their workload and enter the assessment using the +/- buttons on the control unit. After the driver validates the entry via the mute button, the second "relative" scale appears (fig. 3). This scale is also composed of 5 levels, but this time, the driver is asked to evaluate the workload variation in terms of the previous assessment. The values of the scale range from one extreme ("the load is much lower than the last assessment") to the other ("the load is much higher than the last assessment"). The assessment moments are triggered spatially so that the results are time independent and comparable between subject. During the experiments, the time between the response and the sound stimuli was also recorded.

Method

Participants

The 14 participants (11 males and 3 females) were primarily university students, and they were not remunerated for the experiments. All participants indicated that they drove at least 5000 km per year and that they had been licensed to drive for more than 5 years. The mean age of the participants was 24.26 years (SD: 3.5).

Experimental run

The 32-km circuit was composed of 5 segments. There were 2 motorway segments, with 2 traffic densities ("low" and "high"); 2 major road segments, with 2 traffic densities ("low" and "high"); and 1 mountain road segment that cut through a village. This last segment didn't have any traffic. The roads were extracted from a real map of the Saint-Etienne (France) area. In order to improve the interaction between the subject and the traffic, the experimenter took control of a vehicle on the motorway during a high traffic density segment. All the participants started on the high traffic motorway segment. The driving simulator does not allow the sequence of the road segments to be changed, thus it was impossible to avoid any potential "order" effect.

Procedure

The experiment was conducted in 4 phases. The first phase allowed the drivers to familiarize themselves with the SHERPA driving simulator. In the second phase, the drivers were equipped with an ASL 5000 eye tracker, and were allowed to familiarize themselves with the ACC and the ISA systems. In the third phase, the first half of the subjects were asked to drive the circuit using the ACC; the second half were asked to drive the same circuit, but without using the ACC. This protocol was reversed in phase 4, with the first group not using the ACC, while the second group did.

After each experimental phase, the drivers were asked to fill out one or more questionnaires from the 4 questionnaires used for this experiment. The first questionnaire asked for general information about the drivers and their driving habits (QA). The second asked about physical fatigue (QB), and the last two asked about the driving, with and without ACC (QC1 and QC2). The last two questionnaires also contained a Modified Cooper Harper scale, which was filled out *a posteriori* as the drivers watched a replay of the video tapes. The experiment is summarized in the following figure (fig. 4).

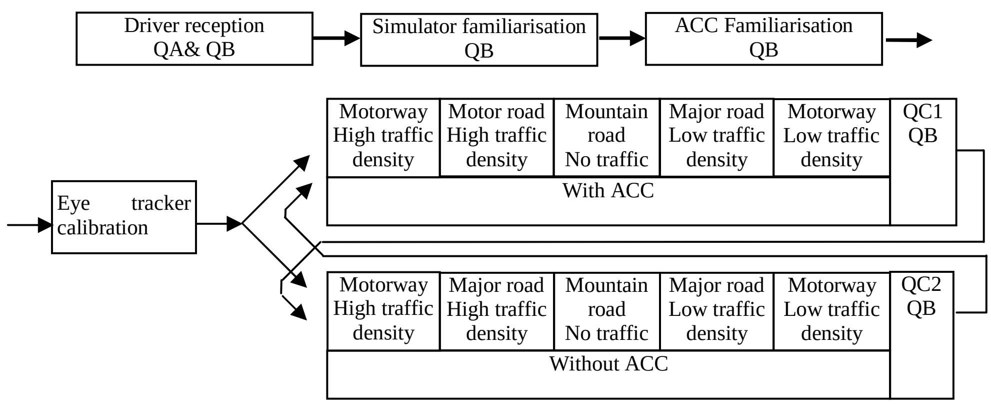

Fig. 4. Finetuning the experiment

Measured variables

During the simulation, several kinds of data related to the interactions between the driver, the vehicle and the environment were recorded automatically, including driver control input, vehicle speed and position. Driver postures, the front view of the simulation, spontaneous verbalizations and eye-gaze positions were also recorded with a video recorder. In this paper, mainly the results obtained with the ISA and the Cooper Harper scales issued from the on-line assessment and the questionnaires are discussed.

Results

Results from the ISA absolute scale

During the run, the drivers were asked to use the ISA interface to evaluate their workload twice on each segment, except on the major road segment with a high traffic density, during which only one evaluation was possible. Each measurement was characterized according to the road profile (H: motorway, M: major road, R: mountain road), the traffic density (LT: low traffic density, HT: high traffic density), and ACC use (with or without ACC).

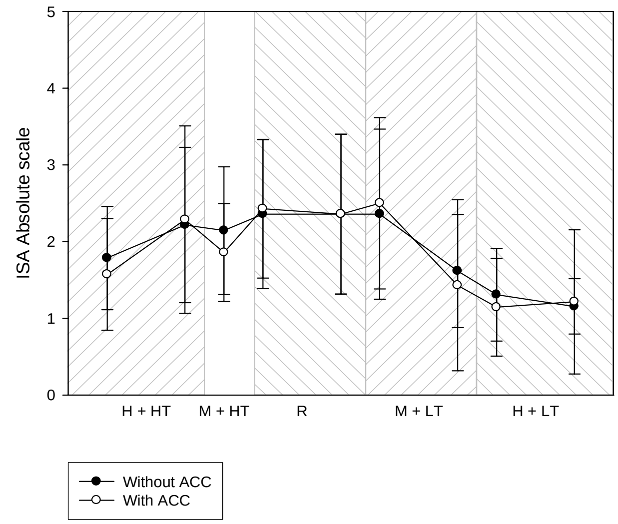

Fig. 5. Results of the ISA absolute scale (mean values obtained for the 14 drivers)

Figure 5 shows that the subjective workload is identical on the motorway with high traffic density and the major road with high traffic density. The workload increases slightly on the mountain road, which yields the same assessment as the major road with low traffic density. This high value was unexpected, and could be interpreted as an edge effect. However, since it was not possible to avoid the possibility of an "order" effect, this explanation remains a hypothesis. On the next road segment—the motorway with a low traffic density—the workload decreases, becoming lower than for the situations with high traffic densities.

The ANOVA analysis shows that traffic density has a significant effect on driver workload (p=0.0146). It also shows that the road profile has a significant effect on the workload (p=0.0071). In addition, there seems to be an important interaction between road profile and traffic density, given that high traffic density produces only a small variation in the workload on the motorway and major road, which is not true for low traffic density. This could be interpreted to mean that traffic density is more important than the road profile.

As expected, the mountain road profile produced the highest workload level of the run. Though this workload is lower when the ACC system is used than when it is not, the differences are not significant. The higher workload difference is encountered on the major road with high traffic density.

Results from the ISA relative scale

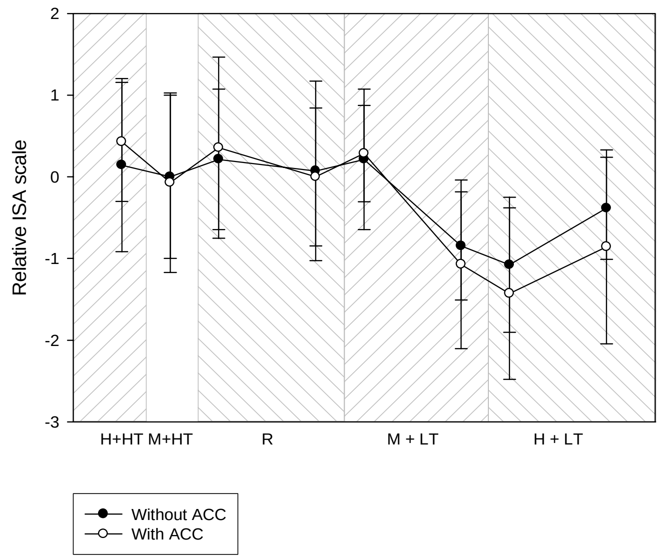

Fig. 6. Results of ISA relative scale

Figure 6 shows that the results obtained with the relative scale are somewhat similar to those obtained with the absolute scale. The workload increases on the motorway, doesn’t increase on the major road segment, and then increases again at the beginning of the mountain road. This same pattern can be observed on the first assessment for the major road with low traffic density, but this time the workload increases while it didn't change in the absolute evaluation. The workload then decreases for the rest of the run.

The ANOVA analysis for the relative scale shows that the effect of traffic density on driver workload is significant (p=0.00003). However, unlike the absolute scale, the effect of the road profile isn’t significant.

Using the results of the relative scale, an absolute scale was reconstructed, the first value being the first assessment on the absolute scale. The successive assessments obtained with the relative scale were then added to the previous value of the reconstructed absolute scale (Fig. 7). The figure obtained is quite similar to figure 5, except that the workload decreases to a greater extent than on the reconstructed scale. The workload moves into the negative numbers for the last segment on the motorway. This could indicate that drivers were more sensitive to the decrease in the workload, meaning they were better able to assess the variations in the workload than the level of workload.

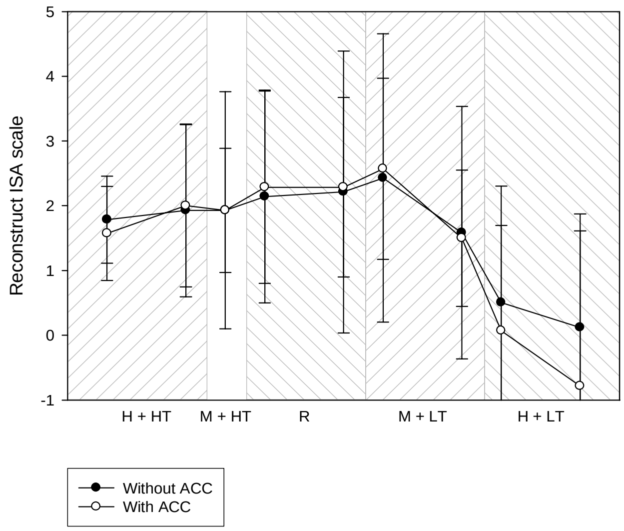

Fig. 7. Result of the reconstruction of the absolute scale

The effect of the ACC on the relative and reconstructed scale cannot be discussed since the results were not significant. The only visible effect is a greater decrease in the workload for the last segments when the ACC is used.

Comparison with the Cooper Harper scale

None of the assessment methods used in this experiment showed a significant variation in the workload with regard to ACC use. Therefore, the effect of ACC use, as examined in this experimental protocol, can not be discussed. The absolute scale appears to be more sensitive to the interaction between road profile and traffic density, with the results indicating that traffic density outweighs road profile. Though the relative scale is also sensitive to traffic density, it does not show a significant effect for the road profile. These results are the opposite of those obtained with the Modified Cooper Harper scale, which highlights a sensitivity to the road profile. This sensitivity is perhaps due to a “memory effect” linked to the *a posteriori* assessment. The on-line ISA assessment methods, on the other hand, are more sensitive to the traffic and better indicate decreases in the workload.

Discussion

The on-line ISA method underlines driver workload variations due to the effect of the driving conditions (i.e., traffic density and road profile). It also classifies the mountain road used in this experiment as requiring the most effort from the driver, and the motorway as requiring the least effort. These results are the same as those from the Cooper Harper scale. However, the ISA assessment also highlights the interaction between the effect of road profile and the effect of traffic density, an interaction in which the traffic density is more important. The Cooper Harper method is not as sensitive to driving conditions because it is subject to a “memory effect” stemming from the fact the method was used for *a posteriori* assessment. The relative ISA scale is more sensitive to workload decreases, but the results on this scale are not significantly different from the results on the absolute scale. On the contrary, it appears to be less sensitive to the road profile than the Cooper-Harper method.

The effect of the ACC used in this experiment can not be discussed in terms of the results from the three scales (Cooper Harper, ISA absolute and ISA relative). The workload assessments appear to be more influenced by driving conditions than by ACC use. These results partially contradict those from a previous study [9], in which ACC use appears to decrease mental effort when used on the motorway, but seemingly has no effect when used on rural roads.

The second step in achieving our long term goal of developing a non-invasive workload assessment technique based on multivariate classification techniques using objective measurements of elements in the driver-vehicle-environment (DVE) system will be to develop a workload assessment technique based on the automatic classification of the instantaneous driving situation. The results of the subjective workload measurement presented in this paper will be used as reference in that classification process. Each driving situation will be characterized as a vector describing driving performance (e.g., speed, lateral position, steering wheel position), driving environment (e.g., type of road, traffic characteristics) and driver status (e.g., eye gaze position). This classification process will allow vector membership values to be attributed to different classes built around a subjective workload indicator. An experimental assessment will then be conducted to validate this workload "indicator".

References

[1] Adaptive Integrated Driver-Vehicle interface project, (2004), Review of existing tools and method, http://www.aide-eu.org/

[2] European commission, (2005), Recommendations from the *e*Safety-HMI Working Group. Available at:
http://europa.eu.int/information_society/activities/esafety/forum/hmi/index_en.htm

[3] ISO 9241-11

[4] W. Piechulla, C. Mayser, Gehrke, W. König, (2003), Reducing drivers' mental workload by means of an adaptive man-machine interface, in Transportation Research Part F 6 (2003,) p 233-248

[5] D. Shaefer, M. S. Straussberger, (2003), Assessment of Mental Workload Part 1 : Basic of Workload Assessment, International Summer School on Aviation Psychology, Graz, Austria, Retrieved March, 23, 2004 MA: http://wundt.uni-graz.at/isap3/downloads/schaefer.pdf

[6] G. E. Cooper and R. P. Harper, (1969), The use of pilot rating in the evaluation of aircraft handling qualities, Moffet Field, NASA Ames Research Center, Aerospace Human Factors Research Division.

[7] A.J. Tattersall and P.S. Ford, (1996), An experimental evaluation of instantaneous self-assessment as a measure of workload, in Ergonomics, vol. 39, n° 5, pp 740-748, 1996.

[8] N. Tricot, M-P. Pacaux and J-C. Popieul, (2004), “Human Machine Cooperation: An automotive application” in Proceedings of 9th IFAC/IFIP/IFORS/IEA SYMPOSIUM, Analysis, Design and Evaluation of Human-Machine System, September 2004

[9] J. Törnros, L. Nilsson, J. Östlund, A. Kircher, (2002), Effects of ACC on driver behavior workload and acceptance in relation to minimum time headway. Proceedings of the 9th workload congress on intelligent transport systems, Chicago, Illinois, USA.