Motorsport Driver Workload Estimation in Dual Task Scenario

A Methodology for Assessing Driver Workload in a Racing Simulator

Luca Baldisserri, Riccardo Bonetti, Francesco Pon
Direzione Gestione Sportiva - Ferrari Driver Academy
Ferrari spa
Maranello (MO), Italy
luca.baldisserri@ferrari.com, riccardo.bonetti@ferrari.com, francesco.pon@ferrari.com

Leandro Guidotti, Maria Giulia Losi, Roberto Montanari, Francesco Tesauri
RE:Lab srl
Reggio Emilia, Italy
leandro.guidotti@re-lab.it, giulia.losi@re-lab.it, roberto.montanari@re-lab.it, francesco.tesauri@re-lab.it

Simona Collina
Università Suor Orsola Benincasa
Napoli, Italy
simona.collina@unisob.na.it

Abstract — **The study aims to define a methodology to assess driver workload by means of a racing simulator applying the dual task paradigm. The experimental plan consists of a series of secondary tasks (i.e., math calculation, imaginative, monitoring and communication tasks) to be performed in a driving simulated context. The preliminary results evidenced that the methodology is a fundamental step forward in order to improve driving abilities of young professional drivers. The described methodology puts the basis for the definition of training programmes for managing high workload racing situations.**

Keywords - Distraction; driver; dual task; motorsport; multitasking; racing; simulator; stress; track; training; workload.

I. INTRODUCTION AND BACKGROUND

This work and the reported experimental plan are intended to define a methodology for assessing the level of cognitive load of young motorsport drivers involved in advanced training programs for single-seater racing cars. The aim of the described methodology is to evaluate potential variation in performance due to the increase of mental workload. Drivers involved in test sessions are engaged in driving in a racing simulator (i.e., primary task) while they are asked to perform also one or more secondary tasks.

Studying the performance of drivers and their "multitasking" skills implies the measurement of their performance first, and the identification of their level of adaptability to increasing workload. The ability and adaptability of multitasking are prominent features of many professions, where the simultaneous management of multiple activities is required and it is sometimes in conflict or in competition over physical and cognitive resources. When multitasking, people can become overloaded as working memory and attentional resources are exhausted [2]. They might also become anxious and frustrated when task challenges outweigh cognitive resources [10]. Furthermore, overall performance can be negatively affected when the demands of one task interfere with those of another task [5]. For instance, a common multitasking everyday situation is the concurrent use of mobile phones while driving, which have been demonstrated to impair driving by, for instance, delaying break reaction times and affecting object detection [4]. Although many studies have shown a generalized impact on attention related to driving distraction tasks, recent research has focused on "supertaskers" [6], by examining individual differences in the performance of multitasking.

In order to highlight abilities of "supertaskers", this study carried out on the short-term working memory [1] that constitutes a field of fundamental analysis to make predictions about the level of resource adaptability. Different skills are required for performing tasks simultaneously, and one must activate different cognitive structures according to tasks, such as the central executive system, the visual-spatial notebook and the phonological loop.

The presented methodology aims to measure the level of racing driver cognitive workload while driving in a simulated circuit, establishing a relationship between the observed cognitive workload and the variation of the driving performance in different scenarios of the dual task paradigm.

The paper is structured as follows: Section II presents the methodology and the design of experiment in details. Section III describes the test protocol and the execution of the experiment. Section IV presents the results of the test. In Section V, the results, the presented methodology, and the related impact are discussed. Conclusion and references close the paper.

II. METHODOLOGY

The methodology for the evaluation of cognitive workload in racing environment has been designed to reach two main goals. On the one side, the objective is to identify the cognitive profile of each individual driver, by assessing the cognitive load. On the other side, once the driver cognitive profile has been outlined, a specific training programme can be identified in order to enhance the abilities and the adaptability of drivers in complex situations. This objective can be achieved by making some target activities automated according to the driver cognitive profile, thus reducing the cognitive and attentional resources allocated to those specific tasks.

In order to outline the cognitive profile, two different kinds of test can be assigned: (1) computer-based tests about cognitive resources and (2) tests in driving and not driving conditions in simulators. Concerning the former point, Individual Differences Measures (IDMs) can be collected assigning to users a series of computer-based tasks to be performed concerning each cognitive structure (i.e., central executive system, visual-spatial notebook, phonologic loop) [2].

In this research, we focused on the latter point, i.e., on the methodology for the assessment of the cognitive workload in driving a racing simulator applying a dual task paradigm. The Dual-Task Paradigm [3] is characterized by a series of secondary tasks assigned in concurrency with the primary driving task.

The reference methodology is the Multi-Attribute Task Battery (MATB) [2], which is used widely in the avionic domain. MATB consists in 4 different tasks simultaneously submitted to users at flight simulators. According to this methodology, the types of secondary tasks drivers might be asked to carry out while driving are reported below.

- Perform complex math calculation (e.g., to count backwards of 7 steps from an assigned number).
- Imaginative task (e.g., mental imagery task).
- Monitoring of system status and tracking (e.g., to indicate an event by pushing a button on the steering wheel or to check the number of engine revolutions).
- Communication task (e.g., to listen to instructions and questions via radio and to answer properly).

Each task has 4 different levels of complexity: automatic (0), low (1), medium (2), high (3).

The Dual-Task Paradigm is aimed to collect objective and subjective measures as well as qualitative and quantitative data to be analysed for assessing if the driver performance improve or decrease, whenever the cognitive workload changes itself due to the variation in the difficulty of the secondary tasks as schematised in Fig. 1.

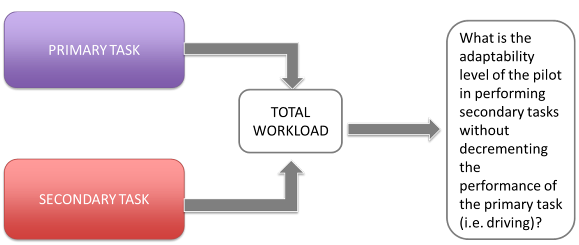

Figure 1. Double task paradigm.

The evaluation of the cognitive workload of racing drivers allows researchers to assess the relationship between workload and changes in the driving performance and then to plan training programmes that may take the cognitive profile of each driver into account for the improvement of his adaptability to multitasking. The estimation of the residual cognitive abilities while performing a Dual-Task Paradigm tests is identified as the result of the comparison between the baseline performance (e.g., the lap time in the driver comfort range) and the performance in terms of lap times carried out experiencing a secondary task.

The methodology applied for the assessment of the cognitive profile of racing driver comes from the assumption that all track, racing, and practice conditions shall be taken into account, and that the driver physical status can change during the driving session. These aspects may be responsible for a significant variation in the level of mental workload while driving. Hence, the following scenarios and conditions have been identified as relevant for test sessions.

- Monitoring of internal and external conditions of the cockpit (e.g., weather, engine revolutions, fuel level).
- Managing radio communication with garage.
- Visualizing data on the steering wheel display.
- Psychophysical degradation related to a high level of mental workload (e.g., sweating or tiredness).

The experimental design has been tailored also by considering some requirements of the driving academy involved in research and, of course, the experimental requirements. They are briefly described below.

- Users should be racing drivers of single-seater vehicles with experience in motorsport domain and should be involved in training programmes.
- Users should be young (i.e., 14 to 22 year old) in order to evaluate future training programs.
- Test sessions need to be carried out in a controlled environment as a lab.
- The main asset required is a single-seater driving simulator for racing. A fully dynamic and immersive racing simulator would be best.
- Material for interviews and collection of data (e.g., datasheet, notebooks, test protocols, chronometer, questionnaires) should be prepared in advance and available at sessions.
- Additional displays should be necessary if visual secondary tasks (e.g., reading of a value; acknowledge following a visual warning) are experienced.
- Radio equipment for remote communication between driver and interviewers should be present.

As far as it concerns the primary task, users are asked to drive in a simulated selected track at the maximum of their skills and performance. The driving performance, measured in terms of lap time within the comfort range of lap times, is considered as baseline. Users are then asked to perform secondary tasks aimed at stimulating each interested cognitive structure (i.e., phonologic loop, visual-spatial notebook and central executive system), either individually

 ISBN: 978-1-61208-340-7

or in association. The goal assigned to the driver is to maintain a level of performance within the comfort lap time range, both in cases of exclusive primary task and in dual task conditions.

Relevant dependent variables are collected in terms of subjective and objective measures. Subjective measures consist of both qualitative data (i.e., post-task comments) and quantitative data, such as NASA-TLX (i.e., Task Load Index) questionnaire, about mental workload [11], and SEQ (i.e., Single Ease Question), about perceived difficulty to carry out tasks [8]. Quantitative objective data are collected by measuring the driving performance in terms of lap time in baseline as well as in dual task conditions, by identifying explicit driving errors (e.g., off-track, over-steering) or modifications in driving behaviours (i.e., by collecting most relevant indicators such as the usage of throttle, brakes, steering wheel, and gear shift), and the number of right/wrong answers reported to secondary task, when asked. The number of answers is considered a measure of efficiency, whereas the speed and correctness in answering are measures of effectiveness. Also direct observation of driving style could be considered, if quantitative data are not available.

According to the identified relevant variables, appropriate tools are used for the collection of measures. In particular, “think-aloud” comments will be written down by observers, numeric values will be reported on scales as answers to questionnaires, and vehicle data will be recorded by the telemetry software of the driving racing simulator.

Different types of data analysis can be performed:

- Analysis and clustering of significant comments, annotations, and answers to open-ended questions.
- Descriptive statistical analysis describing each aspect of interest (e.g., frequency of answering, rate of perceived difficulty in driving, the reported subjective mental workload).
- Inferential statistical analysis of the impact of the dual task on the primary task (i.e., performance in lap time) carried out for subjects and for items.
- Analysis of statistical regression to assess which one of the variables considered is predictive of the performance of the other one, by correlating both the linguistic and the visual-spatial tests to see how much they influence the driving task.

Testing the driver performance in a racing simulator environment cannot disregard some constraints that might affect the testing procedure. In particular, a higher workload level can arise compared to the experience of driving on track with a real single-seater vehicle. This increased workload might be due to the lack of feedbacks from the external environment and to the effort of the racing driver who artificially recreates mentally inputs in order to generate specific driving behaviours. Furthermore, the selection of young racing drivers implies different technical skills and different levels of automations while driving. Depending on the track configuration and on the technical skills of the driver, a track can be more or less difficult to be covered in terms of mental workload. For instance, the "Fiorano" track in Italy can be considered familiar for the Ferrari Driver Academy racing drivers and it is assumed to imply a medium level of mental workload compared to the "Monza" track, which generates a variable level of mental workload depending on the track segments and on the number and type of corners. Other minor aspects could impact the workload, for instance the use of a standardized rather than a customized steering wheel for each racing driver can influence the level of familiarity and the automatic driving procedures. Moreover, the variable and permanent conditions of real tracks are missing, such as the high or low grip feeling, the weather conditions, and the presence of other vehicles that can only be simulated. It is not possible to simulate this variability and to control it as an independent variable in order to increase the difficulty level due to changing external context conditions.

III. TESTING OF METHODOLOGY

The test protocol has been designed for sessions on a fully dynamic and immersive single-seater driving simulator for racing, as shown in Fig. 2.

Figure 2. Racing driving simulator facilities

The test session has been carried out on the 13th of March 2013 involving one young professional driver of the Ferrari Driver Academy, 19 years old. The simulated single-seater vehicle has been his "F3 Euro Series" in the simulated track of "Autodromo Nazionale Monza", Italy (5793 km, 8 corners, clockwise). The track is represented in Fig. 3.

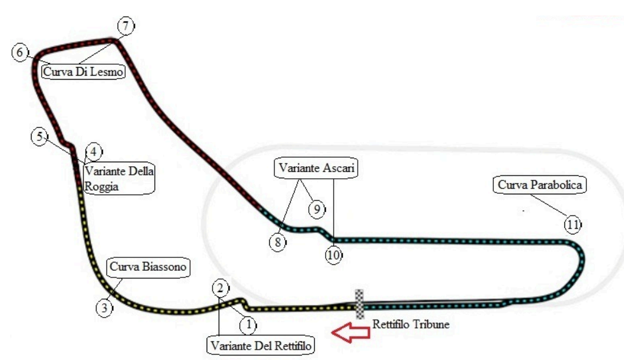

Figure 3. Autodromo Nazionale Monza

 ISBN: 978-1-61208-340-7

The test session lasted about 120 minutes and it consisted of 40 laps with 3/4 minutes of stop between each test run.

The test protocol has been designed considering 1 run of warm-up, aimed at identifying the "comfort lap time range" for the driver and 1 run for each task session. Each run consists of 1 out-of-garage lap and 3 timed laps.

The comfort lap time range (i.e., minimum lap time and maximum lap time) for the involved user is within 1:46.200 - 1:47.200 in Monza track, considering a comfort variance of 2%. The baseline sessions, in which the driver is asked to focus only on the primary task, are followed by SEQ and NASA-TLX questionnaires. After that, a sequence of randomly selected test run including secondary tasks is assigned to the driver. The user performs, firstly, only the secondary task, in not driving conditions, and he is then asked to fill in SEQ. Then, he experienced the secondary task in driving conditions, followed by SEQs referring both to primary and to secondary tasks, and by the NASA-TLX questionnaire about perceived mental workload. The baseline is repeated three times, at the beginning, in the middle and at the end of the test session. In the performed test session, the randomized task sequence has been: B, T1, T3, T2, B, T5, T4, B, which are described below.

- T1 = Math identification if an assigned number is grater or less than 45 (in not driving and driving conditions).
- T2 = Count backwards of 3, step by step, starting from an assigned number (in not driving and driving conditions).
- T3 = Count backwards of 7, step by step, starting from an assigned number (in not driving and driving conditions).
- T4 = Count backwards of 13, step by step, starting from an assigned number (in not driving and driving conditions).
- T5 = Count the number of letters of a given word (in not driving and driving conditions).
- B = Baseline run.

The user was alone in the driving simulator room. During test sessions, the researcher was able to communicate with the user through one-way at a time radio. As in real "F3 Euro Series" racing context, user had to press a button on the steering wheel in order to enable communication and let researcher hear him while speaking. Also, the researcher had to press button to be heard by user. These actions are mutually exclusive. The researcher has been in charge of the task assignment, such as to provide to the user the list of numbers according to the given task.

After each test phase, SEQ has been provided to the driver. It is a 7-points scale item from "Very difficult" to "Very Easy" to be self reported by the user. The objective of the measure is to evaluate the perceived easiness of the primary task (i.e., driving performing a lap time within the comfort range) not experiencing (i.e., in baseline) or experiencing the secondary tasks, and also the perceived easiness of the secondary task itself, in stationary and driving conditions. SEQ has concerned the following topics.

In baseline - "How do you evaluate the driving task on track? (i.e., how difficult or easy it has been?)"

After experiencing the secondary task stationary - "How do you evaluate the 'secondary task'? (i.e., how difficult/easy has it been?)"

After experiencing the dual task while driving - "How do you evaluate the driving task on track while you are performing the 'secondary task'? (i.e., how difficult/easy has it been?)"

After experiencing the dual task while driving - "How do you evaluate the 'secondary task' while driving on track? (i.e., how difficult/easy has it been?)"

After each driving session, both in baseline and in dual task, the user is also asked to fill in the six scales of NASA-TLX questionnaire. RTLX (i.e., Raw NASA-TLX) version of the questionnaire has been considered [11].

The driving performances have been recorded by Atlas software for telemetry. It collected data from different channels (e.g., brake, accelerator, steering wheel angle, speed, engine revolution, gear, wheel speed, etc.) and the press of the buttons on the steering wheel.

IV. RESULTS

The data analysis has shown that it is possible to point out specific remarks about dual task performance in the motorsport driving domain by carrying out test sessions applying the described methodology.

With respect to lap timing, a remarkable impact on performance has been registered according to the type of secondary task. Timings in dual task conditions are higher (M_a= 01:47.5) in comparison with those in baseline conditions (M_a= 01:46.7). In three occasions, the lap time is higher than the comfort range threshold (i.e., between 01:46.2 and 01:47.2), as shown in Fig. 4 and in Table I.

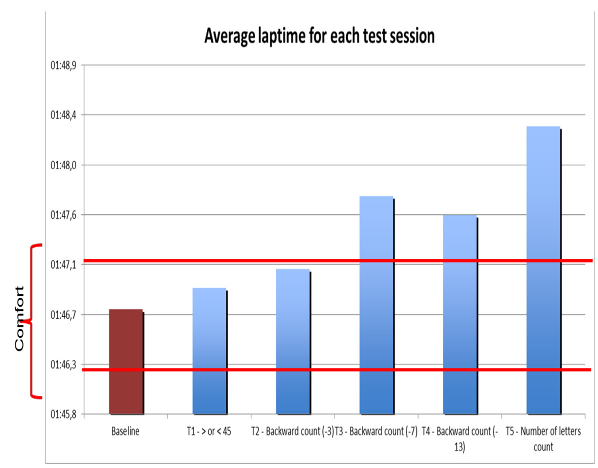

Figure 4. Average lap time

TABLE I. AVERAGE LAP TIME

Test session	Average laptime
Baseline	01:46,7
T1 - > or < 45	01:46,9
T2 - Backward count (-3)	01:47,1
T3 - Backward count (-7)	01:47,7
T4 - Backward count (-13)	01:47,6
T5 - Number of letters count	01:48,3

The perceived mental workload by NASA-TLX and the self-reported changes depend on the type of secondary task, as shown in Fig. 5 and Table II.

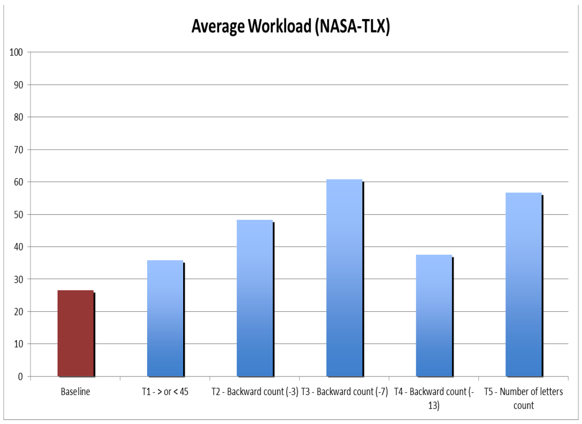

Figure 5. Average Workload (NASA-TLX)

TABLE II. NASA-TLX RESULTS

Test session	Average Workload (NASA
Baseline	26,67
T1 - > or < 45	35,83
T2 - Backward count (-3)	48,33
T3 - Backward count (-7)	60,83
T4 - Backward count (-13)	37,50
T5 - Number of letters count	56,67

The perceived easiness while performing the dual task changes among the tasks themselves and it changes for the same task in stationary or driving conditions, as shown in Fig. 6.

The secondary tasks that are considered the most demanding and that seem to have a greater impact on lap times and on perceived mental workload are the backward counting of -7 steps (NASA-TLX score 60,83 and easiness 3/2/2) and the identification of the number of letters in a word (NASA-TLX score 56,67 and easiness 3/4/3). Nevertheless, an effect of self-learning may have happened while performing the secondary tasks.

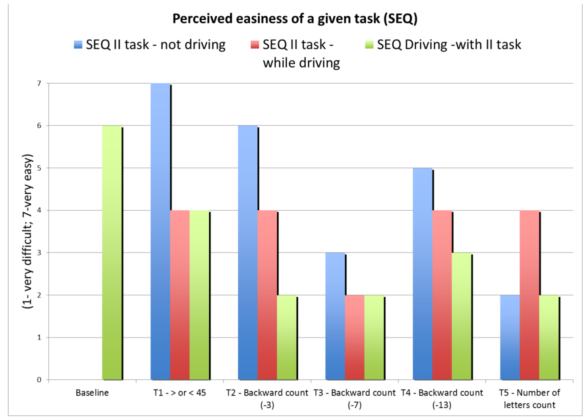

Figure 6. Perceived easiness of tasks

The times for answering to assigned sub-tasks reflect the difficulties perceived while performing the tasks, as shown in Fig. 7.

All the surveys and the results seem to be consistent with the perceived easiness, the errors occurred in the answers provided to secondary tasks (Fig. 8), workload, and the level of performance (i.e., lap time).

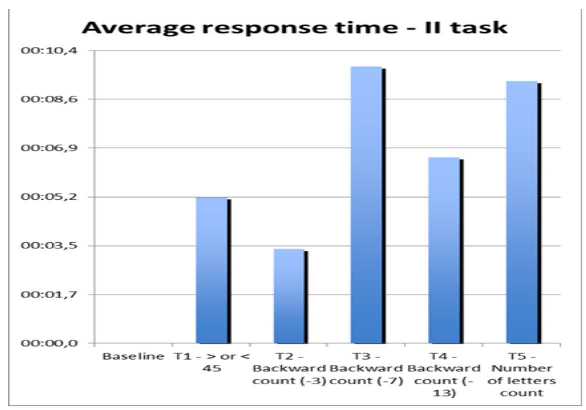

Figure 7. Average response time

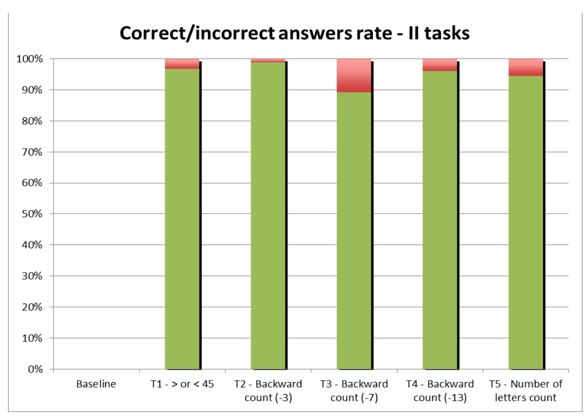

Figure 8. Correct answer rate

TABLE III. ERRORS AND RESPONSE TIME

Test session	Item Number	Errors	% errors	Average time per item
Baseline				
T1 - > or < 45	62	2	3,23%	00:05,2
T2 - Backward count (-3)	96	1	1,04%	00:03,3
T3 - Backward count (-7)	33	4	12,12%	00:09,8
T4 - Backward count (-13)	49	2	4,08%	00:06,6
T5 - Number of letters count	35	2	5,71%	00:09,3

The user involved in the test session has a maximum error rate counting back of 7 steps task (i.e., 12,12 % of error rate) and he takes more time (i.e., 00:09.8) to carry out sub-tasks rather than other dual tasks, as reported in Table III.

V. IMPACT AND DISCUSSION

A methodology to analyse the multitasking ability of a young motorsport driver while performing a dual task paradigm on a racing simulated track has been designed. A preliminary test has been carried out. Researchers collected and analysed data to evaluate if any degradation could be detected in the driving performance due to the concurrency with increasing demanding cognitive tasks.

Findings cannot be considered statistically relevant due to the narrow sample of users, but it has served the scope to identify and to test the designed methodology for data collection and data analysis. Thus, it can be considered an indicator of the expected impact of cognitive workload on the driving performance in a racing context.

Driver has autonomously and constantly performed the requested dual tasks. The secondary tasks seemed to induce a diversified level of workload pressure, pointed out by the self-reported data, perceivable also by the tone of the voice and by the speed of the answers provided by the user.

In some conditions, driver has not immediately answered to or performed the secondary task, but he took some time to complete the task assigned. From the observation, it seemed that, when approaching the corners of the Monza track, the driver was focused exclusively on the primary task and he started performing the secondary task only after the completion of the manoeuvre.

Learning strategies in the answering to dual task requests (e.g., to count -30 or -3 when asked to count back of 13) does not seem to influence the performance of the dual task and the global impact it has on the primary task. Driver encountered an initial physical difficulty, reported as a speech comment, when pushing the “radio” button while changing the gear or when approaching the corners. Furthermore, possible relevant changes in the driving performance by analysing the telemetry channels shall be assessed and interpreted. The survey in baseline conditions and double task will allow an interpretation of the driving style in the different conditions of mental workload.

VI. CONCLUSION

Research findings shall be considered as interpretations referred to collected data and to the descriptive results presented. There are no specific considerations and interpretations about the performance and driving style in different test sessions. There are no interpretations related to or statistic interferences among the data collected. For practitioners and researchers, it shall be highlighted that motorsport, together with the avionic domain, is a well suited context for the goal of testing multitasking ability of young drivers and the impact of workload on the driving performance. Training programmes can be easily defined thanks to the results gathered by applying the methodology, with the aim of improving the performance of the driver by increasing his/her multitasking ability and the ability to face concurrent task on cognitive resources. Although only one user has been involved in the tests performed so far, the expected impact has been verified by the collected data: the young professional driver registered large variations in the driving performance in terms of lap time, due to the secondary tasks. In addition to that, the methodology applied has been solid, coherent and correct in terms of results. It is also in line with the studies that have been carried out in the last years concerning the dual task paradigm. In order to enhance the obtained results, the objective will be to consolidate the methodology by involving a larger sample of drivers in test sessions as a way to identify areas of improvement and to evaluate if such improvements have been reached thanks to the training activity.

VII. REFERENCES

[1] A. Baddley, Human Memory: Theory and Practice, Lawrence Erlbaum Associates, London, UK, 1990.

[2] B. Morgan, S. D’Mello, R. Abbott, G. Radvansky, M. Haass, and A. Tamplin, "Individual Differences in Multitasking Ability and Adaptability", Human Factors: The Journal of the Human Factors and Ergonomic Society, vol. 55, 2013, pp. 776-788.

[3] C.D. Wickens, "Processing resources and attention", Multiple Task Performance, Taler & Francis, Ltd., Bristol, UK, 1991, pp. 3-34.

[4] D. L. Strayer, F. A. Drews, and W. A. Johnston, "Cell phone induced failures of visual attention during simulated driving", Journal of Experimental Psychology: Applied, vol. 9, 2003, pp. 23-52.

[5] E. M. Altmann and W. D. Gray, "An integrated model of cognitive control in task switching", Psychological Review, vol. 115 (3), 2008, pp. 602–639.

[6] J. M. Watson and D. L. Strayer, "Supertaskers: Profiles in extraordinary multitasking ability", Psychonomic Bulletin & Review, vol. 17 (4), 2010, pp. 479-485.

[7] J. R. Comstock, and R. J. Arnegard, The Multi-Attribute Task Battery for human operator workload and strategic behaviour research, National Aeronautics and Space Administration Technical Memorandum No. 104174, Washington D.C., USA, NASA, 1992.

[8] J. Sauro and J. S. Dumas, "Comparison of three one-question, post-task usability questionnaires", Proceedings of CHI, ACM, Boston (MA), USA, 2009, pp. 1599-1608.

[9] L. Angell et al., Driver Workload Metrics Project, Task 2 Final Report, CAMP Driver Workload Metrics, Report No: DOT HS 810 635, U.S. Department of Transportation NHTSA, USA, 2006.

[10] M. Csikszentmihalyi, Beyond boredom and anxiety, Jossey-Bass, San Francisco (CA), USA, 1975.

[11] S. G. Hart, “NASA-TLX Load Index; 20 years later”, Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, Santa Monica (CA), USA, 2006, pp. 904-908.