Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
INGEGNERIA CIVILE, CHIMICA, AMBIENTALE E DEI MATERIALI

Ciclo XXXI

Settore Concorsuale: 08/A3 Infrastrutture e sistemi di trasporto, estimo e valutazione

Settore Scientifico Disciplinare: ICAR/04 Strade, Ferrovie ed Aeroporti

A MULTIDISCIPLINARY RESEARCH APPROACH FOR EXPERIMENTAL APPLICATIONS IN ROAD-DRIVER INTERACTION ANALYSIS

Presentata da: Arianna Bichicchi

Coordinatore Dottorato
Prof. Ing. Luca Vittuari

Supervisore
Prof. Ing. Andrea Simone

Co-Supervisore
Dott. Ing. Valeria Vignali

Esame finale anno 2020

Abstract

This doctoral dissertation represents a cluster of the research activities conducted at the DICAM Department of the University of Bologna during a three years Ph.D. course.

In relation to the broader research topic of “road safety”, the conducted activities focused on the investigation of the interaction between the road and the drivers according to human factor principles and supported by the following strategies: 1) The multidisciplinary structure of the research team covering the following academic disciplines: Civil Engineering, Psychology, Neuroscience and Computer Science Engineering. 2) The development of several experimental real driving tests aimed to provide investigators with knowledge and insights on the relation between the driver and the surrounding road environment by focusing on the behaviour of drivers. 3) The use of innovative technologies for the experimental studies, capable to collect data of the vehicle and on the user: a GPS data recorder, for recording the kinematic parameters of the vehicle; an eye tracking device, for monitoring the drivers’ visual behaviour; a neural helmet, for the detection of drivers’ cerebral activity (electroencephalography, EEG). 4) The use of mathematical-computational methodologies (deep learning) for data analyses from experimental studies.

The outcomes of this work consist of new knowledge on the casualties between drivers’ behaviour and road environment to be considered for infrastructure design. In particular, the ground-breaking results are represented by:

- the reliability and effectiveness of the methodology based on human EEG signals to objectively measure driver’s mental workload with respect to different road factors;
- the successful approach for extracting latent features from multidimensional driving behaviour data using a deep learning technique, obtaining driving colour maps which represent an immediate visualization with potential impacts on road safety.

Table of contents

List of figures

Acronyms

AI: Artificial Intelligence

ANN: Artificial Neural Network

DAE: Denoising Autoencoder

DL: Deep Learning

IVS: Instrumented Vehicle Study

KPCA: Kernel-Principal Component Analysis

ML: Machine Learning

MWL: Mental Workload

NDS: Naturalistic Driving Study

PCA: Principal Component Analysis

RSA: Road Safety Audit

RSI: Road Safety Inspection

SDAE: Stacked Denoising Autoencoder

SER: Self-explaining roads

S-NDS: Semi- Naturalistic Driving Study

VRUs: Vulnerable Road Users

WL: Workload

Keywords

Driving Behaviour

Eye Tracking

Human Factors

Machine-Learning

Mental Workload

Road Safety

Road-Driver Interaction

1. Introduction

1.1 Motivation

1.1.1 Problem statement

This doctoral dissertation deals with the topic of Road Safety, that field of research that studies methods and measures used to prevent road users from being killed or seriously injured.

The importance of this subject results evident from the yearly statistics on road fatalities and injuries: despite the increasing diffusion of road safety policies, safety measures as well as stronger safety laws and public awareness safety campaigns, the number of deaths and injuries in the EU is undoubtedly too high. The World Health Organization recently estimated that 25,100 people still lost their lives on EU roads in 2018 and about 135,000 were seriously injured (WHO, 2018).

As this is an unacceptable and unnecessary human and social price to pay for mobility, the European Commission has recently put forward a new and ambitious road safety agenda, which includes the challenging long-term goal of moving close to zero fatalities and serious injuries in road transport by 2050. The stand-out ambition of such target emphasizes the importance of pushing forward actions and strategies addressed to find more effective and efficient ways to reduce accidents and save lives on roadways.

1.1.2 Approach

Many investigations have proved that individual road accidents are complex events involving a variety of contributing factors mainly attributable to three categories represented by the infrastructure, the vehicle and the driver. Traditionally, the analysis of risk examines the three factors separately and there is a tendency by researchers and practitioners to look for one or a few factors, when in actual fact they should be analysing multiple factors. Accordingly, the traditional approach to road safety tackles the following well-known interventions:

a. Safer infrastructures, through planning and design;

b. Safer vehicles, through better crashworthiness, active vehicle safety, and vehicle inspections;

c. Human behaviour, through legislation, enforcement, and campaigns.

Nevertheless, according to the Venn's diagram typically used for the representation of the road safety paradigm, how do the factors relate between them is a less explored but fundamental understanding, allowing to introduce a broader perspective including the combination of the factors contributing to the occurrence and to the severity of a crash (Figure 1).

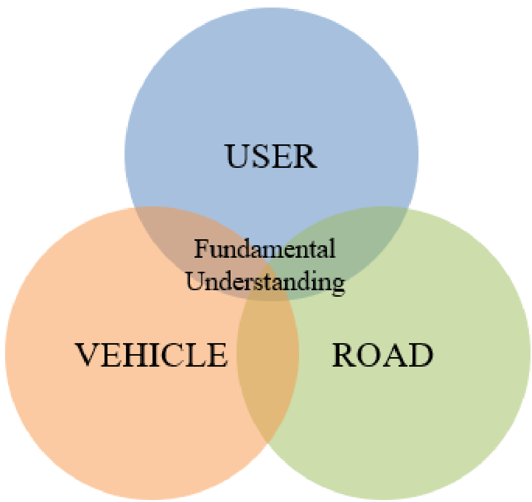

Figure 1. Venn's diagram representing the road safety paradigm.

In relation to this, William Haddon developed a matrix that identifies risk factors before the crash, during the crash and after the crash, in relation to the person, vehicle and environment (Figure 2). Haddon described road transport as an ill-designed "man-machine" system in need of comprehensive systemic treatment and interactions between different components must be taken into account (Haddon, 1980). The essence of using a system approach is to consider the fact that road traffic injuries are a multidimensional problem that require a comprehensive view when examining the determinants, consequences and solutions. Reducing road traffic crashes to one "cause" only imply to consider the components of the system – human, infrastructure and vehicle factors – as independent. Measures addressing any one component can thus be implemented separately but the opportunities to influence one type of factor through another (for example, to obtain more appropriate driver behaviour through changes in road design) are entirely ignored.

		FACTORS		
PHASE		HUMAN	VEHICLES AND EQUIPMENT	ENVIRONMENT
Pre-crash	Crash prevention	Information Attitudes Impairment Police enforcement	Roadworthiness Lighting Braking Handling Speed management	Road design and road layout Speed limits Pedestrian facilities
Crash	Injury prevention during the crash	Use of restraints Impairment	Occupant restraints Other safety devices Crash protective design	Crash-protective roadside objects
Post-crash	Life sustaining	First-aid skill Access to medics	Ease of access Fire risk	Rescue facilities Congestion

Figure 2. Haddon's matrix identifying risk factors before the crash, during the crash and after the crash, in relation to the person, vehicle and environment.

In conclusion, as crash factors relate to human as well as to physical and technical components of the road and transport system, a detailed analysis of road crashes and consequently of "road safety" requires a multidisciplinary approach (WHO, 2006), which is the major scope of this dissertation.

1.2 Research objectives

Road engineering is the body of research traditionally appointed to tackle the challenge of finding evidence in the relationship between infrastructure and road crashes aiming to define indications, rules and good practices in road design and maintenance.

Considering that the infrastructure factor is estimated to contribute, alone or in combination with other factors, for the 30% to total yearly crashes in Europe (EU Commission, 2018) and that human error is the main cause of the 57% of road accidents and a contributing factor in over 90% of them (Treat, Tumbas, McDonald, Shinar, & Hume, 1979), a "broader perspective" for the role own by road engineers is represented by the analysis of the interaction occurring between the infrastructure and the driver, which is possible by including ergonomics principles to the traditional analyses. This research approach deals with the commonly known "Human Factors engineering", a science dealing with the application of information on physical and psychological characteristics to the design of systems (in this case the road or the vehicle) for human use.

As the knowledge of Human Factors can help to better understand how road users behave and then address adjustments to the design of the environment according to human capabilities (Theeuwes, 2012), the strong ambition carried out with the investigation and application of those principles to road design is the reduction of both the whole number of accidents occurring on roads and the severity of those accidents.

With reference to the problem statement described above, the main objectives of the research project presented in this dissertation are:

- The identification and acquisition of the transversal knowledge and requirements needed to convert behavioural studies in engineering measures. In particular, are focused the causes of the manifestation of certain behaviours and the cause-effect mechanisms that regulate man-road interaction. That purpose, which is hard to reach because of the complexity of the arguments, is enabled throughout the cooperation between multidisciplinary research teams including psychologists, neuroscientists and computer science engineers sharing their knowledge in order to provide qualitative advances in the research.
- The definition of engineering procedures that allow the use and application of Human Factors principles in road design and analysis (e.g. road safety audit – RSA; road safety inspection - RSI). Indeed, after understanding the process that regulate man-road interaction, the research project wants to identify the procedures that allow the use and application of Human Factors principles in road design (or implement the ones that already exist).

The consolidated experience in the development of road safety studies owned by the DICAM Department of the University of Bologna allowed to identify the proper methodologies for the research program, identified mainly with observational studies on existing roads. In addition, the multidisciplinary collaboration with national and international research institutions, provided the opportunity to address complex and interconnected challenges of the future of road safety where it is no longer possible for these issues to be solved in a single discipline or profession.

2. Research program

This dissertation, structured in the form of a “collection of papers”, presents a wide part of the activities carried out in the doctoral course which converged into five journal papers. Those papers are an integral part of the dissertation and the reason why most of details are intentionally not recalled by the chapters is because publications are conceived to describe the methodologies, the numerical or statistical analyses and the results of the research in the most effective and clear way, representing lasting contribution to science.

The author, thus, recommend the reading of full papers included in the Annexes section to catch most of the details on the methodologies and results. Vice versa the chapters, as the overall thesis, represent the tool aimed at ensuring that the multidisciplinary research performed can be put into perspective and the pieces of work performed are connected and aimed at a same final scope.

2.1 List of publications

The order of the listed papers follows the order of the activities indicated in the research program and discussed in the manuscript.

- Vignali V., Bichicchi A., Simone A., Lantieri C., Dondi G., Costa M., (2019). *Road sign vision and driver behaviour in work zones*. Transportation Research Part F: Traffic Psychology and Behaviour, 60, pp. 474-484. ISSN 1369-8478, https://doi.org/10.1016/j.trf.2018.11.005.
- Costa M., Bonetti L., Vignali V., Bichicchi A., Lantieri C., Simone A., (2019). *Driver's visual attention to different categories of roadside advertising signs.* Applied Ergonomics, 78, pp. 127-136. ISSN 0003-6870, https://doi.org/10.1016/j.apergo.2019.03.001.
- Costa M., Bichicchi A., Nese M., Lantieri C., Vignali V., Simone A., (2019). *T-junction priority scheme and road user’s yielding behavior.* Transportation

Research Part F: Traffic Psychology and Behaviour, 60, pp. 770-782. ISSN 1369-8478. https://doi.org/10.1016/j.trf.2018.12.009.

- Di Flumeri G., Borghini G., Aricò P., Sciaraffa N., Lanzi P., Pozzi S., Vignali V., Lantieri C., Bichicchi A., Simone A. and Babiloni F., (2018). *EEG-Based Mental Workload Neurometric To Evaluate the Impact of Different Traffic and Road Conditions in Real Driving Settings*. Frontiers in Human Neuroscience, 12, 509. 10.3389/fnhum.2018.00509.
- Bichicchi A., Belaroussi R., Simone A., Vignali V., Lantieri C. and Li, X. (2020). *Analysis of Road-User Interaction by Extraction of Driver Behavior Features Using Deep Learning*. IEEE Access, 8, pp. 19638-19645. 10.1109/ACCESS.2020.2965940

2.2 Author contribution

The contribution of the author to papers consisted:

- in the definition of the experimental protocol, development of the experimental activities, data analyses and paper writing (for papers n. 1, 2, 3, 5);
- in the definition of the experimental protocol, development of the experimental activities, minor data analyses and minor paper writing (for paper n. 4).

2.3 Research activities

As indicated in the following Figure 3, the research framework relies on a series of original research activities aimed to understand in-depth the interaction between the road and the road users.

The activities have a common experimental strategy represented by the conduction of observational studies consisting in Instrumented Vehicles (IVS) tests with a standard experimental protocol, whereas the ground-braking multidisciplinary level of the activities includes insights dealing with the following subjects:

- Civil engineering (i.e. road/traffic engineering);
- Psychology;
- Neuroscience;
- Computer Science.

Therefore, the different features belonging to each activity consist in a different experimental query, rationale, investigated data, technologies involved and data analyses.

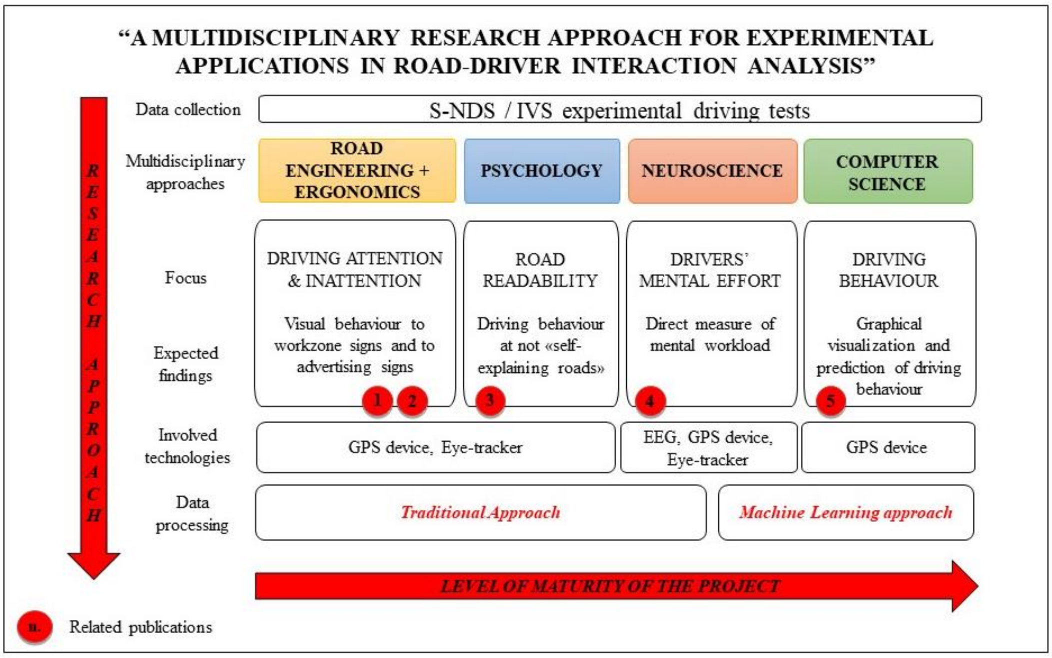

Figure 3. Framework of the doctoral thesis.

The **research activity I** deals both with road engineering and ergonomics and is related to the understanding of drivers' visual behaviour towards the road environment. Following the experiences of previous activities investigating the visual behaviour of drivers and the identification of the elements of road environment that most attract the drivers and how they affect them (Vignali, 2019), in this activity the focuses of a twinning experimentation are workzone signs and advertising signs.

Basically, the dual scope of the activity is:

- To quantify and understand the effectiveness of workzone signs as elements of attention, indicating potential danger situation along the roadway, and potentially allowing to define new positioning/designing methodologies and best practice measures for those road elements for safer roads (*1st publication*);
- To quantify and understand the potential distraction offered by advertising signs and billboards on the roadway (*2nd publication*).

In both cases, the comparison between the visual behaviour data and the drivers' performance, allowed to understand and make considerations on the effects induced by the visualization of both categories of signs.

The **research activity II** deals with psychology, in particular with the "self-explaining road" concept. This concept is used to define road design principles in accordance with drivers' expectations. In the conducted experimentation, T-intersections defined as "not-explicative" are investigated to understand drivers' compliance to yielding rules (*3rd publication*) and make considerations on road design characteristics.

The **research activity III** deals with neuroscience. More precisely, the concepts applied in this experimental study belong to the field of "neuroergonomics", the research branch aiming to study the relationship between the human behavior and the brain at work (Parasuraman & Rizzo, 2008). Neuroergonomics provides a multidisciplinary translational approach that merges elements of neuroscience, cognitive psychology, human factors and ergonomics to study brain structure and functions in everyday environments. Applied to the scenario of driving a car, a "neuroergonomics" approach allows to investigate the relationship between human mental behavior, performance and road safety, providing a deeper understanding of human cognition and its role in decision making and possible errors (Lees, Cosman, Lee, Rizzo, & Fricke, 2010). The activity consisted in a pilot test involving an EEG device aiming to tackle the challenge of a real time measurement of driver's workload, the commitment to which drivers are subjected while driving, and its comparison to visual behaviour and driving performance (*4th publication*).

The **research activity IV** deals with Computer Science and, differently from the previous activities, tackles the issue of the processing of data. Machine Learning techniques are envisaged as an effective strategy to enhance research thanks to a more effective processing of the large amounts of data coming from sensors and data recording devices, which are generally difficult and expensive activities. The conducted study represents, thus, one of the first attempts of a deep learning application for analyzing kinematic data from driving datasets (5th *publication*).

2.4 Data collection

Traditionally, three methods are commonly used by road safety researchers to collect data useful for the conduction of road-driver analyses:

- Self-reports;
- Simulation of the driving task;
- Observation through naturalistic driving studies (NDS).

The scientific community divides on the reliability of self-reported driving habits: some studies positively consider them for detecting Human Factors data related to driving behaviour (Åberg et al., 1997; Prabhakar et al., 1996; Ulleberg, 2002), whilst others consider their use only as a complement of driving simulator studies and NDS (Kapitaniak, 2015). Considering, instead, driving simulators or direct observation methods, one of the arguments criticising is that there is a possibility that drivers could be more disciplined than they would be when they feel they are not being observed (Ulleberg, 2002).

In this research program, the data collection strategy is the same for all the aforementioned activities. The choice of the author’s research group consisted in the conduction of a particular typology of driving observations called “Instrumented Vehicle” studies (IVS) or “Semi-Naturalistic Driving” studies (S-NDS), which refer to driving tests where subjects are requested to drive in real traffic but in a special, highly equipped vehicle and with an experimenter on-board. With difference to NDS studies, who have no experimental control and undergo the impact of the long duration and high costs, IVS have a great potential in providing insights into the actual real-world behaviour of road users with acceptable time and costs. The experimental conditions and related biases who may affect the outcomes are reasonably foreseen a priori and thus limited by means of proper experimental protocols (e.g. provision of adaptation time to drivers), enabling the investigators to understand with reliability what happens on the road in real traffic situations (Barnard, 2016; Kim, 2017).

3. Activity I: Experimental study on drivers' visual behaviour

3.1 Introduction

3.1.1 Visual behaviour

Driving is a task strongly linked to human perception, being both a sight-based and an experienced-based activity, which means that car drivers are both influenced by what they see (road and road environment) and what they have seen.

Literature suggests that the 90% of the information that drivers use in the accomplishment of that task is visual (Castro, 2009) and, according to Kapitaniak, it is substantially attributable to navigation (to go from one place to another), to driving (e.g. selecting a lane) and for checking the vehicle (i.e. setting speed, braking and turning) (Kapitaniak, 2015). Even some other research affirms that a similar quantification is difficult (Sivak, 1996), most people would nevertheless agree on the undiscussed importance of the visual information collected and used by drivers.

The continuous search for useful information occurs according to drivers' own visual strategy; the importance of those strategies and of the related behaviors relies in the possibility to gain insight for understanding the task of vehicle driving as-a-whole, in the broader view of their consequences in terms of safety.

The method of tracking eye movements is currently the preferred method chosen for the study of different types of cognitive strategies, in particular visual strategy. Eye tracking, in fact, has been gaining in popularity over the past decade as a window into observers' visual and cognitive processes. Although the number of publications on the analysis of eye movements while driving has increased, it appears that much of the research in this area is not published basically because the experimentations with eye tracking devices undergo several limitations and the access to visual information by users can be very difficult (occurring conditions such as darkness, inclement weather, glare due to the sun during sunset; physical characteristics of test participants; etc.). When those limitations are, somehow, overcame, researchers typically analyse eye movements in terms of the following parameters:

- Gaze points: are the instantaneous spatial locations of the visual axis landing on the stimulus. As such, they have an (x, y) coordinate and a timestamp corresponding to its measurement. To understand their entity, with devices operating at 300 Hz gaze points are spaced a mere 3 milliseconds apart;
- Fixations (the period of time where the eye is kept aligned with the target for a certain duration, allowing for the image details to be processed, i.e. longer gazes);
- Saccades (are the type of eye movement used to move the fovea rapidly from one point of interest to another, representing thus the rapid movements between fixations).

Accordingly, the main indicators evaluated for the assessment of visual strategies are:

- the number of fixations;
- the fixations length (time);
- the exploration areas of the functional vision field (Areas of Interest).

Other analysis metrics include average length of fixations, saccadic speed, saccadic amplitudes, and various transition-based parameters between fixations and/or regions of interest, pupil diameter.

On the basis of the purposes of the present research, the most essential characteristics for understanding cognitive and visual processing behaviour are fixations, which are identified according to a fixed time length and considers the following hypotheses:

- little or no visual processing can be achieved during a saccade;
- smaller eye movements that occur during fixations, such as tremors, drifts, and flicks often mean little in higher-level analyses.

After the identification of fixations, a qualitative analysis of fixations and saccades must be conducted. This analysis consists in the assignment of a category (attention/inattention) to the single element visualized in the fixation. Even if driving attention and inattention are generally associated to mental workload (as afterwards discussed in Activity), the importance of such approach is undiscussed, as it is one of the most effective methodologies enabling to verify road readability and consistency.

In general terms, the categorization of “target points” includes the following elements:

- Attentional target points: road, circulating vehicles and pedestrians, vehicle mirrors, road signs and markings, etc.
- Inattentional target points: environment, internal vehicle, passengers, advertising signs and billboards, etc.

3.1.2 The role of road signs

"Reading traffic sign information correctly is crucial. It helps the transport operator (car driver, pilot or train driver) to anticipate future situations, make decisions, and start to carry out appropriate motor responses" (Castro, Horberry, & Tornay, 2004).

The reported citation explains with simple words a crucial ergonomics principle, called "priming effect". In general terms, this phenomenon consists in anticipating useful information about upcoming issues. Declining the question to the driving activity, it has been proven that being warned beforehand about something enables people to react more quickly, inducing more correct driving behaviour (Charlton, 2006; Crundall & Underwood, 2001). The principal elements implementing that "priming effect" in roadways are undoubtedly temporary and permanent signage, who aims to actively protect vehicles, passengers, workers and site equipment.

In the view of foregoing, studies concerning the quantification of signage visualization are as much crucial and cannot ignore the relationship occurring between signs and the driver (consider, for example, the effect that driver familiarity with the road may have on the capacity to recall signs).

In this sense, also considering that literature refers to a series of principles (conspicuity, size and position of devices in relation to the required decisions or actions, visibility, comprehensibility, credibility) all differently considered in previous studies, the fundamental questions to whom answers are searched within those studies are the following:

- Do road users refer to road signs?
- Are road signs effective?

One of the first studies - conducted in 1966 by Johansson and Rumar - attempted to answer the above mentioned questions by trying to find the number of road signs recorded by a driver during the course of a car journey under the most favorable conditions and to find to what degree road signs act as signals, e.g. affect driving behavior. Although some participants of the study recorded 90% of the signs, the authors conclude that drivers cannot be expected - even under optimal condition (e.g. no distraction) to record every sign (Johansson and Rumar, 1966). Hereafter, another study reported that drivers are able to recall most of the traffic signs they pass if they are motivated enough (Summala and Näätänen, 1974), whereas others report that memory for signs is very often poor (Fisher, 1992; Costa et al., 2014).

According to Martens, the former requirement is that road signs have to be conspicuous and the latter requires that the road sign's message needs to be significant (Martens, 2000). Furthermore, it has been argued that driving behavior is based on expectations and consequently, road signs should be placed where the road users expect them (Borowsky, Shinar, & Parmet, 2008). An interesting study by Viviani and Edquist the ratio between measurements of the maximum distance from which a driver can read the sign and the distance from which he/she actually begins to read it, provide indication of proper location of traffic signs (Viviani, 1990; Edquist, 2011).

Another related problem to be solved is: how can road signs affect drivers who are not actively searching for them? This is especially important for road users in familiar environments. Johannson and Rumar (1966), Martens and Fox (2007) demonstrated that drivers on familiar roads do perceive new road signs but don't respond to them. As road users are likely to oversee important elements of a traffic scene, important elements have to be designed in a way that considers human limitations, captures the road users' attention and are in accordance with the road users' expectations (e.g. place a road sign where road users would expect it).

3.2 Objectives and methods

In the light of the several aspects concerning effectiveness of road signage above described, the objective of the here-presented experimental activity is to understand and quantify the drivers’ visual behaviour towards specific signs proper of the road environment.

A crucial type of signage is represented by workzone signage and, accordingly to the following considerations, has been object of an experimental study aimed to investigate their capability to be visible and easily readable to drivers, as afterwards discussed:

- road work zones are unsafe locations as they disrupt the drivers' expectations about the road geometry, meaning that they have to make sudden adjustments to their driving speed.
- the risk of accidents and severity of collisions increases noticeably in road work zones, compared to the area's pre-work conditions (Yang et al., 2015);
- roadwork signs are supposed to cover an important role in the passive protection of vehicles, passengers, workers and site equipment.

Vice versa, a singular signage category is represented by advertisements signs and, accordingly to the following considerations, has been object of an experimental study as afterwards discussed:

- Hughes and Cole (1986) and Castro and Horberry (2004) have demonstrated that drivers pay more attention to advertisement than to road signs.
- Consequently, many drivers don't perceive or respond to road signs that are in close proximity to an advertisement. When traffic signs are erected, it should be ensured that they are (a) not close to any possible sources of distraction and (b) placed where drivers would expect them.

According to the premise, two experimental studies were conducted. Drivers who participated to both experimental tests underwent several signs, respectively work zone signs and advertising signs, and their visual behaviour was recorded by means of the eye tracking equipment together with the recording of driving performance in terms of kinematic variables by means of a GPS device (e.g. distance of fixation). Fixation rate, length of fixations, fixation distance and driving speed were assessed in both tests.

3.3 Outcomes

Regarding workzone signs, the visual behaviour analyses revealed that drivers glanced at both the temporary and the permanent signs along the sites with a similar 40% frequency. About the single/multiple temporary signs, isolated single signs in work zones caught more attention by the drivers (in terms of both frequency and average duration of the fixations) than a sequence of signs along a work zone. The statistical analysis shows that drivers’ familiarity of the route did not influence road signs fixation frequency, as did instead the vehicle speed. Visible workers activity on the work zone slightly anticipated the distance of first-fixation to the road signs, probably because the presence of dynamic elements on the visual scene increase the conspicuity and detectability of the work zone, but with no influence on speed. Age and poor expertise were predictors for higher speed reductions, but not for fixation rates. Overall, the outcome that had a direct implication for road safety and for future road design is the comparison between the fixation distance and the correspondent stopping distance. The sight distance was frequently lower than stopping distance, revealing inadequate effectiveness of signage positioning.

Regarding advertising signs, as previous research has mainly investigated driver's visual attention to billboards, which represents only one category of advertising signs, in this study, driver's visual attention was assessed in a semi-naturalistic driving setting for six categories of roadside advertising signs: vendor signs, billboards, movable display boards, single and multiple commercial directional signs, and gas price LED displays. Additionally, the role of clearance from the road, elevation, height, width, surface, number and size of characters, total number of characters, side of the road (driving side, opposite side), context (rural, urban), were considered: fixation rate was also significantly influenced by clearance from the road and number of characters.

Indeed, the 24% of the roadside advertising signs were fixated and fixation rate was significantly influenced by sign category. When assessing the distracting potential of an advertising sign (i.e., fixation rate and fixation duration) the critical factors that emerged were the clearance from the road and the amount of text included in the advertising sign. Distraction increases when the advertising sign is placed near the road. The text included in the advertisement induce the driver to read, an activity that requires much more time than capturing the graphical content.

In consideration of the outcomes of the twin activities, the following considerations and recommendations may be outlined:

- Road signage is, generally, poorly considered by drivers;
- As for the revealed inefficiency of work zone signals, the need of improved regulations is envisaged;
- Advertising signs could have a significant distracting potential. Vendor signs, in particular, tend to be more frequent than billboards, and in many cases their size, visual complexity, and textual content is higher, determining a serious distraction source for drivers. In addition, regulations should focus also on vendor signs, commercial directional signs, gas price LED displays, and movable display boards since all compete with the driver's attention undermining traffic safety.
- The visual complexity of roadsides could contribute to attract the driver's attention, diverting it from the driving and resulting in safety issues that should be assessed and controlled.

Besides the increased awareness gained with this study, a limitation is envisaged in the reduced sample dimension which is expected to be increased in future experimentations.

4. Activity II: Experimental study on self-explaining road design

4.1 Introduction

For the evaluation of the influence of road design on road safety, it is important to consider the dual role to which the infrastructure is called to fulfil:

- “Active” role, as well-designed and properly maintained roads can reduce the probability of road traffic accidents;
- “Passive” role, as roads can reduce the severity of accidents that do happen.

Respectively to those roles, find place two concepts, “self-explaining roads” and “forgiving roads”, whose difference has been clearly defined by CEDR (La Torre, 2013). In this section, only the former principle, Self-Explaining roads, will be analysed.

4.1.1 Self-explaining roads

Anticipating drivers’ expectations is one strategy to provide for human information processing limitations. To give some examples of drivers’ short-term expectancies it is possible to consider the following:

1. After driving a few miles on a gently winding roadway, upcoming curves will continue to be gentle;
2. After traveling at a relatively high speed for some considerable distance, drivers expect the road ahead will be designed to accommodate the same speed.

In relation to this has been described the principle of Self-Explaining Roads (SER), which refer to those roads characterized by a precise geometric identity that can suggest to the driver the most correct and safe driving style. Since the aim is to provide road users with road infrastructures that are safe and operationally efficient, the needs and limitations of road design, road signage and users need to properly integrate between them.

Theeuwes and Godthelp sustain that are safe only the roads that are self-explaining: this means that there users know how to behave because they can understand the infrastructure in the way it was designed and also because the road has features that are in line with their expectations (Theeuwes and Godthelp, 1992). The fact that road "explain itself" is a concept only related to design characteristics, not on external agents such as signals and signals for traffic regulation. Design characteristics, in fact, allow to maximize the amount of information coming from the environment and useful to the user for driving.

The self-explaining roads approach exploits roads characteristics focusing on three key principles to influence drivers: functionality, homogeneity and predictability. Homogeneity is guaranteed by the creation of well-defined road categories that prevent large differences in speed between vehicles, direction and flow between each hierarchical level. Predictability, instead, means that the appearance and perception that each road category communicate must be characteristic and exclusive, so that the user's behavior is confident and comfortable. In this way, depending on the situation, an individual should be able to identify road type.

By creating road categories and designing roads in a distinctive way for each category, users are able to be automatically informed about the context in which driving activity takes place and about the appropriate rules or behaviors to be followed (speed limits, the possibility of overtaking, presence of tight turns). Additionally, with categorization an individual is brought to act from experience: when, for example, he understands that he is on a motorway, he will automatically know how to behave (in addition to the rules) also according to his personal history of driving on motorways, with precautions, greater or lesser alert, personal experience. To sum up, a self-explaining road must have the following characteristics:

- exclusive elements and different behavior in each road category;
- behavior suggested strictly by the physical elements of the road;
- crossings, sections, and curves highlighted differently in each category;
- no quick transition from one category to another, but the change must be clearly indicated and perceived;

- indications about the name of the road category and its behaviour to be followed (standards);
- good night visibility of the elements;
- at last, the road geometry, who must suggest the most appropriate speed for the road stretch.

These needs have been improved over the years by the evolution of road engineering, the renewed design study and traffic calming techniques. Lately, focus is given to those measures capable of psychologically influence the user in terms of sensory and cognitive perception.

SER approach is related with the principle of road consistency. Design consistency is defined as the relationship between the geometric characteristics of a highway and those conditions the driver expects to encounter. When the design is consistent with what the driver expects to find, the highway is also consistent. This reduces the possibility of driving errors and unsafe manoeuvring (Castro et al. 2008). The design consistency features mainly concern width of carriageway, road markings, signing, and use of street lighting.

In addition, self-explanation is enhanced when a design is strongly based on an affordance analysis of its components. Affordance, a concept firstly introduced by Gibson (1979), indicates the physical properties of an object, place or situation that suggest to the user how to manipulate it properly. Affordance, therefore, indicates the self-explication in each object or context. It can be that an object has a good affordance, when it is immediate and intuitive to use it; vice versa, when being in front of an object does not automatically understand how to use it, it is said that this object is characterized by poor affordance. This idea can be adapted to the road allowing to consider that the lack of vision of a road sign (see *Chapter 3*), the ambiguity of a certain intersection and the poor visibility, are all indicators of little or no affordance.

Road elements can function as affordances that serve as built-in instructions that can guide driving behaviour, either implicitly or explicitly (Walker, Stanton, & Chowdhury, 2013; Weller et al., 2008). They include road markings, delineated lane width, roadside objects, pavement evenness, but also more intrinsic properties such as the geometry of road intersections (Charlton, 2007; Elliot, Mccoll, & Kennedy, 2003; Weller et al., 2008). For example, narrow lanes and street vegetation close to road shoulders have been shown to reduce vehicle speed by reducing the perceived road width (Ewing & Dumbaugh, 2009).

To sum up, it is therefore important that in the design continuum the ambition is to create road infrastructures that have adequate affordance, with the aim of creating road routes that are easy to understand and to use.

4.1.2 Intersections

The role of intersections in the self-explaining road design is fundamental and hereby discussed.

A road intersection is defined as the area where two or more roads intersect, allowing a partial or total exchange of vehicular currents through devices and equipment designed to limit mutual interference for vehicles in transit.
Among the simpler intersections find place the T-junctions, who consist of a simple confluence between two roads that generate a cross of contrasting flows. The main aspects which it is necessary to address an ergonomics assessment are: arrangement of arms, optimization of visibility conditions, tricks for left and right turn maneuvers, measures in favor of weak users, traffic lights.
According to FHWA, intersections are one of the most complex traffic situations that road users encounter. The intersections represent singular points of the road system and can constitute an element of criticality both for the quality of the circulation and the phenomena of co-management that can arise, and for the increase of accidents generally associated with them.
Intersections represent a disruption in the driving task, generating a variation in the required effort. Their characteristics determine the magnitude of this variation: for example, according to Werneke and Vollrath (2012) and in reference to the contents of *Paragraph 3.1.1,* some intersection characteristics have effects on how drivers allocate their visual attention.

At a general level, the envisaged criteria for the good design of the intersections are (the first three points relate to the concept of self-explaining design):

- ease of understanding of the geometry;
- intuitiveness of the maneuvers;
- adequate affordance for all types of users;
- good conditions of visibility both day and night;
- maintenance of a speed level the most homogeneous possible;
- few points of conflict;
- little chance of colliding with obstacles along the way;
- sharp angles in the nodes in such a way as to avoid the dangerous collisions on the sides that would result from 90° angles.

As for the arrangement of the arms at the intersection, who belongs to the geometry criteria, it is to be considered the angle at which the segments meet. The most advantageous solution would be to intersect the road segments at an angle of 90 degrees, since it guarantees the following possibilities: minimizing the area of the intersection itself, since more acute corners would create an area of longer intersection, increasing the travel time of the crossing area and thus the permanence of vehicles in a dangerous area; improves visibility conditions because in a right-angle intersection you have a good view in all directions, while a more sharp angle would ensure an excellent vision on one side but a very reduced one on the opposite side. In order to create a proper arrangement of the arms at an intersection it is advisable to maintain a section of at least 20 m at the intersection, which as mentioned must have an angle as close to 90 degrees as possible. If this is not possible, for example, because of the territorial conformation, the angle must be at least between 75 and 105 degrees.

The way the arms are disclosed can also influence the perception of precedence of one section of road over another: a road with precedence should be straight, or at least present less structural variations than its subordinate, also to meet the perceptual rules according to our cognitive system.

4.2 Objectives and methods

An observational field study on the effects of self-explaining roads on behaviour has been attempted. The presented study aims to verify the theory that affordance, the self-explanation inherent in every object or context, is the cause of the quality of the interaction between the object/context and the user. If it is adequate and clearly expresses the functions of the object the interaction will be valid and free of errors, vice versa the user's behavior will be inadequate and counterproductive.

In the carried-out activity the affordance is tested by associating road linearity to the perception of priority in unsignalized T-intersections, thus investigating priority perception according to the self-explaining road approach.

A T-junction is a grade three-way intersection between three road segments (arms) where two arms belong to a straight road. Authors focused on T-junctions because, when they are controlled by a yield or stop sign, priority could be assigned according to two very different and asymmetric schemes: in one case priority could be assigned to the vehicles approaching from the straight through road (priority-to-straight-arm condition) and in the other case priority is assigned to vehicles approaching from the intersecting arms (priority-to-intersecting-arm condition).

In particular, authors start from the assumption that having to yield along the straight road of the junction would lead to unsafe behaviors, whereas having to yield at the intersecting road would prompt a much safer behavior (i.e., speed reduction, enhanced visual attention to the intersection). A schematic simplification representing the not-auto-explicative priority scheme of a T-junction is shown in Figure 4.

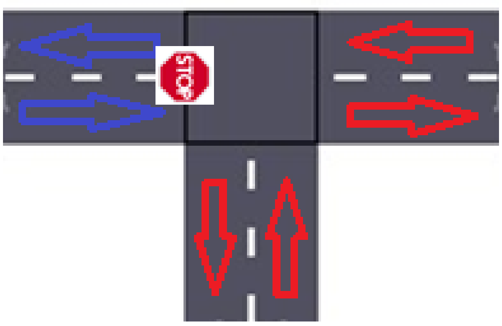

Figure 4. Scheme of an intersection with a not-autoesplicative priority rule.

To the contrary, the priority-to intersecting-arm condition mismatches suggestions inducing a perception of having priority that would clash with the actual need to yield. Also, is supposed that this mismatch would result in a driver's unsafe behavior when approaching the intersection.

In the conducted analysis the type of road design, self-explained or not (control), was established as an independent variable, while the dependent quantitative variable is the number of vehicles that perform, at the analysed intersection, a certain type of behavior, correct (safe) or incorrect (unsafe). Two experimental roads (not-self-explaining, used as experimental intersections) and two corresponding roads in a similar context (self-explaining, used as "monitoring" intersections) were examined (an example is shown in Figure 5). Chosen samples of drivers were analysed on video and their behaviour, in terms of yielding/decelerating or not, was recorded.

In the two analyses, yielding behavior was assessed together with approaching speed and gaze behavior towards the critical areas of the intersection.

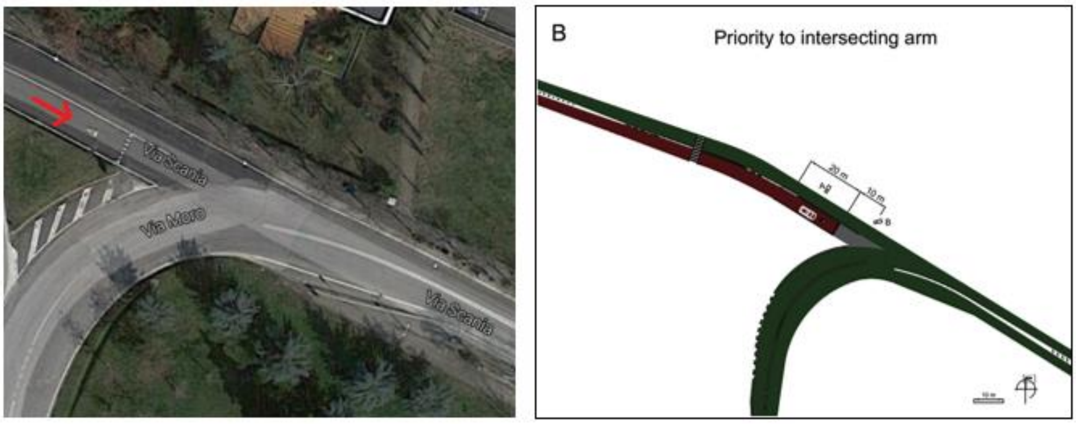

Figure 5. One of the experimental intersections analysed in the paper showing the priority given to the intersecting arm (via Moro).

4.3 Outcomes

The statistical analysis reveals a significant difference both in the explicative intersection and the monitoring one, thus demonstrating the causality between the above-mentioned variables. On the non-explicative intersections results a much greater number of vehicles that implements transgressive behaviors with respect to the laws in force because they are misled by an intuitive design.

More in particular, the results show a significant speed reduction and an increase of driver's visual inspection to the intersection area in the priority-to-straight-arm condition in comparison to the priority-to-intersecting-arm condition. The eye movement analysis showed that total fixation time towards the intersection critical area and horizontal eye movements were significantly higher in the priority-to-straight arm condition, revealing drivers' uncertainty.

The results emphasize the importance of considering perceptual affordances and expectations for priority in intersection design to increase drivers' compliance to yielding rules. This is only one of the potential affordances that could influence the perception of priority, and future investigations should focus on other affordances.

For example, a road with a larger cross-section could be perceived as having priority over a road with narrow cross section, or, in case the intersection arms are not planar, a road that is more elevated could be perceived as having priority over an intersecting road that is less elevated. The approaching curvature and geometry of a road to the intersection could also play a significant role for the affordance of priority. For example, if a road is connected to the intersection by a straight line the perception of having priority could be significantly higher than a condition in which a road is connected to an intersection with a curve. In roundabouts, for example, road users enter with a curvilinear trajectory and this contributes to a speed reduction and hence to an increased safety in comparison to standard signalized intersections (Elvik, 2003; Gross et al., 2013; Jensen, 2013). Also, Stephens et al. (2017) proposed two intersection designs that succeeded in speed reduction eliminating the possibility for a road user to cross the intersection along a straight trajectory, introducing islands that induced curvilinear trajectories.

In conclusion, self-explaining roads may be considered as the greatest result of ergonomics applied to the road environment, and the study of affordance the most suitable method for every situation with interventions that make the most of the perception of the driver, manipulating or even tricking him in following driving rules and adopting the most suitable driving style.

5. Activity III: Experimental study on drivers' workload

5.1 Introduction

5.1.1 The concept of "workload"

The mental workload is an important and central construct in ergonomics and human factor research. At a general level, it represents a complex construct that is assumed to be reflective of an individual's level of attentional engagement and mental effort (Wickens, 1984). Even if nowadays still not exists a universal definition of workload, many have been given during the last decades:

- "Mental workload refers to the portion of operator information processing capacity or resources that is actually required to meet system demands" (Eggemeier et al., 1991);
- "Mental workload is a hypothetical construct that describes the extent to which the cognitive resources required to perform a task have been actively engaged by the operator" (Gopher and Donchin, 1986);
- "The reason to specify and evaluating the mental workload is to quantify the mental cost involved during task performance "in order to predict operator and system performance" (Cain, 2007).

Such definitions show that the mental workload may not be a unitary concept because it is the result of different interacting aspects.

In a more pragmatic sense, the word "workload" identifies the load of the activity brain which a person is subject while performing an action; its measurement, indeed, essentially represents the quantification of mental activity resulting from performing a task and, in addition, the different actions conducted at the same time.

Although several studies about the workload we don't have a complete knowledge of this subject yet. The complexity of workload estimation relies in the fact that the WL depends on many factors, such as circumstances and users' skills, behaviors, and perception (Hart and Staveland, 1988). Intuitively, it is possible to affirm that more difficult is the task to perform, more workload is needed. When the workload is too high, some users might elaborate the information with significant delay or may not react at all. On the contrary, when the workload is too low the user tends to be boring and may make some errors. Several empirical investigations have indicated that performance declines at either extreme of the workload demand continuum, that is when the event rate is excessively high as well as when the event rate is extremely low (Borghini et al., 2012). The presumed general rule, thus, is that it is important to preserve a good level of a user's mental workload, avoiding mental under- or overload state, with the aim to maintain an optimal level of performance and reducing the probability of errors commission (Parasuraman and Hancock, 2001)

The most famous relation that connects the workload and the performance level is described by the Yerkes-Dodson's Law in 1908 (Figure 6, Yerkes and Dodson 1908). From their study is possible to consider that the performance increases with the mental request until when the workload becomes too high that leads to a reduction of the performance.

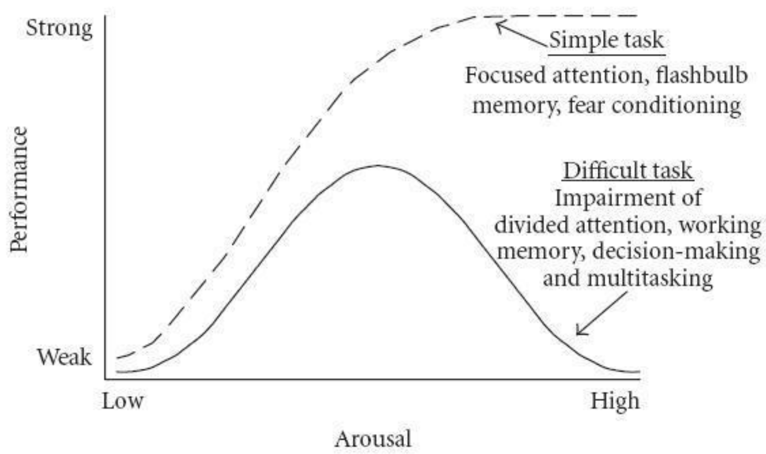

Figure 6. Yerkes-Dodson's Law describing the relationship between arousal and performance in a general context.

Another relation between task demand and task performance has been described by Meister (Figure 7), who defined three regions, region A, B and C:

- Region A is described as low operator workload with high performance. An increase in demands does not lead to performance decrements;

- In region B the level of performance declines with increased task demands;
- In region C extreme levels of load have diminished performance to a minimum level and performance remains at this minimum level with further increases in demand.

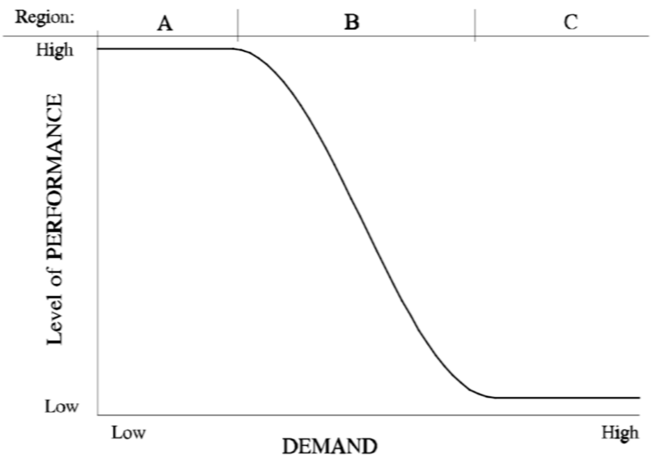

Figure 7. Meister regions describing the relationship between task demand and performance in a general context.

According to this model, a primary-task workload measure, i.e. a measure of performance, will only be sensitive to variations in levels of workload (region B). In region A, performance remains constant and independent from variations in demand, while in region C performance will stay at a minimum level, independently from demand. Extreme levels of load resulting in overload can be situated in the C-region, but it is not clear where the domain of underload is.

McCracken and Aldrich give a proper definition of the resources involved in each task (McCracken and Aldrich, 1984):

- Visual and Auditory resources (external inputs);
- Cognitive resource (as the component that describes the level of information required);
- Psychomotor resource (as the component that describes the real action that the user has to perform in order to complete the task).

One of the most relevant theorems about the workload is the Multiple Resource Theory (MRT) by Wickens. The MRT proposes that the subject does not have one single information processing source that can be used, but several different groups of resources that can be consulted simultaneously.

According once again to Wickens, workload and performance are described by a model where only one dimension of mental workload is displayed and described among 6 regions (Figure 8):

- In region D (D for deactivation) the user's state is affected;
- In region A2 performance is optimal, the operator can easily cope with the task requirements and reach a (self-set) adequate level of performance;
- In the regions A1 and A3 performance remains unaffected but the operator has to exert effort to preserve an undisturbed performance level;
- In region B this is no longer possible and performance declines, while in region C performance is at a minimum level: the operator is overloaded.

In this model what is depicted denotes the overall or sum relation between demand, workload and performance. Then, auditory task demands, visual task demands, and central demands do not necessarily have to be in the same region, which is in accord with Wickens' MRT theory (Wickens, 1984).

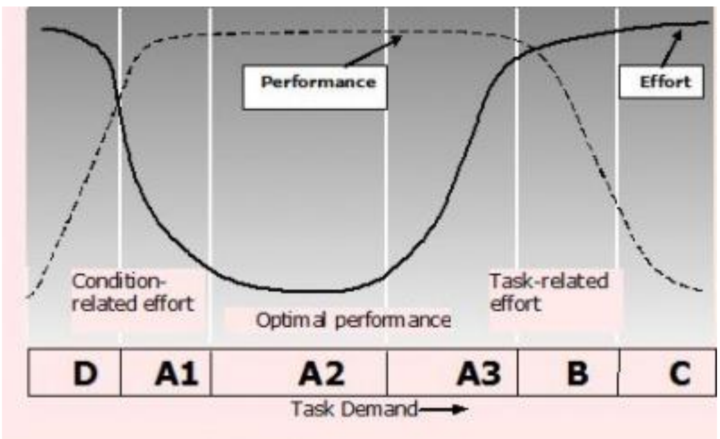

Figure 8. Wickens' Multiple Resource Theory (MRT) describing performance and effort in relation to task demand.

Wickens' theory considers the performance decrement as a lack of these different sources and describes subjects as having limited capability for processing information. In this view, excessive workload caused by a task using the same resource can cause problems and result in errors or slower task performance. For example, if a second task is being performed at the same time of the first and that makes demands on the same component, the result may be excess workload.

5.1.2 Workload and the driving task

The focus of this paragraph is to consider the application of the general theories to the driving task. For our purpose, a good definition of workload regarding the road environment was given by Messer (1980): "Driver workload is defined as the time rate at which drivers must perform a given amount of driving tasks that increases with the increase of the complexity of highway geometric features". Additionally, Knowles defined workload as consisting of the answer to two questions: "How much attention is required?" and "How well will the operator be able to perform additional tasks?". The latter is a definition particularly fitting to driving environment, given that the activity consists of many overlapping tasks, each requiring a portion of the drivers' attention.

Several are the studies conducted in last decades in relation to the concept of mental WL and to the assumption that the road environment affects the mental demand. To this regard is possible to consider that:

- different studies strictly correlate the total amount of brain activity to the general complexity of the route;
- WL represents an approach to measure or rate the road design, especially in terms of consistency (e.g. evaluation of a LOC – Level of Consistency) (Knowles, 1963, Messer, 1980). Also, Kanellaidis demonstrated that design consistency (see also *Paragraph 4.1.1*) is indirectly associated with how drivers maneuver geometric features, while drivers WL is directly related to it. Messer indicated that driver WL increases with reductions in sight distance and increasing complexity of geometric features.

The estimation of drivers' WL is generally associated also to drivers' attention, who is divided into three tasks (control, guidance and navigation) and is not fully within drivers' conscious control, being in general, highly automated tasks. Attentional levels are in general associated to the principles of mental overload and mental underload (Brookhuis & De Waard, 2010; De Waard & Brookhuis, 1991): both results avoidable as the former is leading to distraction, the latter to state of drowsiness. In addition, the following considerations may be carried out:

- Mental overload occurs when drivers have to attend to more one than one task. Various studies have for example demonstrated that a secondary task such as mobile phone usage or listening to the radio causes a decrease in brake reaction time (Brookhuis et al. 1991, Irwin et al. 2000, Consiglio et al. 2003). Roadway design considerations for reducing driver workload are:
    - Presenting information in a consistent manner to maintain appropriate workload;
    - Presenting information sequentially, rather than all at once, for each of the control, guidance, and navigation tasks; and,
    - Providing clues to help drivers prioritize the most important information to assist them in reducing their workload by shedding extraneous tasks.
- Mental underload occurs in the case of a lack of mental demands can, which can be as detrimental to performance as overload (Branscome & Grynovicki, 2007). Mental underload is especially likely to occur when the driving environment is predictable, and it is associated with both predictability and monotony. In that sense, has been described the phenomenon of the "highway hypnosis" (Tejero & Chóliz, 2002), firstly introduced by Williams (1963), postulating that prolonged driving in a monotonous environment leads to a trance like state who may also lead to drowsiness.

5.1.3 Measurement methods

Workload quantification, according to mental workload theories, assumes the following hypotheses:

- people have a limited cognitive and attentional capacity;
- different tasks will require different amounts (and perhaps different types) of processing resources;
- two individuals might be able to perform a given task equally well, but differently in terms of brain activation (Baldwin, 2003; Wickens, 1984).

Over the years, many different measurement methods have been developed according to the following three macro-categories:

a. *Physiological measurements*: are related to assessing the function of major human organ systems and provide useful information regarding how the human body responses in different external conditions. The human body's responses seem to be the most accurate measurement methods. Due to the fact that some physiological reactions may be influenced by the previous driver's experience each measure could be related to the physical and mental status of the user.

   Physiological measurements can be subdivided into five different areas:
   - heart's activity;
   - respiratory activity;
   - visual activity;
   - speech measure;
   - mental activity.

   Although these measurements require the subject to wear intrusive sensors, they provide an objective assessment for drivers' attention level e.g. the variability of heart rate, breath rate or pupil diameter.

b. *Subjective assessments.* The most common techniques are scales for the subjective mental workload. Their examples are:
   - the subjective workload assessment technique (SWAT);
   - the rating scale mental effort (RSME);
   - the NASA task load index (NASA-TLX).

The most commonly used tool to conduct the subjective evaluation of workload is the NASA-TLX. It consists in a multidimensional assessment questionnaire that rates perceived workload in order to assess a task, system, or team's effectiveness or other aspects of performance. It was developed in the 80' by the Human Performance Group at NASA's Ames Research Centre over a three-year development cycle that included more than 40 laboratory simulations (Hart and Staveland, 1988). It has been used in a variety of domains, including aviation, healthcare and other complex socio-technical domains. The total workload score is calculated by the combination of six factors:

- Mental Demand: How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
- Physical Demand: How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
- Temporal Demand: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
- Performance: How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
- Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?
- Frustration: How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Subjects have to rate each factor (from 0 to 100) and then factors are compared in pairs to each other, and the number of times each factor is chosen is the weight by which the previous rates is multiplied. The linear combination of the weighted rates is then divided by 15 (number of total comparisons) and the NASA-TLX total score (value from 0 to 100) is provided. The test, if submitted immediately after the completion of the task, allow to capture accurate workload perception.

5.2 Objectives and methods

In the experimentation object of this study, it has been investigated the possibility to evaluate the mental workload experienced by a car driver, by means of his/her EEG activity in real driving settings. The rationale was to use the EEG measures, instead of other neurophysiological or subjective measures, because of their specific suitability in objectively assessing human mental states and to test a machine learning algorithm (see also Chapter 6) for the calculation of a EEG-based Workload index, already validated for other operational settings (e.g. Aviation; Aricò, Borghini, Di Flumeri, Colosimo, Pozzi, et al., 2016).

The real driving context represents the first added value of this study because, so far, all the works related to workload investigation by using EEG have been performed in simulator, or in poor realistic settings. It is important to prove the effectiveness of EEG-based metrics in real contexts, since it has been proven that same experimental tasks are perceived differently, in terms of mental workload, if performed in a simulator or in real environment (De Winter et al., 2014).

The second added value of the work is the possibility of validation of a brand-new machine-learning approach for EEG based WL index, also through the integration with:

- Eye-Tracking technology, to provide evidence of the complementarity of the obtained insights and the possibility to make considerations on visual behaviour;
- GPS device, to evaluate driving performance in terms position, speed, acceleration;
- A further analysis based on the NASA-TLX questionnaire, to assess if the neurophysiological measures were consistent with the perception of the workload.

The study, indeed, explored the potential of integrating the above-mentioned new methodologies with traditional approaches in order to enhance and extent research on drivers' behaviors and road safety.

5.3 Outcomes

The experimental investigation followed by the analysis of its results confirmed the validation of the envisaged methodology. The results demonstrated the reliability and effectiveness of human EEG signals for measuring drivers' mental workload and the ability to provide insights about human mind while dealing with tasks that are difficult or even impossible to obtain by using traditional approaches.

The electroencephalographic technique resulted an appropriate solution to evaluate the mental workload in realistic and operational settings and capable to be integrated in passive BCI systems.

The workload scores (WL score) have been used to evaluate the impact of different factors, that are the road complexity and the traffic as well as specific events along the driving experience and the outcome consists in their consistency.

Nevertheless, the main limit that affects the present study is the algorithm calibration with data coming from the task itself and recorded in very similar conditions. From one side, it could be argued that in everyday life context such a calibration would be unfeasible; from the other side it could be argued that the proposed algorithm is not classifying the targeted mental state, i.e., mental workload, but only two conditions that are very similar. Regarding the calibration, actually it is one of the main still open issues in transferring machine learning approaches from research to applied field: several solutions have been explored, such as cross-task calibration or employment of unsupervised algorithms, but the problem is still open and needs further investigation. In conclusion, other than the specific obtained results, the present work breaks new ground for the integration of these new methodologies, i.e., neurophysiological measures, with traditional approaches in order to enhance and extend research on drivers' behaviors and road safety.

6. Activity IV: Analysis of drivers' behaviour with deep learning techniques

6.1 Introduction

With the constant increasing amounts of available data collected with the here-presented experiences, one of the main problems faced is the difficulty and time-wasting issue represented by data analyses. It is evident, indeed, that new approaches of data analysis are a necessary ingredient for improve the future research.

Computers often undertake those problems simpler than human beings and Artificial Intelligence ecosystem is the strategy universally adopted for tackling data analyses.

6.1.1 *Principles of Deep Learning and Neural Networks*

Nowadays Artificial Intelligence, Machine Learning and Deep Learning are easily confused and often seem to be used interchangeably thinking that they refer to the same subject. Actually, they are not quite the same thing and, as a compulsory premise, terminologies are hereafter clarified and graphically represented in Figure 9:

- Artificial Intelligence regards a width area of computer science aims to reproduce the human behavior on the machines;
- Machine Learning is a sub-set of Artificial Intelligence where computer algorithms are used to autonomously learn from data and information. Machine learning provides to computers the ability to learn from data without being explicitly programmed and likely most of us are benefiting from that. A more general definition given by Arthur Samuel is: "Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed" and that is the reason of Machine Learning success over the past twenty years. This self "training" involves feeding huge amounts of data to the algorithm and allowing the algorithm to adjust itself and improve; Neural Network or Artificial Neural Network is a set of algorithms used for modelling the data using graphs of Neurons, capable to solve many kinds of problems. A Neural Network is based on a collection of connected units called Neurons (Figure 11)

which loosely attempt to model the neurons in a biological brain and, according to the neural network structure, the model can act as one or more algorithms working together;

- Deep Learning is a sub-area of Machine Learning using “Deep Neural Networks”, i.e. a neural network consisting of more than one hidden layer.

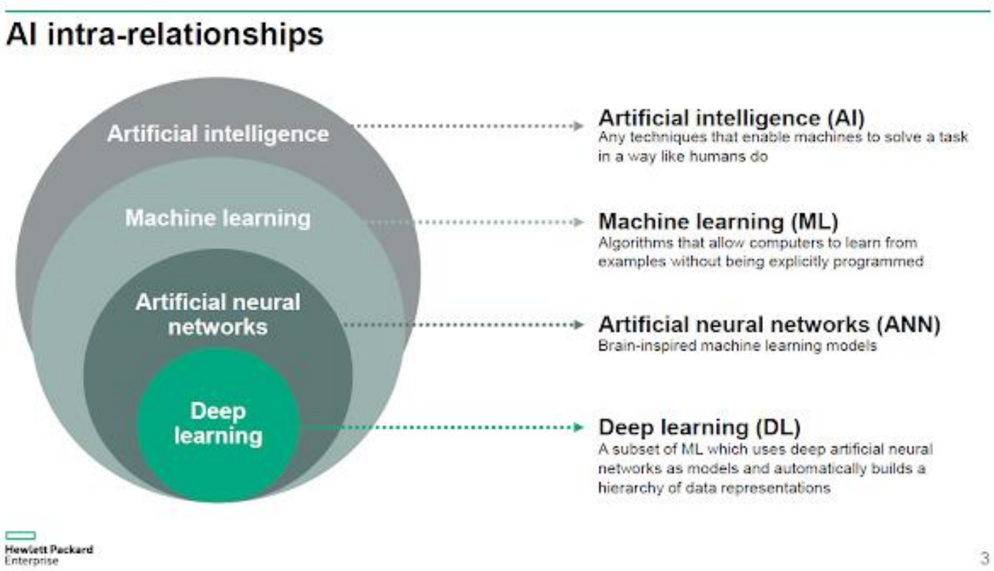

Figure 9. Venn's diagram for Artificial Intelligence, Machine Learning and Deep Learning.

As a common peculiarity, learning algorithms are intended to be computational models of natural learning, or in other words, models that attempt to reproduce the learning process of the brain, whose first mathematical modelling dates back to 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pitts.

Today, the learning algorithms reproducing human performance on complex tasks are nearly identical to the old first ones, and the modifications occurred substantially consist in the approach used to train the Artificial Intelligence systems. Additionally, the development occurred is given by the fact that we can provide these algorithms the resources they need to succeed. With “Big Data”, indeed, learning algorithms have been boosted as they have been provided with a huge amount of data to feed their knowledge, and, additionally, exists computational resources to run many complex models today on the contrary of fifty years ago. Technology improvements (computer even faster, larger memory), in fact, are leading the success of the Artificial Intelligence.

The learning algorithms in machine learning are several and can be addressed to two main typologies, Supervised and Unsupervised Learning (Figure 10), afterwards presented.

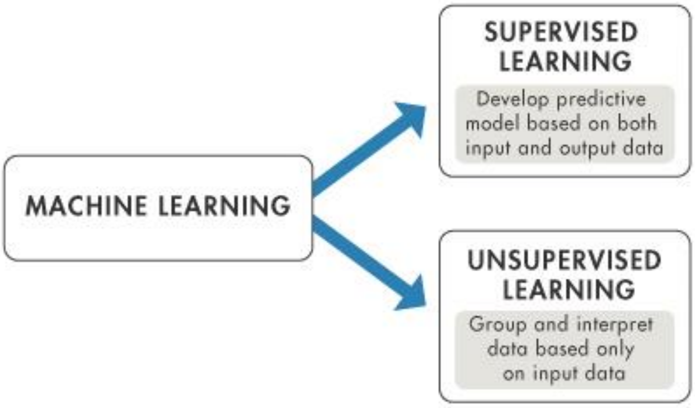

Figure 10. Differentiation between Supervised and Unsupervised learning.

- Supervised Learning:

The most widely used machine-learning methods are supervised learning methods. Supervised learning systems, including spam classifiers of e-mail, face recognizers over images, and medical diagnosis systems for patients, all exemplify the function approximation problem where the training data take the form of a collection of (x, y) pairs and the goal is to produce a prediction y* in response to a query x*. Supervised Learning Algorithms experience a dataset contains features where each example is associated a label (or target). So, roughly speaking, they are able to analyze a dataset and learn how to associate the input *x* to a given output *y* (Godfellow, 2011).

- Unsupervised Learning:

Unsupervised Learning Algorithms experience a dataset containing many features and learn useful properties of the structure of the dataset. They are focused on find out correlations, hidden patterns and insights between the examples of the dataset.

Unsupervised training occurs when you do not provide the expected outputs to the algorithm and usually, when we are going to handle a specific problem, we have a set of many algorithms that we can use to solve it, which algorithm to use it's due the boundary conditions, restriction and to the initial assumptions.

6.1.2 How Neural Networks work

The simplest form of a Neural Network is given by the basic computational unit that, given an input $\boldsymbol{x} = [\ x_1, x_2, \dots , x_n\ ]$ and a term $\boldsymbol{b}$ called Bias Unit (typically set equal to +1), produces an output (Figure 12):

$$\boldsymbol{y} = f\,(\boldsymbol{Wx} + \boldsymbol{b})$$

where $f : R \rightarrow R$ is called "Activation Function".

The Activation Function can be a linear or non-linear function (Linear function, Step function, Hyperbolic Tangent (tanh), Sigmoid function, Rectified Linear Units, Softmax function). Using the Sigmoid Function is obtained exactly the input-output mapping algorithm defined by Logistic Regression Algorithm.

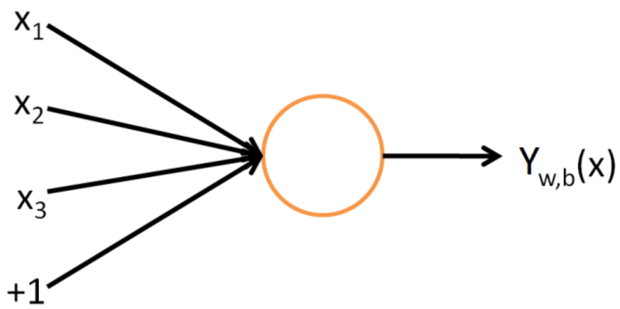

Figure 11. Graphical diagram of a single neuron.

Many simple "neurons" stacked together are called "Layer". Conventionally, in a neural network the first layer is called "Input Layer", the last layer "Output Layer" and the middle layer "Hidden Layer".

A typical example of Neural Network is showed in Figure 123, where n_l is the number of layers we label L1 as *Input layer*, L2 as *Hidden layer* and L3 as *Output layer*.

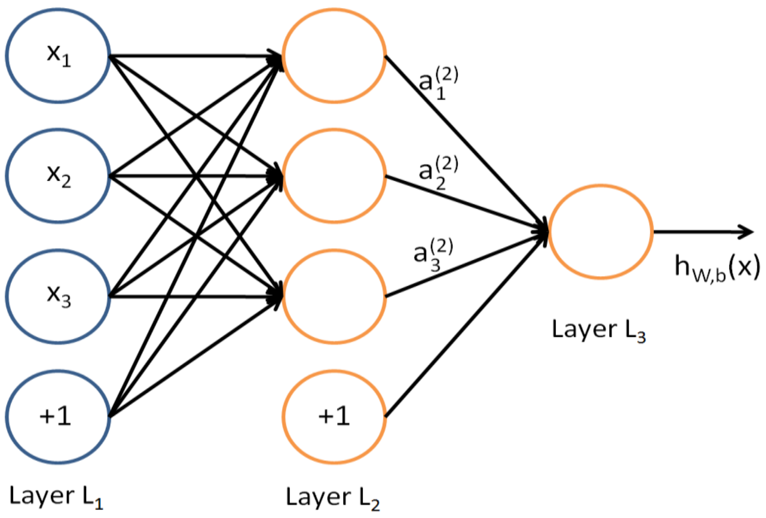

Figure 12. Graphical diagram of a Simple Feed-Forward Neural Network.

The neural network has parameters that describe each connection between the units of the different layers, in the case above:

$$(\boldsymbol{W}, \boldsymbol{b}) = [\ \boldsymbol{W}_{(1)}, \boldsymbol{b}_{(1)}, \boldsymbol{W}_{(2)}, \boldsymbol{b}_{(2)}\ ]$$

Where:

- $\boldsymbol{W(l)}_{ij}$ indicates the weight parameter associated to the connection between unit j in layer l and unit i in layer $l+1$ and
- $\boldsymbol{b(l)}_i$ indicates the Bias associated to the unit i in level l.

In this way the mathematical notation of the model is:

$$a^{(2)} = f(\ \boldsymbol{W}_1 * \boldsymbol{a}^{(1)} + \boldsymbol{b}_1)$$

$$a^{(3)} = f(\ \boldsymbol{W}_2 * \boldsymbol{a}^{(2)} + \boldsymbol{b}_2)$$

More generally, given an input vector $\boldsymbol{x} = \boldsymbol{a}^{(1)}$, it results possible to compute the layer's activations $\mathbf{a}^{(l)}$ given the activation of the previous one :

$$z^{(l+1)} = \boldsymbol{W}_l * \boldsymbol{a}^{(l)} + \boldsymbol{b}_l$$

$$a^{(l+1)} = f(z^{(l+1)})$$

The *architecture* of the model may be changed by simply modified by adding multiple hidden layers to the previous neural network (examples in Figure 13).

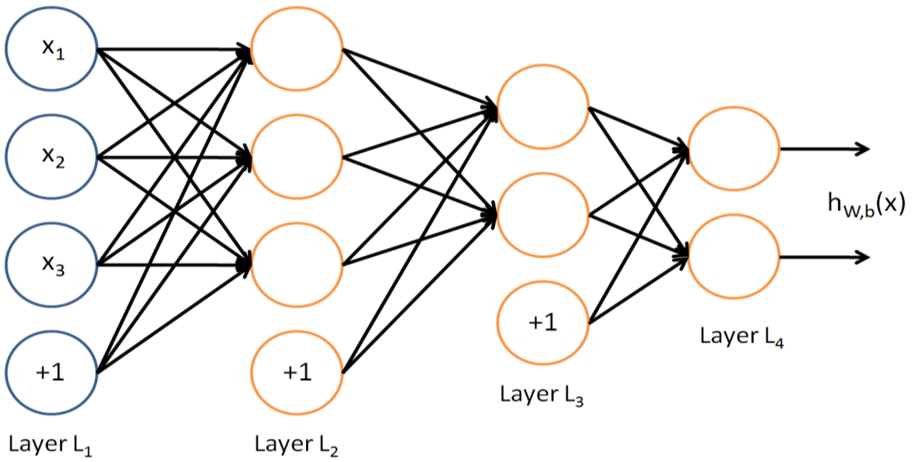

Figure 13. Graphical diagram of a Feed Forward Neural Network with 2 hidden layers.

Since in the described networks there are no backward connections, loops or cycles, this is an example of “Multilayer Perceptrons” or well-known as “FeedForward Neural Network”. Mathematically we can look at a feedforward network as a set of function in which the model describes how those are composed together:

$$f(x) = f_{(3)}(\ f_{(2)}(f_{(1)}(\mathbf{x})))$$

Where each function represents a single layer and the overall length of the chain gives the “Depth” of the model.

More than one hidden layer identifies a neural network as *Deep Neural Network* (Figure 14).

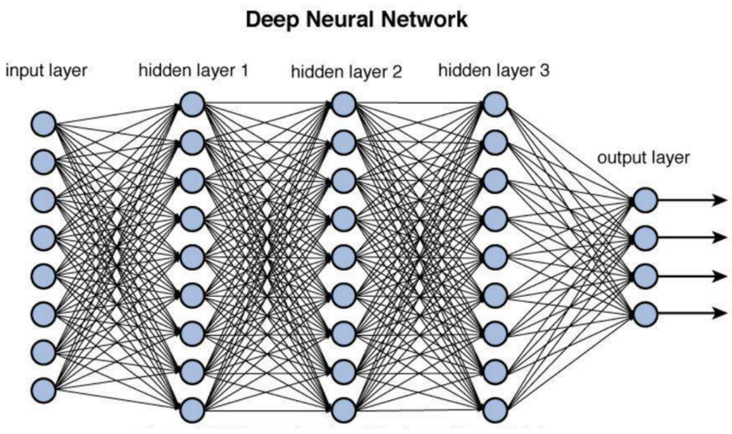

Figure 14. Graphical diagram of a Deep Feed Forward Neural Network.

In practice, Deep Neural Network is used to approach complex problem where a common single-hidden-layer neural network may fail. In deep-learning networks each layer of nodes trains on a distinct set of features based on the previous layer's output, and the further you advance into the neural net, the more complex the features your nodes can recognize.

The DNNs are aimed to overcome lack of the common neural networks in order to handle complex problem, with high-dimensional datasets and billions of parameters that pass through non-linear functions. These networks are in fact capable of discovering latent structures within unlabelled, unstructured data, which is the vast majority of data in the world.

When training on unlabelled data, each node layer in a deep network learns features automatically by repeatedly trying to reconstruct the input from which it draws its samples, attempting to minimize the difference between the network's guesses and the probability distribution of the input data itself (Heaton, 2013).

Restricted Boltzmann machines, Autoencoder and Convolutional neural networks for examples, create so-called reconstructions in this manner.

6.1.3 Autoencoder

An autoencoder is a neural network that is trained to attempt to copy its input to its output. In the middle, it has just one hidden layer ***h***, where its output wants to be the same as the input (Figure 15).

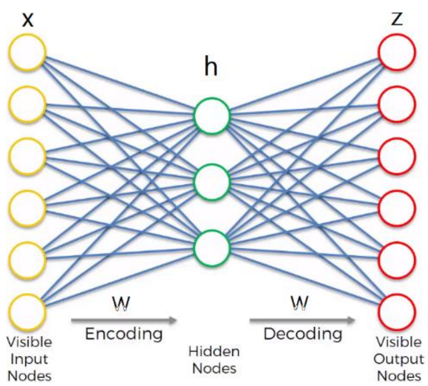

Figure 15. Graphical diagram of an Auto-Encoder.

The autoencoder neural network may be viewed as consisting of two parts:

- Encoder Function: $\boldsymbol{h} = f(\boldsymbol{x})$
- Decoder Function: $\boldsymbol{r} = g(\boldsymbol{h})$

The structure of an autoencoder forces the model to choose which aspects of the input data should be copied. Typically, there is no interest in the output of the decoder; instead, it is interesting to train the autoencoder to perform the input by copying task that will result in ***h*,** taking on useful properties.

The learning process is described simply as minimizing a ***Loss*** (or Cost) ***Function:***

$$L(\boldsymbol{x}, g(f(\boldsymbol{x})))$$

where L is a loss function that penalizes the mean squared error between the input data and the reconstructed same data:

$$MSE = [\, g(f(\boldsymbol{x})) - \boldsymbol{x} \,]^2$$

Improvements in the learning process can be obtained by changing the Cost Function or changing the reconstruction error term of the Cost Function.

The denoising autoencoder is an autoencoder that receives a corrupted data point as input and is trained to predict the original, uncorrupted data point as its output. DAE minimizes:

$$L(\boldsymbol{x}, g(f(\tilde{\boldsymbol{x}})))$$

where $\tilde{\boldsymbol{x}}$ is a copy of $\boldsymbol{x}$ that has been corrupted by some form of noise.

6.1.4 Stacked Autoencoder

Stacked (Denoising) Autoencoders have multiple layers but its training is not same as a multi layered neural Network.

Autoencoders can be stacked to form a deep network by feeding the latent representation autoencoder found on the layer below as input to the current layer. The unsupervised pre-training of such an architecture is done one layer at a time.

Each layer is trained as an autoencoder by minimizing the error in reconstructing its input (which is the output code of the previous layer). Once the first l layers are trained, we can train the l+1-th layer because we can now compute the code or latent representation from the layer in Figure 16:

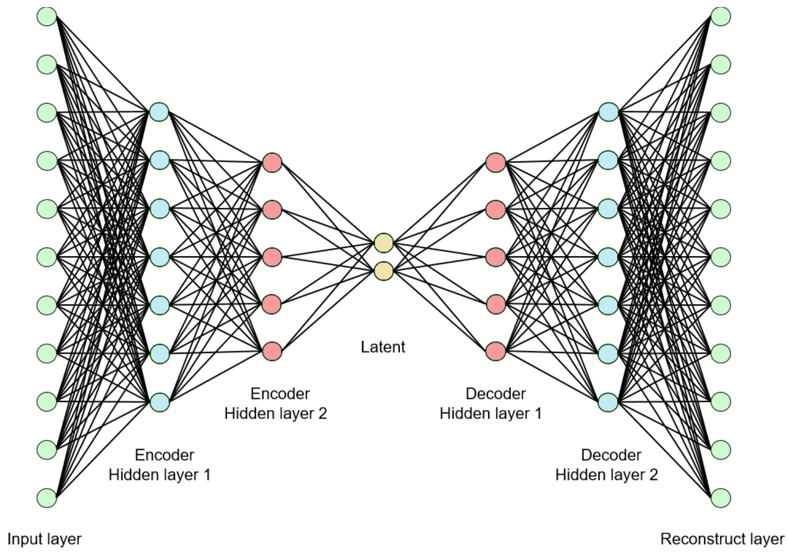

Figure 16. Graphical diagram of a Stacked-Autoencoder.

To produce better results, after the completion of the phase of training, fine-tuning using backpropagation can be used to improve the results by tuning the parameters of all layers are changed at the same time.

6.2 Objectives and methods

The activity mainly takes inspiration from a research by Liu et al. (2014). In order to exploit the large availability of real-time vehicle trajectory data which, in general, are difficult and very costly to treat (see *Paragraph 2.4*) and with limited accessibility to the general research community, the aims of this activity is to apply a machine learning technique for the graphical visualization of latent features extracted by observational driving studies. More specifically, distinctive driving behaviour patterns are searched among kinematic data obtained from sensors connected to the vehicle.

The data were recorded for ten users driving a car for 6 laps on a round circuit route. The obtained dataset thus consisted in a multi-dimensional time series dataset, including driving kinematic parameters recorded by a GPS device. In particular, the considered dimensions were:

- Distance;
- Vertical Speed;
- Longitudinal speed;
- Vertical acceleration;
- Longitudinal acceleration;
- Heading;
- Absolute height;
- Relative height;
- Latitude;
- Longitude;
- Radius of turn;
- Centre line deviation;
- Combo G (ratio of the components of the acceleration).

In many cases, the recorded dimensions of the time-series data were not independent from the others (e.g. lateral acceleration results highly correlated to the radius of turn or to heading). Thus, essential latent features from measured driving behaviour data were extracted and reduced while preserving information.
A Deep Sparse Autoencoder (*Paragraph 6.1.3*52) was chosen to extract hidden features from the driving behaviour data and the Python libraries used are Numpy, Scipy, Scikit-learn and Theano.

In order to visualize the latent features time-series we used a method called "Driving Colour Map" that maps the extracted 3-D hidden features to the Red Green Blue (RGB) colour space. A driving colour map is produced by placing the colour in the corresponding position in the map.

6.3 Outcomes

The experiment shows that the features extraction method based on the DSAE facilitates a visualization of driving behaviour better than conventional methods. From a theoretical point of view, the low-dimensional time series of latent features extracted using DSAE resulted useful for driving behaviour visualization. Feature extraction resulted robust against defects and outliers. This is a direct consequence of the training method used on the DSAE, namely the back-propagation method that minimize the square error between the input data and the reconstructed data.

The research demonstrated also that dataset with high correlated inputs features obtained best results in term of defects reparability and latent features extraction.

By reducing the dimension of feature vectors while retaining the information contained within the time-series data, the proposed approach is able to reduce the computational cost of possible post-processing tasks, such as prediction or classification.

From a practical point of view, considering the impacts on road safety of driver behaviour recognition from large datasets, is possible to evaluate this first attempt as a successfully one, as resulted the capability of the method to recognize both road complexity and external events (as pedestrians) in the majority of cases.

The envisaged study limitation consists on the necessity to involve different categories of experimental variables in order to go beyond the limit of using only one typology of data (kinematic data). In particular, it is expected that physiological drivers' measurements (i.e. oculometry, direct measure of workload) and road conditions, if implemented, would add significance to the graphical outputs.

7. General conclusions

As road safety has become a significant public health issue and it has been found that driver is a contributing factor in the 90% of cases, a sound understanding of driver behaviour and attitudes towards the road infrastructure and environment may help in the development of effective countermeasures and in the definition of new road design principles. For this reason, this dissertation investigated the interaction between the road and the drivers according to human factor principles and assessing the effects that some road-related factors have on driving behaviour and performance, also addressing some of the gaps currently present in the literature on the casual links between drivers' behaviour and road features.

In particular, the Activities I and II investigated the road characteristics (road signs, geometry and readability) relating them with drivers' expectations. The Activity III focused on the driver and the effort spent in particular conditions of driving. The analysis of the relation between driving conditions, effort and performance allowed to validate the methodology used for the measurement of the mental workload resulting an important achievement for the possibility to integrate new indicators, i.e., neurophysiological measures, to traditional approaches in order to enhance and extend research on drivers' behaviors and road safety. At last, the Activity IV consisted of a first attempt to analyse data from semi-naturalistic driving studies with deep learning techniques. This latter result paves the way for more effective and reliable data analyses.

Overall, it is possible to consider that the strengths of the presented work are multiple and represented by:

- A multidisciplinary approach to the investigation of drivers' behaviour;
- The results reliability and significance, founded on the data collection i.e. real experimental driving test, who do not undergo simulator studies limitations;
- A technological innovation upscale, represented by the high use of technologies for the experimental studies.

Considering the increasing research maturity achieved over the doctoral period, the author envisages the following as the ground-breaking results of the overall research. The first is represented by the reliability and effectiveness of the methodology allowing the measurement of a “workload” index (i.e. the mental commitment to which drivers are subjected while driving) with EEG device. Hence, this method enabled to evaluate the impact of different factors, specifically the road complexity, the traffic intensity, and external events (a pedestrian crossing the road and a car entering in the traffic flow).

The second is represented by a novel successful approach consisting in the application of a Deep Learning methodology allowing to extract low-dimensional time series of latent features from multidimensional driving behaviour data using a Deep Sparse AutoEncoder. This revealed useful for obtain graphical visualization of driving behaviour representing a potential tool for driver behaviour recognition from large datasets.

Regarding future research, fine-tuning activities are undoubtedly required (mainly on experimental protocols e.g. sample type and dimension, road typology, etc.) and additional variables might be added (e.g. vehicles’ automation, adaptive or not, which is already underway from the research group).

Even if the research path allowing a sound understanding of the occurring phenomena in driving safety is still long and complex, the achieved findings allow the definition of best practices and guiding principles for the road design, which should no more be considered just a road engineering task but the outcome of a multidisciplinary process.

References

(1) Åberg, L., Larsen, L., Glad, A. & Beilinsson, L. (1997). Observed Vehicle Speed and Drivers' Perceived Speed of Others. Applied Psychology, 46, 287-302.

(2) Aricò, P., G. Borghini, G. Di Flumeri, A. Colosimo, S. Pozzi, F. Babiloni (2016). Chapter 10 - A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks. In: Editor(s): Damien Coyle. Progress in Brain Research, Elsevier, Volume 228, Pages 295-328, https://doi.org/10.1016/bs.pbr.2016.04.021.

(3) Barnard, Yvonne & Utesch, Fabian & Nes, Nicole & Eenink, Rob & Baumann, Martin. (2016). The study design of UDRIVE: The Naturalistic Driving Study across Europe for cars, trucks and scooters. European Transport Research Review. 8. 10.1007/s12544-016-0202-z.

(4) Baldwin, C. (2003). Neuroergonomics of mental workload: New insights from the convergence of brain and behaviour in ergonomics research. Theoretical Issues in Ergonomics Science. 4. 132-141. 10.1080/14639220210159807.

(5) Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F. (2012). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci. Biobehav. Rev. 44, 58–75. doi:10.1016/j.neubiorev.2012.10.003.

(6) Borowsky A., Shinar D., & Parmet Y. (2008) “The relation between driving experience and recognition of road signs relative to their locations”, in Human Factors, 50, 173-182.

(7) Brookhuis, K.A., de Vries, G., de Waard, D. (1991). The effect of mobile telephoning on driving performance. Accident Analysis and Prevention 23 (4), 309–316.

(8) Brookhuis, K.A., de Waard, D. (2010). Monitoring drivers’ mental workload in driving simulators using physiological measures. Accident Analysis &

Prevention, Volume 42, Issue 3, pp. 898-903. https://doi.org/10.1016/j.aap.2009.06.001.

(9) Bishop, C.M. (2006). Pattern recognition and machine learning. Springer

(10) Cain, B., (2007). A Review of the Mental Workload Literature.

(11) P. Calderó-Bardají *et al*., "Detection of steering direction using EEG recordings based on sample entropy and time-frequency analysis," *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, 2016, pp. 833-836. doi: 10.1109/EMBC.2016.7590830

(12) Camlica, Z., Hilal, A., Kulić, D. ( ). Feature abstraction for driver behaviour detection with stacked sparse auto-encoders.

(13) Castro, C. (2009). Human Factors of Visual and Cognitive Performance in Driving. Boca Raton: CRC Press, https://doi.org/10.1201/9781420055337

(14) Castro, C., Horberry, T., & Tornay, F.J. (2004). The Effectiveness of Transport Signs. DOI:10.1201/9780203457412.ch4

(15) Charlton, S. G. (2007). Delineation effects in overtaking lane design. Transportation Research Part F: Traffic Psychology and Behaviour, 10(2), 153–163. https://doi.org/10.1016/J.TRF.2006.09.003

(16) Charlton J.L., Oxley J., Fildes B., Oxley P., Newstead S., Koppel S., O'Hare M. (2006). Characteristics of older drivers who adopt self-regulatory driving behaviours. Transportation Research Part F: Traffic Psychology and Behaviour. 9(5):363–373.

(17) Consiglio, W., Driscoll, P., Witte, M., Berg, W.P. (2003). Effect of cellular telephone conversations and other potential interference on reaction time in a braking response. Accident Analysis and Prevention 325, 495–500.

(18) Costa, M., Simone, A., Vignali, V., Lantieri, C., Bucchi, A., Dondi, G. (2014). Looking behavior for vertical road signs. Transport. Res. F Traffic Psychol. Behav. 23, 147–155.

(19) Crundall, D., Underwood, G. (2001). The Priming Function of Road Signs. Transportation Research Part F: Traffic Psychology and Behaviour, 4, 187-200. http://dx.doi.org/10.1016/S1369-8478(01)00023-7

(20) De Waard, D. (1996). The measurement of drivers' mental workload. Doctor of Philosophy, University of Groningen.

(21) De Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. Transportation Research Part F: Traffic Psychology and Behaviour, 27, 196– 217. https://doi.org/10.1016/j.trf.2014.06.016

(22) A. Diaz Alvarez, F. Serradilla Garcia, J. E. Naranjo, J. J. Anaya and F. Jimenez (2014). Modeling the Driving Behavior of Electric Vehicles Using Smartphones and Neural Networks. *IEEE Intelligent Transportation Systems Magazine*, 6, 3, pp. 44-53. doi: 10.1109/MITS.2014.2322651

(23) Dong, W., Yuan, T., Yang, K., Li, C., Zhang, S. (2017). Autoencoder regularized network for driving style representation learning. In Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17), Carles Sierra (Ed.). AAAI Press 1603-1609.

(24) Edquist J, Horberry T, Hosking. S, Johnston I. (2011). Effects of advertising billboards during simulated driving. Appl Ergon., 42(4):619–26, http://dx.doi.org/10.1016/j.apergo.2010.08.013

(25) Eggemeier, F.T., Wilson, G.F., Kramer, A.F., Damos, D.L. (1991). Workload assessment in multi-task environments, in: Multiple-Task Performance. Taylor&Francis, London, pp. pp. 207–216.

(26) Elliot, M. A., Mccoll, V.A., & Kennedy, J. V. (2003). Road design measures to reduce drivers' speed via 'psychological' processes: A literature review (No.TRL564).

(27) Elvik, R. (2003). Effects on road safety of converting intersections to roundabouts: Review of evidence from non-U.S. studies. Transportation Research

Record: Journal of the Transportation Research Board, 1847, 1–10. https://doi.org/10.3141/1847-01

(28) EU Commission webpage. Funded European projects on "Self-Explaining Roads". https://ec.europa.eu/transport/road_safety/specialist/knowledge/road/designing_for_road_function/self_explaining_roads_en.

(29) European Commission (2019). Handbook on the External Costs of Transport. https://ec.europa.eu/transport/themes/sustainable/studies/sustainable_en

(30) European Commission (2019). Publication of preliminary road safety statistics 2018: http://europa.eu/rapid/press-release_IP-19- 1951_en.htm

(31) Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety. Journal of Planning Literature, 23(4), 347–367. https://doi.org/10.1177/0885412209335553

(32) Farlow, S. J. (1981). The GMDH Algorithm of Ivakhnenko, The American Statistician, 35:4, 210-215, DOI: 10.1080/00031305.1981.10479358.

(33) Fisher, J. (1992). Testing the Effect of Road Traffic Signs′ Informational Value on Driver 530 Behavior. Human Factors, 34(2), 231–237.

(34) FHWA Office of Safety Intersection Safety Website. Crash facts. http://safety.fhwa.dot.gov/intersection/crash_facts/#data

(35) Garach Morcillo, L., Francisco Javier Calvo Poyo, Miguel Pasadas Fernández, Juan de Oña (2014). Measurement of Road Consistency on Two-lane Rural Highways in Granada (Spain), Procedia - Social and Behavioral Sciences, Volume 162, pp. 237-242. https://doi.org/10.1016/j.sbspro.2014.12.204.

(36) Goodfellow, I. (2011). Deep Learning (Adaptive Computation and Machine Learning series)

(37) Gopher, D., Donchin, E. (1986). Workload: An examination of the concept, in: Handbook of Perception and Human Performance, Vol. 2: Cognitive Processes and Performance. John Wiley & Sons, Oxford, England, pp. 1– 49.

(38) Gross, F., Lyon, C., Persaud, B., & Srinivasan, R. (2013). Safety effectiveness of converting signalized intersections to roundabouts. Accident Analysis and Prevention, 50, 234–241. https://doi.org/10.1016/j.aap.2012.04.012

(39) Haddon, W. (1980). Advances in the epidemiology of injuries as a basis for public policy. Public Health Report, 95:411–421.

(40) Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload (pp. 139-183). North-Holland: Elsevier Science.

(41) Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York: Springer.

(42) Heaton, J. (2013). Artificial Intelligence for Humans - Volume 1: Fundamental Algorithms. CreateSpace Independent Publishing Platform.

(43) Hughes, P. K., Cole, B. L. (1986). What attracts attention when driving?, Ergonomics, 29:3, 377-391, DOI: 10.1080/00140138608968272

(44) Irwin, M., Fitzgerald, C., Berg, W.P. (2000). Effect of the intensity of wireless telephone conversations on reaction time in a braking response. Perceptual and Motor Skills 90, 1130–1134.

(45) Jensen, S. (2013). Safety effects of converting intersections to roundabouts. Transportation Research Record: Journal of the Transportation Research Board, 2389, 22–29. https://doi.org/10.3141/2389-03.

(46) Johansson, G., Rumar, K. (1966). Drivers and road signs: A preliminary investigation of the capacity of car drivers to get information from road signs. Ergonomics, 9(1), 57-62. http://dx.doi.org/10.1080/00140136608964342

(47) Kanellaidis, G. (1996). Human Factors in Highway Geometric Design. Journal of Transportation Engineering, Vol. 122, No. 1, American Society of Civil Engineers, pp. 59–66. 32.

(48) Kapitaniak, B., Walczak, M., Kosobudzki, M., Jóźwiak, Z., Bortkiewicz, A. (2015). Application of eye-tracking in the testing of drivers: A review of research.

International journal of occupational medicine and environmental health. 28. 941-54. 10.13075/ijomeh.1896.00317.

(49) Khandil F.I., Rotter, A., Lappe M. (2009). Driving is smoother and more stable when using the tangent point. J Vis.; 9(1),1–11, 50.

(50) Khandil F.I., Rotter, A., Lappe M. (2010). Car drivers attend to different gaze targets when negotiating closed vs. open bends. J Vis., 10(4), 24, 1–11, http://dx.doi.org/10.1167/10.4.24.

(51) Kim, N., Jeong, K., Yang, M., Oh, Y., & Kim, J. W. (2017). "Are you ready to take-over?": An exploratory study on Visual Assistance to enhance driver vigilance. In CHI 2017 Extended Abstracts - Proceedings of the 2017 ACM SIGCHI Conference on Human Factors in Computing Systems: Explore, Innovate, Inspire (pp. 1771-1778). Conference on Human Factors in Computing Systems - Proceedings; Vol. Part F127655. Association for Computing Machinery.https://doi.org/10.1145/3027063.3053155

(52) Knowles, W.B. (1963). “Operator Loading Tasks.”. Human Factors, 31, 155–161.

(53) Knox, S. W. (.). Machine Learning: a concise introduction.

(54) Kramer, U. (1985). “On the application of fuzzy sets to the analysis of the system driver vehicle-environment,” Automatica, vol. 21, no. 1, pp. 101–107.

(55) La Torre, F. (2013). CEDR 2013/09 Forgiving roadsides design guide.

(56) Lees, M. N., Cosman, J. D., Lee, J. D., Fricke, N., Rizzo, M. (2010). Translating cognitive neuroscience to the driver's operational environment: a neuroergonomic approach. Am. J. Psychol, 123(4), 391--411.

(57) Liu, H., Taniguchi, T., Tanaka, Y., Takenaka, K., Bando, T. (2017). Visualization of Driving Behavior Based on Hidden Feature Extraction by Using Deep Learning. IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 9, pp. 2477-2489. doi: 10.1109/TITS.2017.2649541

(58) Martens, M. H. (2000). Assessing Road Sign Perception: A Methodological Review. Transportation Human Factors, 2(4), 347. doi:10.1207/STHF2-4_4

(59) Martens, M. H., Fox, M. R., J. (2007). Do familiarity and expectations change perception? Drivers' glances and response to changes. Transportation Research Part F: Traffic Psychology and Behaviour, 10 (6), pp. 476-492, 10.1016/j.trf.2007.05.003

(60) McCarthy, J. (1958). Programs with Common Sense. Proc. Teddington Conference on the Mechanization of Thought Processes.

(61) McCracken, J.H., & Aldrich, T.B. (1984), Analysis of selected LHX mission functions implications for operator workload and system automation goals, (Anacapa Sciences, INC. Military programs, Alabama).

(62) McCulloch, W. S., Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity.

(63) Meister, D. (1976). Behavioral Foundations of Systems Development. New York, NY: Wiley.

(64) Messer, C. J. (1980). "Methodology for evaluating geometric design consistency" Transportation Research Record, vol.757, U.S.A.

(65) Messer, C.J. (1980). "Methodology for Evaluating Geometric Design Consistency." Facility Design and Operational Effects. Transportation Research Record 757, Transportation Research Board, National Research Council, Washington D.C. pp. 7–14.

(66) Muhlrad N, Lassarre S. (2005). Systems approach to injury control. In: Tiwari G, Mohan D, Muhlrad N, eds. The way forward: transportation planning and road safety. New Delhi, Macmillan India Ltd., 52–73.

(67) Naatanen, R., Summala, H. (1974). A model for the role of motivational factors in drivers' decision-making. Accident Analysis and Prevention, 6(3-4), 243-261. http://dx.doi.org/10.1016/0001-4575(74)90003-7

(68) Parasuraman, R., Hancock, P. (2001). Adaptive control of mental workload, in: Stress, Workload, and Fatigue. pp. 305–320.

(69) Parasuraman, R., and Rizzo, M. (2008). Neuroergonomics: The Brain at Work. Oxford: Oxford University Press.

(70) Pecchini D., Roncella R., Forlani G., Giuliani F. (2014). Measuring driving workload of heavy vehicles at roundabouts.

(71) Prabhakar, T., Lee, S., & Job, R. (1996). Risk-taking, optimism bias and risk utility in young drivers.

(72) Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229.

(73) Schölkopf B., Smola A., Müller KR. (1997) Kernel principal component analysis. In: Gerstner W., Germond A., Hasler M., Nicoud JD. (eds) Artificial Neural Networks — ICANN'97. ICANN 1997. Lecture Notes in Computer Science, vol 1327. Springer, Berlin, Heidelberg.

(74) Sivak, M. (1996). “The information that drivers use: is it indeed 90% visual?” Perception, vol. 25, no. 9, pp. 1081–1089.

(75) Sparse autoencoder (2011), CS294A Lecture notes.

(76) Stephens, A. N., Beanland, V., Candappa, N., Mitsopoulos-Rubens, E., Corben, B., & Lenné, M. G. (2017). A driving simulator evaluation of potential speed reductions using two innovative designs for signalised urban intersections. Accident Analysis and Prevention, 98, 25–36. https://doi.org/10.1016/j.aap.2016.09.022.

(77) Taniguchi, T., K. Furusawa, H. Liu, Y. Tanaka, K. Takenaka, and T. Bando (2016). “Determining utterance timing of a driving agent with double articulation analyzer,” IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 3, pp. 810–821.

(78) Tejero, P., Chóliz, M. (2002). Driving on the motorway: The effect of alternating speed on driver's activation level and mental effort. Ergonomics. 45. 605-18. 10.1080/00140130210145882.

(79) Theeuwes, J. (1998). Self-explaining roads: Subjective categorisation of road environments. In A. G. Gale, I. D. Brown, C. M. Haslegrave, & S. P. Taylor (Eds.), Vision in Vehicles VI. North-Holland: Netherlands.

(80) Theeuwes, Jan & Van der Horst, Richard & Kuiken, M. (2012). Designing safe road systems: A Human Factors perspective.

(81) Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., & Hume, R. D. (1979). Tri-level study of the causes of traffic accidents. Executive summary. Retrieved from https://trid.trb.org/view.aspx?id=144095.

(82) Trivedi M.M., Gandhi T., McCall J. (2007). Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety. IEEE Intell Transp Syst., 8(1):108–20, http://dx.doi.org/10.1109/TITS.2006.88944.

(83) Ulleberg, P. (2002). Influencing subgroups of young drivers and their passengers: motivational influences of personality traits on risk-taking attitudes and driving behaviour. Institute of Transport Economics, Norwegian Centre for Transport Research.

(84) Ulleberg, P., Rundmo, T. (2002). Risk–taking attitudes among young drivers: The psychometric qualities and dimensionality of an instrument to measure young drivers' risk–taking attitudes. Scandinavian Journal of Psychology.

(85) Vignali V., Cuppi, Acerra E., Bichicchi A., Lantieri C., Simone A., Costa M., (2019). Effects of median refuge island and flashing vertical sign on conspicuity and safety of unsignalized crosswalks. Transportation Research Part F: Traffic Psychology and Behaviour, 60, pp. 427-439. ISSN 1369-8478. https://doi.org/10.1016/j.trf.2018.10.033.

(86) Viviani P. (1990). Eye movements in visual search: Cognitive, perceptual and motor aspects. In: Kowler EE, editor. Eye movements and their role in visual and cognitive processes. Amsterdam: Elsevier Science; p. 353–93.

(87) Walker, G. H., Stanton, N. A., & Chowdhury, I. (2013). Self-explaining roads and situation awareness. Safety Science, 56, 18–28. https://doi.org/10.1016/j.ssci.2012.06.018.

(88) Wang Y., Reimer, B., Mehler, B., Zhang, J., Mehler, A., Coughlin, J. F. (2010). The impact of repeated cognitive tasks on driving performance and visual attention.

(89) Wegman, F. (2017). The future of road safety: A worldwide perspective, IATSS Research. Volume 40, Issue 2, Pages 66-71. ISSN 0386-1112 https://doi.org/10.1016/j.iatssr.2016.05.003.

(90) Weller, G., Schlag, B., Friedel, T., & Rammin, C. (2008). Behaviourally relevant road categorisation: A step towards self-explaining rural roads. Accident Analysis and Prevention, 40(4), 1581–1588. https://doi.org/10.1016/j.aap.2008.04.009.

(91) Werneke, J., Vollrath, M. (2012). What does the driver look at? The influence of intersection characteristics on attention allocation and driving behavior. Accident Analysis & Prevention, Volume 45, pp. 610-619, https://doi.org/10.1016/j.aap.2011.09.048.

(92) Wickens, C.D. (1984). "Processing resources in attention", in R. Parasuraman & D.R. Davies (Eds.), Varieties of attention, (pp. 63–102). New York: Academic Press.

(93) Williams, G. W. (1963). Highway hypnosis: An hypothesis. International Journal of Clinical and Experimental Hypnosis, 11(3), 143-151. http://dx.doi.org/10.1080/00207146308409239

(94) Wold, S., Esbensen, K. Geladi, P. (1987). “Principal component analysis” Chemometrics and intelligent laboratory systems, vol. 2, no. 1, pp. 37–52.

(95) World Health Organisation (2018). “Global Status Report on Road Safety”: https://www.who.int/violence_injury_prevention/road_safety_status/2018/en

(96) World Health Organization (2006). Road traffic injury prevention: training manual. https://apps.who.int/iris/handle/10665/43271.

(97) Yang, H. Ozbay, K., Ozturk, O. Xie, K. (2015). Work Zone Safety Analysis and Modeling: A State-of-the-Art Review. Traffic Injury Prevention, 16:4, 387-396, DOI: 10.1080/15389588.2014.948615

(98) Yerkes, R.M., Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology, 18, 459-482.

(99) Young M.S., Stanton N. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. Human Factors.

(100) Zhao, C., Zheng, C., Zhao, M., Liu, J. (2010). Classifying driving mental fatigue based on multivariate autoregressive models and kernel learning algorithms. 3rd International Conference on Biomedical Engineering and Informatics, Yantai. pp. 2330-2334. doi: 10.1109/BMEI.2010.5639579.

List of Annexes

1. Vignali V., Bichicchi A., Simone A., Lantieri C., Dondi G., Costa M., (2019). *Road sign vision and driver behaviour in work zones*. Transportation Research Part F: Traffic Psychology and Behaviour, 60, pp. 474-484. ISSN 1369-8478, https://doi.org/10.1016/j.trf.2018.11.005.
2. Costa M., Bonetti L., Vignali V., Bichicchi A., Lantieri C., Simone A., (2019). *Driver's visual attention to different categories of roadside advertising signs.* Applied Ergonomics, 78, pp. 127-136. ISSN 0003-6870, https://doi.org/10.1016/j.apergo.2019.03.001.
3. Costa M., Bichicchi A., Nese M., Lantieri C., Vignali V., Simone A., (2019). *T-junction priority scheme and road user's yielding behavior.* Transportation Research Part F: Traffic Psychology and Behaviour, 60, pp. 770-782. ISSN 1369-8478. https://doi.org/10.1016/j.trf.2018.12.009.
4. Di Flumeri G., Borghini G., Aricò P., Sciaraffa N., Lanzi P., Pozzi S., Vignali V., Lantieri C., Bichicchi A., Simone A. and Babiloni F., (2018). *EEG-Based Mental Workload Neurometric To Evaluate the Impact of Different Traffic and Road Conditions in Real Driving Settings*. Frontiers in Human Neuroscience, 12, 509. 10.3389/fnhum.2018.00509.
5. Bichicchi A., Belaroussi R., Simone A., Vignali V., Lantieri C. and Li, X. (2020). *Analysis of Road-User Interaction by Extraction of Driver Behavior Features Using Deep Learning*. IEEE Access, 8, pp. 19638-19645. 10.1109/ACCESS.2020.2965940

Transportation Research Part F 60 (2019) 474–484

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Road sign vision and driver behaviour in work zones

Valeria Vignali [a,*], Arianna Bichicchi [a], Andrea Simone [a], Claudio Lantieri [a], Giulio Dondi [a], Marco Costa [b]

[a] Department of Civil, Chemical, Environmental and Material Engineering, University of Bologna, Italy
[b] Department of Psychology, University of Bologna, Italy

ARTICLE INFO

Article history:
Received 16 May 2018
Received in revised form 9 November 2018
Accepted 10 November 2018
Available online 17 December 2018

Keywords:
Roadwork signs
Roadwork zones
Vertical road signs
Driver perception
Eye tracker
Speed
Fixations

ABSTRACT

The effectiveness of roadwork signs on drivers' safety is a poorly investigated topic. The present study examined visual fixations of 29 participants to work zone signs, while driving 27 km along rural roads. The drivers' visual fixations on the work zones signs were recorded with an eye tracking device, synchronized to a GPS recorder that collected kinematic data. The routes crossed 23 roadwork zones, including a total of 69 vertical work zone signs. Visual behaviour to roadwork signs were compared to visual behaviour to permanent vertical signs. The results revealed that drivers glanced at both temporary and permanent signs along the roadwork areas with a similar 40% frequency. In addition, they glanced at single roadwork signs more often and for longer than at multiple-roadwork signs. The main findings of this paper lead to conclude that driver behaviour, investigated by comparing instant speed and visual fixations, is frequently unsafe.

1. Introduction

1.1. Safety in roadwork zones

Roadwork zones are unsafe locations, as they disrupt the drivers' expectations about the road geometry, meaning that they have to make sudden adjustments to their driving speed. Recent research seems to agree that the presence of work zones is likely to increase the crash rate (Yang, Ozturk, Ozbay, & Xie, 2015).

Because of ageing roads, maintenance work is becoming ever more common, so that it is possible to affirm that accidents at roadwork sites are likely to increase and, for this reason, countermeasures should be taken to prevent them.

The overall knowledge about work zone safety was mainly referred to main roadways (such as highways and motorways) and major worksites (those that in general relate to standardizing road layouts). There is little research that addressed safety issues in roadworks in rural roads that are simpler, smaller in size and generally short-termed. Despite that, even rural road crashes may have a considerable social and economic cost.

An extensive literature who analyses work zone collisions mostly rely on simple approaches, such as investigating crashes frequency, external factors, characteristics of the work zones and the type of crash. Observational studies that compared crash rate before and during roadworks have been carried out to test the safety level at specific roadway maintenance sites, by assessing the increase in crash frequency caused by roadworks. Khattak, Khattak, and Council (2002) examined the combined effect of increasing length and duration of freeway worksites in California, finding that there was a significant increase

* Corresponding author.
E-mail address: valeria.vignali@unibo.it (V. Vignali).

https://doi.org/10.1016/j.trf.2018.11.005
1369-8478/

in crash rate compared to the baseline. According to the USA Transportation Research Board (TRB), the occurrence of rear-end and fixed-object collisions increases in correspondence with work zones (Campbell et al., 2012). Similarly, recent data on Italian roads indicate that, between 2007 and 2012, there were 762 collisions in roadwork zones (with 21 fatalities and 1252 injuries). Rear-end collisions were the most frequent, followed by single-vehicle accidents and lateral crashes caused by the driver changing lane (La Torre, Domenichini, & Nocentini, 2017). Besides crash frequency, crash severity in roadwork areas was also investigated. A recent study revealed that advanced-warning, activity and termination areas of a work zone were all associated with higher injury severity crashes (Osman, Paleti, & Mishra, 2018).

Other investigations have similarly developed methodologies to predict crash frequency (Crash Prediction Model), adapting the general equations of the Crash Modification Factor (CMF) to roadworks. CMF is a multiplicative factor that computes the expected number and severity of crashes after implementing a given countermeasure at a specific site (AASHTO – American Association of State Highway and Transportation, 2010). Different methodologies and CMF formulations were developed to estimate the expected number of crashes through the use of prediction model weight (Gross, Persaud, & Lyon, 2010). Crash severity relating to work-in-progress zones is connected to different factors, as the vehicle speed or the involvement of road workers. On this latter point, Kröyer, Jonsson, and Várhelyi (2014) found that at increasing speed there is a significant increase in fatal collisions involging pedestrians. More precisely, they found that the risk of fatality in collisions between a car and a pedestrian is 4–5 times higher at 50 km/h than at 30 km/h. Therefore, forewarn drivers about the presence of working areas represent the simpler practice to induce a speed reduction.

However, speed limits are frequently ignored on road sections with hazardous conditions, such as when there is work-in-progress. Bella (2009), for example, simulated a crossover work zone, with the outcome that mean speed was below the limit only when drivers faced physical constraints. It seems that drivers make adjustments to their driving speed in reaction to contextual changes in the road, rather than simply in compliance with the road sign content. This means that drivers are more likely to comply with speed limits if they see that they match a concomitant danger, as workers or police on the road (Blackman, Debnath, & Haworth, 2014a). The drivers' average speed decreased only if they perceive the necessity to do so (Finley, Jenkins, & McAvoy, 2015). A similar study also examined the drivers' subjective evaluation about whether work zone features had any influence over their choice of speed. The feature that was evaluated as most effective was workers activity, police presence and speed feedback displays (Blackman, Debnath, & Haworth, 2014b).

A main factor in determining whether a crash will occur is linked to whether a work zone is easily visible and recognized. Temporary road signs are the most common tools to achieve both work zone conspicuity and legibility (Bella, 2009), because they inform drivers about the oncoming road conditions beforehand. The effectiveness of signage is related to the "priming effect", the ergonomic paradigm consisting in the anticipation of some information (stimulus) that would influence the response to a subsequent stimulus. In this field, the presence of warning signs informs the driver about the upcoming work-sites and get him ready to take the appropriate action before reaching the hazard. Several studies have proved that being warned beforehand about something enables people to react more quickly, inducing a more correct driving behaviour (Charlton, 2006; Crundall & Underwood, 2001). The capacity to respond to the sign is however influenced by the experience of the context and by the overall driving expertise.

Some studies support the theory that even unconsciously perceived signs (i.e., that drivers do not recall later) are effective in terms of reducing speed, as they implicitly warn drivers about hazards, inducing them to exert proper control over their vehicles (Fisher, 1992; Summala & Hietamaki, 1984).

1.2. Readability of roadwork signs

Both temporary and permanent signage cover an important role in the passive protection of vehicles, passengers, workers and site equipment, since they are used to signal work zones. According to this, it is fundamental to consider their capacity to be easily readable.

Firstly, several studies supports the importance of visual graphics in signage equipment (Costa et al., 2014; Ullman & Brewer, 2014; Ullman, Trout, & Dudek, 2009) and, moreover, the European Union has set up standards for vertical road signs, including graphics such as shape, background colour, border colour, size and symbols (Vienna Convention on Road Signs and Signals, 1968).

Literature offers several studies investigating the role of sign visibility and legibility in relation to the sight distance (Costa et al., 2014; Discetti & Lamberti, 2011; Zwahlen, 1995), but none of them evaluated specifically temporary signage.

Regarding sign design (graphic content, positioning and orientation) Lewis (1989) made a great effort in highlighting the importance of a standardization in work zone signs positioning, also in terms of terminology and definitions. A correct positioning, in fact, means that the road can be more easily monitored, which in turn can avoid the problem of not being warned about the potential negative side-effects of the roadworks, which can include traffic jams, which are a major factor in the increased risk of crashes (Beijer, Smiley, & Eizenman, 2004). In addition, a recent study regarding sign positioning, confirmed that it plays a key role as it affected the drivers' perception–response time and speed (Discetti & Lamberti, 2011). A correct positioning practice, also, suggests to avoid sign overcrowding, as 'visual pollution' from roadside information (intended as billboards, warnings and installations) can distract drivers (Edquist, Horberry, Hosking, & Johnston, 2011) or let drivers to lose important information (Liu, 2005).

Besides, the content of roadwork signs is supposedly to be crucial for the comprehension of drivers' reaction. Several ergonomics studies, in fact, confirm that sign effectiveness does not depend solely on the readability, invoking thus the credibility principle. A study measuring vehicle speed in the presence of different signs found that drivers lift their foot from the accelerator more often and more pointedly when they saw signs they considered to be significant (Summala & Hietamaki, 1984).

1.3. Roadwork activity

A relevant factor for the investigation of drivers' behaviour at work zones is the conspicuity of roadworks, by which we mean the visibility of site operations, workers and active vehicles. Visible site activity, in fact, seems to be an essential requirement on speed modulation. According to the results of a recent study (Steinbakk, Ulleberg, Sagberg, & Fostervold, 2017), higher speed was preferred at work zones without visible roadwork activity and roadwork activity was the strongest predictor of preferred speed. An interesting study by Benekohal and Wang (1994), involving more than one hundred drivers, computed the actual speeds that drivers were travelling at when reaching a road site where work-in-progress was clearly indicated, informing them that they were approaching an operational work site. The findings revealed that the drivers' speed adjustment was strictly connected to their initial speed. Also, it was noted that all the drivers, including those speeding, generally reduced their speed and continued to do so while transiting through the work zone. "Extremely" speedy drivers represented an exception, slowing down in the advance-warning area and speeding up immediately after passing it.

Similarly, the drivers' choices of speed were investigated in presence or absence of road workers. Here, the results show that drivers are significantly more cautious in the presence of workers, as they chose to drive more slowly (Blackman et al., 2014b). Another study confirmed that the size of this effect is dependent on whether the workers are conspicuous. If drivers see solitary or small groups of workers, they are less likely to reduce their speed than if they see larger groups of workers (Haworth, Symmons, & Mulvhill, 2002).

1.4. The application of eye tracking techniques to roadwork zone safety

Eye trackers make possible to investigate the integrated and complex relationship between drivers, traffic, environment and road infrastructure (Bucchi, Sangiorgi, & Vignali, 2012; Dondi, Simone, Lantieri, & Vignali, 2011). This technology allows the assessment of fixation events (i.e., when the eyes focus on a specific point of the scene), distinguishing fixations from saccades (i.e., quick movement of the eyes), providing a direct measure of whether signs are glanced. In addition, fixation duration provides important information on the depth of visual processing.

Literature offers a vast body of evidence that eye tracking technology could be exploited to determine how road equipment affects drivers (Costa et al., 2014; Costa, Simone, Vignali, Lantieri, & Palena, 2018; Filtness et al., 2017; Lantieri et al., 2015; Mantuano, Bernardi, & Rupi, 2017; Taylor et al., 2013; Topolšek, Areh, & Cvahte, 2016; Zwahlen, 1995), involving both simulated and real driving environments.

Nevertheless, a few eye tracking applications has been carried out to investigate drivers' visual behaviour at work zones. For example, drivers' gazing patterns were monitored in a virtual scenario where there were traffic signs belonging to the maintenance roadwork operation. The eye tracker recordings proved useful in concluding that repeated exposure to signs was beneficial to drivers and that interference between permanent and temporary signs is to be avoided, as the drivers' attention is split between them (De Ceunynck et al., 2015). Another study focused on temporary dynamic message signs, and it found that drivers spent longer on fixing their gaze on signs that warned about the presence of road workers (Rahman, Strawderman, Garrison, Eakin, & Williams, 2017).

As most of the experimental research were conducted in driving simulators, this paper aims to fill the gap analysing drivers' gaze to roadwork signs in a real driving test.

2. Methods

2.1. Participants

Twenty-nine participants were recruited among researchers, graduate and undergraduate engineering students. Twenty were men (mean age: 32.95 years, *SD:* 11.72, range: 19–56) and 9 were women (mean age: 36.1 years, *SD:* 12.00, range: 22–54). They all held a regular driving license for cars, with a mean driving expertise of 14.39 years (*SD:* 9.95) and a mean value of kilometres per year of 14,770 (*SD:* 9604). Participants had normal vision without glasses or contact lenses, that prevented the recording of eye movements. They were not informed about the true aim of the study, having been told instead that they were testing the use of a mobile eye tracker device during a driving task. At last, their participation was voluntarily.

2.2. Experimental settings

Driving tests were carried out on rural roadways in Northern Italy, throughout the provinces of Bologna and Reggio Emilia. The selected routes typically had high accident rate and many scattered small-sized road maintenance work zones. The road geometry was consistent along the route, and was a single carriageway with two 3.75 m wide lanes, a shoulder width of 1.5 m (not always present), and a signalled speed limit of 70 km/h.

Along the experimental route, drivers encountered 23 small-sized roadwork zones, with no reduction in lane width at either side. Urban roadworks were excluded from the data analysis. In relation to the signs in the work zones, each driver encountered a total of 69 vertical signs, all with static content, belonging to both temporary (yellow background) and permanent (white background) road vertical signs. The signs with a yellow background were mostly warning signs, while the permanent signs were mostly regulatory road signs. Ten of the considered work zones displayed a single sign (roadworks of negligible length), while the remaining 13 work zones displayed multiple signs (more than two, with an average length of the work zone of 152.61 m). The single signs were all placed at road level, mounted on tripods with an elevation of 0.6–1.20 m from the road surface, beyond the edge-line markings. Work zones with more than two signs included both tripod-mounted and pole-mounted signs, the latter with a maximum height of 2.20 m and placed at 0.3–1 m from the roadside, in compliance with Italian regulations (Fig. 1). Roadwork activity, in terms of presence of visible workers or active vehicles, was encountered in 14 sites over the total of 23 included in the study.

2.3. Apparatus

Experimental vehicles were a Fiat Panda and a BMW series 1 car. Data was collected from 9.30 to 12.00 and from 14.00 to 16.30, to avoid peak rush hours. Driving tests were conducted under good weather conditions, with a dry road surface and complete visibility.

The test vehicles were provided with a Racelogic Video V-Box Pro, a GPS data logger capable of detecting and recording kinematic parameters (forward and lateral acceleration, speed). Two cameras and a GPS antenna, connected with cables to the Video V-Box, were positioned on the top of the cars and recorded the external road scenario, as well as data on acceleration, speed and GPS coordinates. Each driver was given a trial run to get used to the car before starting out along the test route. Speed was recorded with an accuracy of 0.1 km/h and distance accuracy was ±50 cm. The recorded data were analysed using Performance Tools software. The eye tracking equipment and the Video V-Box Pro equipment were kept on the back seat of the car and were monitored by one of the experimenters, who was instructed not to talk to the driver except if assistance was requested.

The combined use of eye tracking monitoring and vehicle kinematic data meant allowed an accurate assessment of the driver' behaviour in work zones. Eye-movement data were available for 29 drivers and kinematic data for 28 drivers, due to technical problem to the Video V-Box equipment in one participant.

Eye movements were recorded with an ASL Mobile Eye-XG tracker. Two digital high-resolution cameras were attached to lightweight eyeglasses. One camera recorded the visual scene while the other camera targeted the participant's eye. The eye tracking recordings were only carried out for the driver's right eye and a calibration process was conducted for each participant. The calibration process took place in a parking lot in a stationary car and involved asking the participants to look at a minimum of 15 visual points spread across the whole scene. The calibration points were chosen between the vertexes and the centres of small objects of the driver's visual scene.

During the tests, the eye movement sampling rate was 30 Hz (i.e., 33 ms time resolution). Spatial accuracy was 0.5–1°. The ASL Mobile Eye-XG software allowed the researchers to match the calibrated datasets with the video recordings and to create, for each participant, a video showing the eye-fixations superimposed to the visual scene (example in Fig. 1).

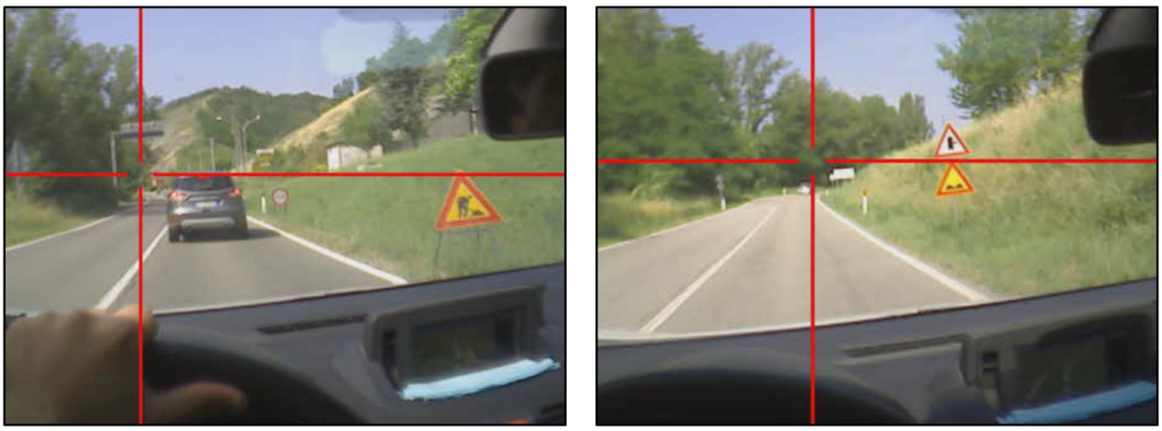

Fig. 1. Tripod-mounted temporary sign (left) and pole mounted temporary sign (right). All temporary signs were triangular in shape, with a yellow background and a red border. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Data analysis

2.4.1. Personal data

Personal data (age, driving licence category, years of car driving, kilometres per year, accident history, prior knowledge of the experimental route) were collected at the end of the experiment, after the driving test. Self-evaluation of driving skills was asked to participants according to four levels: "poor", "average", "good" and "excellent". In particular, 24% of the participants had been responsible of at least one accident and 62% of them had a prior knowledge of the route selected for this study.

2.4.2. Eye-movement data

Drivers' eye fixations on the road signs were assessed through a frame-by-frame analysis of the ASL Mobile Eye-XG video output. Drivers were considered to have fixated a work zone sign if the fixation point (the intersection between the horizontal and vertical line in Fig. 1) was superimposed over the road sign area (AOI: Area of Interest) for at least two frames (66 ms), to avoid the inclusion of saccadic movements. Although research practice normally considers higher temporal thresholds for the definition of a fixation (Holmqvist et al., 2011), the authors' choice was justified by the highly dynamic environment in which eye movements were recorded. Under such conditions, differently from a recording in a virtual environment or in a more controlled setting as in a laboratory, the highly dynamic optical flow of a real driving context implies a rapid sequence of saccades and short fixations (Costa, Simone, et al., 2018). The total fixation duration was computed multiplying 33 ms by the number of frames in which the road sign was fixated.

Once the scorer detected an eye fixation on a work zone sign, the distance of this visual fixation (longitudinal distance on approach to the sign) was acquired by synchronizing the eye tracker video with the Video-V-Box output (Fig. 2). The distance between the first fixation to a road sign and the position where the car was perpendicular to the sign (overtaking the sign) was computed using the Video-V-Box distance parameter. In the case of multiple fixations, the distance was computed considering the first fixation.

Fig. 2. Video V-Box (bottom) and Mobile Eye Tracker (top) synchronization for the computation of the distance of first-fixation to a road sign.

2.4.3. Speed analysis

Speed was entered in the analysis considering these parameters:

– instant speed, as the speed at the time of first-fixation to the road sign;
– approaching speed, as the speed at 100 m before the first sign of the work zone;
– speed reduction, differential between the speed at the time of first-fixation and the speed at the time the driver crossed the road sign;
– work zone speed: the average speed along the whole work zone.

3. Results

3.1. Road sign fixation rates

Table 1 shows the fixation frequency and the absolute frequency of the road signs included in the work zones considered in the study.

In decreasing order, the road signs that received more glances were: Slippery road (64.2%), Uneven road (53.85%), Generic danger (50.41%), Loose chippings (50%). The road signs that were glanced with a percentage lower than 50% were: No overtaking (47.92%), Roadworks (44.01%), Keep left (35.17%), Speed limit (35.17%), Work zone ahead (37.14%), Give priority to vehicles from opposite direction (28.57%), Work zone end (27.78%), Road narrows (22.5%), Modified visibility (14.29%), Hump (0%).

The overall mean fixation percentage, weighted according to the frequency of each sign, was 40.14% (*SD:* 17.09%).

The distinction between temporary and permanent signs was not critical for fixation frequency: (mean value 40.37% and standard deviation 18.51 for temporary signs; mean value 39.78% and standard deviation 14.11 for permanent signs) (Table 2).

Table 1
Fixation frequency and absolute frequency for each road sign included in the work zones.

Road sign	Sign icon	#	Fixation frequency
Roadworks		17	44.01%
Generic Danger		9	50.41%
Hump		2	0.00%
Road Narrows		3	22.50%
Loose Chippings		2	50.00%
Uneven Road		2	53.85%
Modified Viability		1	14.29%
Work Zone Ahead		4	37.14%
Work Zone End		2	27.78%
Slippery Road		2	64.29%
Give priority tovehicles from oppositedirection		2	28.57%
Speed Limit		12	35.17%
No Overtaking		6	47. 92%
Keep Left		5	35.78%
TOTAL		**69**	40.14%

Table 2
Fixation frequency for temporary and ordinary road signs along the work zones considered in the study.

Sign classification	Sign typology	#	Fixation frequency
Temporary Signs	Warning	36	40.37%
(Yellow Background)	Direction	6	
Ordinary Signs	Warning	2	39.78%
(White Background)	Regulatory	25	

Fixations on work zone signs were not influenced by age ($r = 0.07$, n.s.) or gender ($F(1,131) = 0.282$, n.s.). Two linear regression models tested the effects of kilometres per year and years of driving experience on the fixation percentage to the road signs included in the work zones. Both regressions were not significant ($p = 0.58$ and $p = 0.37$ respectively).

Road sign positioning on approach to the work zones was also considered. Specifically, we compared fixation frequency to the first work zone sign and the following road signs. The fixation frequency was higher for the first sign (M: 41.74%, *SD*: 11.80), than for the following road signs (M: 38.94%, *SD*: 17.87).

Fixation rate to the first temporary sign in the work zone was compared considering the presence-absence of visible road-work activity. For work zones with visible activity, the fixation rate to the first temporary sign in the work zone increased to 62.96% (*SD:* 33.95). Chi-square test was used and resulted equal to 5.7273 with a p-value of 0.0167.

3.2. Fixation duration

The distribution of fixation durations to the road signs is shown in Fig. 3. Since the distribution was not normal we report the median as a measure of centrality. The median fixation length was of 132 ms (*SD:* 108.67, mode: 66). The distribution resulted to be highly asymmetrical and positively skewed, with a kurtosis of 14.08 (*SD:* 0.052) and an asymmetry of 3.06 (*SD:* 0.027). Both Kolmogorov-Smirnov and Shapiro-Wilk normality tests were significant ($p < 0.001$), showing that the distribution was not normal.

The average fixation time was also specifically computed considering users' self-evaluation of their driving skills. This was 107.25 ms (*SD:* 31.60) for the drivers who self-evaluated their driving skills as "average", 169.32 ms (*SD:* 104.34) for those professing "good" skills and 156.75 ms (*SD:* 45.34) for those who thought that they had "excellent" driving skills.

At last, results show that the fixation duration was not influenced by drivers' age, gender or prior experience with the experimental route.

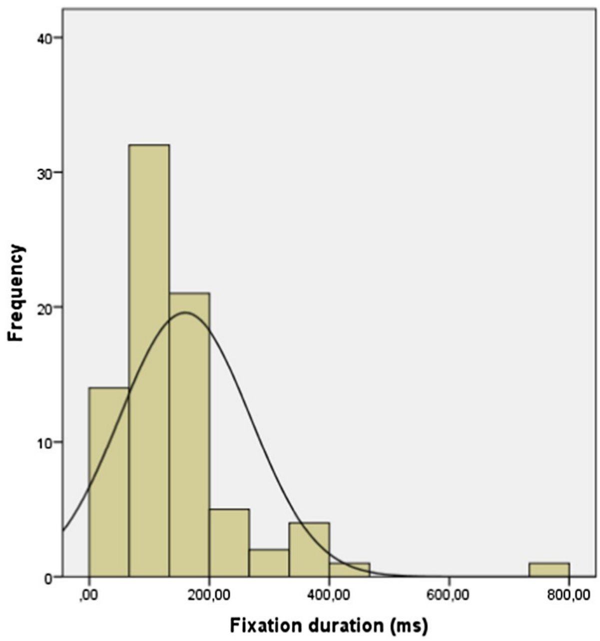

Fig. 3. Distribution of fixation duration to the road signs included in the study.

3.3. Fixation distance

Results revealed that first-fixation to work zones were generally recorded at a mean distance of 43.5 m (*SD:* 32.5, range 15–80), increasing to 48.48 m (*SD:* 34.85) if ongoing activity was present. The difference in first-fixation distances with or without ongoing visible activity was not significant. In addition, the ANOVA test for assessing the effects of visible activity as an independent variable on first-fixation distance resulted as non-significant $F(2, 74) = 2.257$.

Also, the mean distance of first fixation was not significantly different considering work zones with one road signs versus work zones with multiple road signs.

3.4. Speed

In average, drivers fixated the first sign at each work zone at an instant speed of 55.34 km/h (*SD:* 13.92). Speed limit (70 km/h) was exceeded by 14% of the participants.

The correlation between instant speed and distance of first fixation was equal to $r = 0.22$, $p = 0.049$. If work zones are distinguished by visible activity, none relevant relationship with sight distance is obtained (Fig. 4).

To determine whether the first fixations occurred at an instant speed that allowed a safe stop of the vehicle in the case of an unexpected obstacle, the distance of first-fixation was compared to the stopping distance. The latter is dependent on travelling speed and has been evaluated as the sum of the reaction distance (reaction time * initial speed) and braking distance, according to Italian regulations (Ministero delle Infrastrutture e dei Trasporti, 2001; World Road Association, 2003). The results showed that distance of first-fixation exceed stopping distance only in 19.48% of cases.

The other speed parameters were:

- average approaching speed: 55.69 km/h (*SD:* 14.04);
- average speed reduction: −21.89 km/h (*SD:* 26.85);
- average whole work zone speed:52.21 km/h (*SD:* 12.18).

In terms of driving performance, the linear regression between approaching speed and speed reduction after the first fixation in the work zone was non-significant (Fig. 5). To the contrary, the linear regression between approaching speed and the whole work zone average speed was significant ($R^2 = 0.55$ and $p = 0.05$ in Fig. 6).

Speed reduction was significantly related to the driver's age ($r = 0.35$, $p < 0.001$) and driving expertise ($r = 0.340$, $p < 0.001$). The work zone average speed resulted to be related to the drivers' driving expertise ($r = -0.249$, $p = 0.043$).

4. Discussion

The drivers' visual behaviour revealed that work zone signage received very little attention overall, with a mean 40.14% probability of looking at roadwork signs. The frequency was similar for permanent and temporary road signs. In a recent previous study by Costa et al. (2014), that used a similar experimental protocol but focused on roads without work zones, vertical signs were generally looked at with a 25% frequency. This comparison clearly shows that in work zones the frequency of road sign glances was higher than in normal road sections.

Assuming, however, that the work zone sign primary role is to trigger drivers' attention on modified road setting, the signs failed to be glanced on average in 60% of the cases, which is very high. This data is even more significant when considering that the participants wore an eye tracker device, drove an unfamiliar car and knew that their driving behaviour was being studied. This frequency, however, does not take into account a possible involvement of peripheral vision in road sign detection and identification (Costa, Bonetti, Vignali, Lantieri, & Simone, 2018).

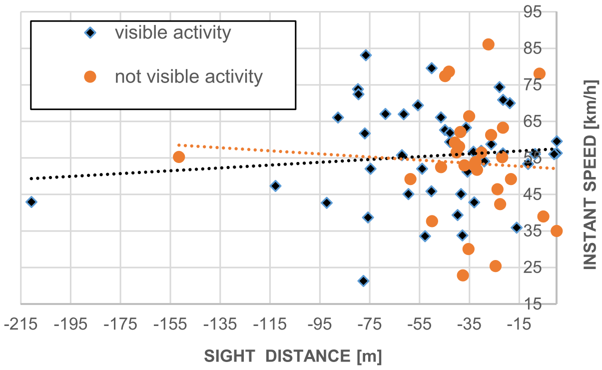

Fig. 4. The distribution of first gaze distance/speed, by visible activity.

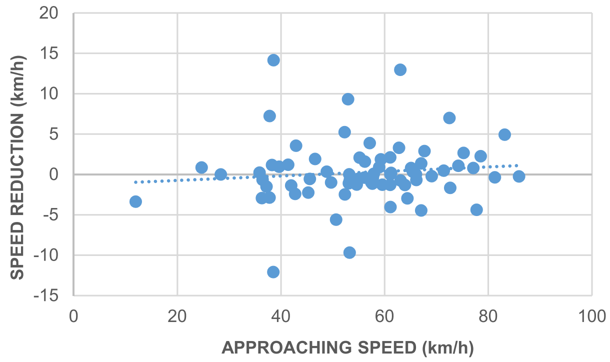

Fig. 5. Relationship between approaching speed and speed reduction.

Gender and age had no influence on fixation frequency, fixation length, fixation distance and speed. For gender, the result confirms previous studies that have monitored eye movements during driving (Costa et al., 2014; Lantieri et al., 2015). The sample, however, was rather low and included only nine females.

Driving expertise had not influence on fixation frequency but correlated significantly with speed reduction approaching the work zone and work zone average speed, coherently with Duncan, Williams, and Brown (1991).

About the knowledge of the route, drivers' experience of the route did not have any influencing effect on the fixation frequency. The novelty effect potentially owned by work zone signs has not influenced fixations neither for the drivers who already experienced that road section, nor for unexperienced drivers. To the contrary, experienced drivers had a higher speed crossing the work zones.

About the driving-skill self-evaluation, the drivers who judged their own driving skills as limited drove more carefully exhibiting lower speed. This result is consistent with a study concerning the reliability of drivers' self-reports (West, French, Kemp, & Elander, 1993). The same category of drivers exhibited also lower fixation times to road signs.

4.1. Work zone features

Concerning work zone features, it is possible to consider that:

- about the single/multiple temporary signs, isolated single signs in work zones caught more attention by the drivers (in terms of both frequency and average duration of the fixations) than a sequence of signs along a work zone. This could be explained by the height of the signs, as single signs frequently were tripod-mounted and positioned at the bottom of the drivers' visual field (0.6–1.20 m from the road surface) and are perceived to be narrower, confirming previous studies (Bella, 2009);
- about the ongoing activity on the work zone, ongoing visible activity on the work zone slightly anticipated the distance of first-fixation to the road signs, probably because the presence of dynamic elements on the visual scene increase the conspicuity and detectability of the work zone. Ongoing visible activity had no influence on speed.

4.2. Safety considerations

The present study addressed the importance of understanding the influence of work zone elements on drivers' road sign vision and behaviour. The comparison between average approaching speed and average speed reduction revealed a useful

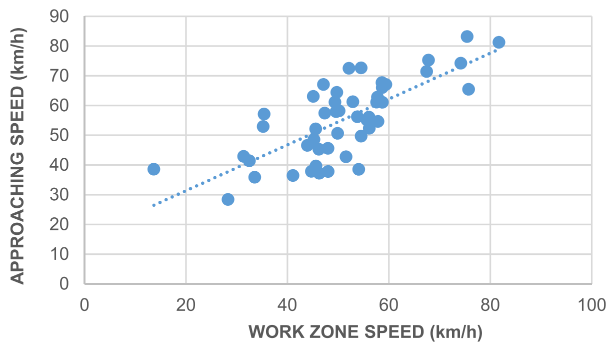

Fig. 6. Relationship between work zone speed and approaching speed.

test of the efficacy of roadworks signage. Age and poor expertise were predictors for higher speed reductions, but not for fixation rates. Also average whole work zone speed resulted adequate.

The analysis that had a direct implication for road safety is the comparison between the fixation distance and the correspondent stopping distance. The sight distance, whose importance has been extensively discussed in literature (Discetti & Lamberti, 2011), was frequently lower than stopping distance. Practically, the inadequate effectiveness of signage would not allow a safe stop in case of a sudden obstacle on the road. The knowledge provided would have a strong practical utility for increasing work zone safety levels using appropriate signalling.

In fact, to generalize the presented outcomes, further driving tests should be devised to include diversified road geometries and work zone settings (as length of the advance-warning area, type of first sign, novel instalments as flashing lights, electronic variable message signs and flaggers). On the contrary, authors highlight the importance of maintain the focus of attention on small work zones and consequently on rural environment scenarios, as the risk of severe crashes has been previously proved (Osman et al., 2018).

The reduced sample dimension represents a further limitation of the presented study and certainly will be considered for future testing.

Acknowledgments

The research was supported by the Alma Idea 2017 Costa grant from the University of Bologna.

References

AASHTO – American Association of State Highway and Transportation (2010). Highway Safety Manual (2010th ed.). Washington, D.C.

Beijer, D., Smiley, A., & Eizenman, M. (2004). Observed driver glance behavior at roadside advertising signs. *Transportation Research Record: Journal of the Transportation Research Board, 1899*, 96–103. https://doi.org/10.3141/1899-13.

Bella, F. (2009). Effects on driver behaviour of different signalling schemes of work zones. *Advances in Transportation Studies, 18*, 55–68.

Benekohal, R., & Wang, L. (1994). Relationship between initial speed and speed inside a highway work zone. *Transportation Research Record, 1442*, 41–48.

Blackman, R. A., Debnath, A. K., & Haworth, N. L. (2014a). Influence of visible work activity on drivers' speed choice at roadworks. In *Proceedings of the 2nd occupational safety in transport conference* (pp. 1–10). CARRS-Q, Queensland University of Technology.

Blackman, R. A., Debnath, A. K., & Haworth, N. L. (2014b). Work zone items influencing driver speeds at roadworks: Worker, driver and expert perspectives. *Australasian road safety research, policing and education conference (RSRPE 2014), 12–14 November 2014, Melbourne, Victoria*.

Bucchi, A., Sangiorgi, C., & Vignali, V. (2012). Traffic psychology and driver behavior. *Procedia-Social and Behavioral Sciences, 53*, 972–979. https://doi.org/10.1016/j.sbspro.2012.09.946.

Campbell, J. L., Lichty, M. G., Brown, J. L., Richard, M. C., Graving, J. S., Graham, J., ... Harwood, D. (2012). *NCHRP Report 600: Human factors guidelines for road systems* Retrieved from (2nd ed.). Washington, D.C: Transportation Research Board.

Charlton, S. G. (2006). Conspicuity, memorability, comprehension, and priming in road hazard warning signs. *Accident Analysis and Prevention, 38*(3), 496–506. https://doi.org/10.1016/j.aap.2005.11.007.

Costa, M., Bonetti, L., Vignali, V., Lantieri, C., & Simone, A. (2018). The role of peripheral vision in vertical road sign identification and discrimination. *Ergonomics*. https://doi.org/10.1080/00140139.2018.1508756.

Costa, M., Simone, A., Vignali, V., Lantieri, C., Bucchi, A., & Dondi, G. (2014). Looking behavior for vertical road signs. *Transportation Research Part F: Traffic Psychology and Behaviour, 23*, 147–155. https://doi.org/10.1016/j.trf.2014.01.003.

Costa, M., Simone, A., Vignali, V., Lantieri, C., & Palena, N. (2018). Fixation distance and fixation duration to vertical road signs. *Applied Ergonomics, 69*, 48–57. https://doi.org/10.1016/j.apergo.2017.12.017.

Crundall, D., & Underwood, G. (2001). The priming function of road signs. *Transportation Research Part F: Traffic Psychology and Behaviour, 4*(3), 187–200. https://doi.org/10.1016/S1369-8478(01)00023-7.

De Ceunynck, T., Ariën, C., Brijs, K., Brijs, T., Van Vlierden, K., Kuppens, J., ... Wets, G. (2015). Proactive evaluation of traffic signs using a traffic sign simulator. *European Journal of Transport and Infrastructure Research (EJTIR), 15*(2), 184–204.

Discetti, P., & Lamberti, R. (2011). Traffic sign sight distance for low-volume roads. *Transportation Research Record: Journal of the Transportation Research Board, 2203*, 64–70. https://doi.org/10.3141/2203-08.

Dondi, G., Simone, A., Lantieri, C., & Vignali, V. (2011). Bike lane design: The context sensitive approach. *Procedia Engineering, 21*, 897–906. https://doi.org/10.1016/j.proeng.2011.11.2092.

Duncan, J., Williams, P., & Brown, I. (1991). Components of driving skill: Experience does not mean expertise. *Ergonomics, 34*(7), 919–937.

Edquist, J., Horberry, T., Hosking, S., & Johnston, I. (2011). Effects of advertising billboards during simulated driving. *Applied Ergonomics, 42*(4), 619–626. https://doi.org/10.1016/j.apergo.2010.08.013.

Filtness, A. J., Larue, G., Schramm, A., Fuller, J., Rakotonirainy, A., Han, C., & Cairney, P. (2017). Safety implications of co-locating road signs: A driving simulator investigation. *Transportation Research Part F: Traffic Psychology and Behaviour, 47*, 187–198. https://doi.org/10.1016/j.trf.2017.04.007.

Finley, M. D., Jenkins, J., & McAvoy, D. S. (2015). Motorists' speed response to nonvariable and variable work zone speed limits and other work zone conditions. *Transportation Research Record: Journal of the Transportation Research Board, 2485*, 70–77. https://doi.org/10.3141/2485-09.

Fisher, J. (1992). Testing the effect of road traffic signs' informational value on driver behavior. *Human Factors, 34*(2), 231–237. https://doi.org/10.1177/001872089203400208.

Gross, F., Persaud, B., & Lyon, C. (2010). A Guide to Developing Quality Crash Modification Factors. Report No. FHWA-SA-10-032. Washington, DC.

Haworth, N., Symmons, M., & Mulvhill, C. (2002). Safety of Small Workgroups on Roadways. Victoria, Australia.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van De Weijer, J. (2011). *Eye Tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Khattak, A. J., Khattak, A. J., & Council, F. M. (2002). Effects of work zone presence on injury and non-injury crashes. *Accident Analysis and Prevention, 34*(1), 19–29. https://doi.org/10.1016/S0001-4575(00)00099-3.

Kröyer, H. R. G., Jonsson, T., & Várhelyi, A. (2014). Relative fatality risk curve to describe the effect of change in the impact speed on fatality risk of pedestrians struck by a motor vehicle. *Accident Analysis and Prevention, 62*, 143–152. https://doi.org/10.1016/j.aap.2013.09.007.

La Torre, F., Domenichini, L., & Nocentini, A. (2017). Effects of stationary work zones on motorway crashes. *Safety Science, 92*, 148–159. https://doi.org/10.1016/j.ssci.2016.10.008.

Lantieri, C., Lamperti, R., Simone, A., Costa, M., Vignali, V., Sangiorgi, C., & Dondi, G. (2015). Gateway design assessment in the transition from high to low speed areas. *Transportation Research Part F: Traffic Psychology and Behaviour, 34*, 41–53. https://doi.org/10.1016/j.trf.2015.07.017.

Lewis, R. M. (1989). Work-zone traffic control concepts and terminology. *Transportation Research Record, 1230*, 1–11.

Liu, Y. C. (2005). A simulated study on the effects of information volume on traffic signs, viewing strategies and sign familiarity upon driver's visual search performance. *International Journal of Industrial Ergonomics, 35*(12), 1147–1158. https://doi.org/10.1016/j.ergon.2005.06.009.

Mantuano, A., Bernardi, S., & Rupi, F. (2017). Cyclist gaze behavior in urban space: An eye-tracking experiment on the bicycle network of Bologna. *Case Studies on Transport Policy, 5*(2), 408–416. https://doi.org/10.1016/j.cstp.2016.06.001.

Ministero delle Infrastrutture e dei Trasporti (2001). Norme Funzionali Geometriche per la Costruzione delle Strade.

Osman, M., Paleti, R., & Mishra, S. (2018). Analysis of passenger-car crash injury severity in different work zone configurations. *Accident Analysis and Prevention, 111*(2018), 161–172. https://doi.org/10.1016/j.aap.2017.11.026.

Rahman, M. M., Strawderman, L., Garrison, T., Eakin, D., & Williams, C. C. (2017). Work zone sign design for increased driver compliance and worker safety. *Accident Analysis and Prevention, 106*, 67–75. https://doi.org/10.1016/j.aap.2017.05.023.

Steinbakk, R. T., Ulleberg, P., Sagberg, F., & Fostervold, K. I. (2017). Analysing the influence of visible roadwork activity on drivers' speed choice at work zones using a video-based experiment. *Transportation Research Part F: Traffic Psychology and Behaviour, 44*, 53–62. https://doi.org/10.1016/j.trf.2016.10.003.

Summala, H., & Hietamaki, J. (1984). Drivers' immediate responses to traffic signs. *Ergonomics, 27*(2), 205–216. https://doi.org/10.1080/00140138408963478.

Taylor, T., Pradhan, A. K., Divekar, G., Romoser, M., Muttart, J., Gomez, R., ... Fisher, D. L. (2013). The view from the road: The contribution of on-road glance-monitoring technologies to understanding driver behavior. *Accident Analysis and Prevention, 58*, 175–186. https://doi.org/10.1016/j.aap.2013.02.008.

Topolšek, D., Areh, I., & Cvahte, T. (2016). Examination of driver detection of roadside traffic signs and advertisements using eye tracking. *Transportation Research Part F: Traffic Psychology and Behaviour, 43*, 212–224. https://doi.org/10.1016/j.trf.2016.10.002.

Ullman, G., & Brewer, M. (2014). Driver perceptions of traffic-calming versus active enforcement efforts in work zones. *Transportation Research Record: Journal of the Transportation Research Board, 2425*, 25–31. https://doi.org/10.3141/2425-04.

Ullman, B. R., Trout, N. D., & Dudek, C. L. (2009). Use of graphics and symbols on dynamic message. *Signs: Technical Report, 7*(2).

Vienna Convention (1968). Protocol on road signs and signals. *Journal of Chromatographic Science*, 51. https://doi.org/10.1093/chromsci/bms220.

West, R., French, D., Kemp, R., & Elander, J. (1993). Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics, 36*(5), 557–567. https://doi.org/10.1080/00140139308967912.

World Road Association (PIARC) (2003). Road safety manual.

Yang, H., Ozturk, O., Ozbay, K., & Xie, K. (2015). Work zone safety analysis and modeling: A state-of-the-art review. *Traffic Injury Prevention, 16*(4), 387–396. https://doi.org/10.1080/15389588.2014.948615.

Zwahlen, H. T. (1995). Traffic sign reading distances and times during night driving. *Transportation Research Record, 1495*, 140–146.

Applied Ergonomics 78 (2019) 127–136

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Driver's visual attention to different categories of roadside advertising signs

Marco Costa[a,*], Leonardo Bonetti[b,a], Valeria Vignali[c], Arianna Bichicchi[c], Claudio Lantieri[c], Andrea Simone[c]

[a] Department of Psychology, University of Bologna, Italy
[b] Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark
[c] Department of Civil, Chemical, Environmental and Material Engineering, University of Bologna, Italy

ARTICLE INFO

Keywords:
Roadside advertising signs
Eye movements
Distraction
Fixation duration
Fixation distance

ABSTRACT

Roadside advertising signs are a salient potential source of driver's distraction. Previous research has mainly investigated driver's visual attention to billboards, which represents only one category of advertising signs. In this study, driver's visual attention was assessed in a naturalistic driving setting for six categories of roadside advertising signs: vendor signs, billboards, movable display boards, single and multiple commercial directional signs, and gas price LED displays. Fixation rate, fixation duration, fixation distance and driving speed were assessed in a sample of 15 drivers along a 30-km route including a total of 154 advertising signs belonging to the six categories described above. The role of clearance from the road, elevation, height, width, surface, number and size of characters, total number of characters, side of the road (driving side, opposite side), context (rural, urban), were also considered. Overall 24% of the roadside advertising signs were fixated. Fixation rate was significantly influenced by sign category, clearance from the road and number of characters. Median value for fixation duration was 297 ms. Fixation duration was significantly influenced by speed, elevation from road level, number of medium size characters, and was higher in the rural context. Median value for fixation distance was 58.10 m, and was significantly influenced by advertising sign category, character count and speed.

1. Introduction

Road advertising signs can adversely affect driver's behavior, representing a salient potential danger for driving. Specifically, advertising signs may distract from the driving task, obstruct visibility at intersections or driveways, present a physical obstacle to vehicles going off road, or interfere with attention to formal traffic signs. In our study, we assessed eye fixations, fixation duration, and distance of fixation to six categories of roadside advertising signs, considering driver's speed, sign placement, size, and other design parameters. Differently from previous research, we explored the impact of six different categories of advertising signs: billboards, vendor signs, single and multiple commercial directional signs, movable display boards, and gas price LED displays.

Studies on accident statistics have often included advertisements as a critical cause. The first example is represented by Ady (1967), who recorded accident rates before and after the installation of billboards on different road segments. One billboard placed at the corner of a sharp bend and brightly illuminated appeared particularly critical. Advertising signs placed on curves are particularly problematic due to their potential distracting effect in a context of high cognitive load (Land and Lee, 1994). However, the positioning of advertising signs on curves is often requested by advertisement agencies because they receive more fixations than advertisements placed laterally on straight road segments (Beijer et al., 2004).

Advertising billboards are highly conspicuous due to their size, coloration, and layout, all designed to be attention-grabbing. Whenever the driver's attention is diverted from the driving task and addressed to irrelevant information such as advertising signs, driving could be impaired (e.g., Lansdown et al., 2015; Wallace, 2003). Wallace (2003), in a review, concluded that there was a real risk of increased accident vulnerability, but that this effect is situation specific, and is due to the additive effect of the advertising sign to other environmental critical aspects.

Many researchers have used driving simulation to study the effect of road advertisement on attentional resources and driving performance, even if the external validity of data obtained with a simulator is reduced in comparison to ecological studies with data collected in real

* Corresponding author. Department of Psychology, University of Bologna, Viale Berti Pichat, 5, I-40127, Bologna, Italy.
E-mail addresses: marco.costa@unibo.it (M. Costa), leonardo.bonetti2@unibo.it (L. Bonetti), valeria.vignali@unibo.it (V. Vignali), arianna.bichicchi2@unibo.it (A. Bichicchi), claudio.lantieri2@unibo.it (C. Lantieri), andrea.simone@unibo.it (A. Simone).

https://doi.org/10.1016/j.apergo.2019.03.001
Received 31 October 2017; Received in revised form 25 February 2019; Accepted 1 March 2019
0003-6870/

naturalistic driving conditions.

The effect of elevation on advertising sign conspicuity was investigated by Crundall et al. (2006) who found that street-level-advertisements received longer fixations that 3-m raised advertisements. In Bendak and Al-Saleh (2010) driving performance was assessed in a driving simulator in two conditions, one with and one without roadside advertising signs. Two driving performance parameters, drifting from lane and recklessly crossing dangerous intersections, were significantly worse in the condition with advertising signs. Similarly, Young et al. (2009) found that the presence of billboards adversely affected driving performance in terms of lateral control. Edquist et al. (2011) examined the level of distraction induce by billboards with the ISO standardized *Lane Change Test*, that involves changing lanes in response to road signs at regular intervals. They found that the presence of billboards diverted eye movements from the road ahead and delayed responses to road signs by 0.5 to 1 s. Marciano and Yeshurun (2012) studied the perceptual load induced by roadside advertisement billboards on driving performance in a simulator. The attentional capture interfered with the ability of participants to distribute their attentional resources efficiently between the road and its sides, resulting in less effective search for critical events that in some occasions ended in accidents.

Other studies have investigated roadside advertisement using ecological settings. Beijer et al. (2004), for example, studied glance behavior of 25 drivers at advertising signs along a Toronto expressway. Averaged glance was 0.57 s in duration. Active signs, containing movable displays, received a significant higher number of glances than passive signs. The number of long glances was also greater for active signs compared to the passive ones. Topolšek et al. (2016) recorded eye movements while participants performed 10 km of urban driving including 56 traffic signs and 31 advertising signs. They found a 0.75 correlation between visual detection of traffic signs and visual detection of advertising signs. Mean detection of advertisements was 0.37. Decker et al. (2015) performed a systematic literature review on the effects of billboards on driver visual behavior. The results highlighted that about 10–20% of all glances at billboards were greater than 0.75 s, and that active billboards drew more and longer glances than passive billboards. Moreover, there is now general agreement that digital billboards attract more attention and cause greater impairment to driving performance when compared with static billboards. In general, rapidly changing stimuli near the roadway may represent a potential safety concern (Belyusar et al., 2016).

Driver's distraction is also significantly affected by the emotional arousal induced by the advertising signs. Chan and Singhal (2013) found that driving performance was differentially affected by the valence of the advertisement emotional content. Megías et al. (2011) investigated the role of the emotional content of roadside advertisements on modulation of attention in a risky driving scenario and the results indicated that the number of fixations and total fixation time elicited by affect-laden advertisements were greater than those elicited by neutral advertisements. Negative pictures received later gaze disengagement than positive and neutral images, and this attentional capture resulted in lower eye fixations to the road.

Signs that are placed in the driver's line of sight are much more fixated in comparison to roadside signs. For example, in the study conducted by Lantieri et al. (2015) eye movements were recorded in a naturalistic driving setting with gateways that marked the transition from rural to urban sections that differed for their design. An extended town sign placed 6.5 m high over the driver's lane, received significantly more visual attention in comparison to traditional town signs placed on the roadside.

Costa et al. (2014) argued that the lateral placement of traffic signs is one of the main reasons for the poor visual detection of vertical traffic signs. Costa et al. (2014), in an ecological setting of 8.34 km driving, found that only 25.06% of vertical traffic signs were glanced by drivers.

Road advertising signs strongly differ for their form, illumination and color (Elliot and Maier, 2014), all critical features that are able to variously affect driver's attention/distraction. Visual processing of signs with logos, for example, is slightly higher than that for guide signs (Kaber et al., 2015). Furthermore, conspicuous colours such as yellow, orange, red or various kinds of illumination or reflective materials can easily capture driver's visual attention (Costa et al., 2016; Lesley, 1995). Similarly, text length, font and size have been shown to have a direct influence on fixation frequency, duration, and distance on print advertisements (Rayner et al., 2001).

Previous research focused quite exclusively on billboards, but road sides host many other types of advertising signs that can contribute to the driver's distraction. The main aim of our study was therefore to assess driver's visual behavior in response to categories of roadside advertisement that were never investigated by previous research. In addition to billboards, we included vendor signs, directional signs for companies, restaurants, commercial activities (distinguishing between single and multiple directional signs), movable display boards and gas price LED displays.

Regulations and restrictions on roadside billboards are highly diversified. In U.S. billboards are regulated by the *Highway Beautification Act* since 1965, with some states in which they are prohibited (e.g., Vermont, Alaska, Hawaii, and Maine), and others in which they are subjected to specific requirements. According to the Italian regulation (Italian Highway Code, 1992) road advertisements outside urban areas, should not exceed 6 m^2, with the exception of vendor signs parallel to the street that should not exceed 20 m. The inferior border should be higher than 1.5 m from the street level. Inside urban areas they are subjected to local regulations. Directional signs for commercial and tourist locations are rectangular with a size between 1×0.2 m and 1.5×0.3 m. Multiple directional signs should not exceed six signs mounted on a same pole. Extra care is requested in the use of the red color to avoid confusion with traffic signs. Roadside billboards should be installed at a distance exceeding 3 m from road shoulders, and at a distance greater than 250 m from intersections. They are prohibited in curves.

Visual attention to advertising signs was assessed by eye movement recording (Land and Tatler, 2009). Eye movements are highly correlated with the path of visual attention (Nuthmann and Einhäuser, 2015; Nuthmann et al., 2017) and therefore the probability to attract fixations could be an indirect measure of a roadside advertisement salience. Eye movements during driving tend to follow specific patterns. Several studies, for example, have found that horizontal eye movements are dominant over vertical movements (Crundall and Underwood, 1998; Mourant and Rockwell, 1972). Most fixations are directed to the focus of expansion (i.e., the point in space where all optical flow vectors intersect), with the driver looking straight ahead. When not looking at the focus of expansion, the eyes usually scan the left or right visual scene with horizontal saccades, looking for driving-relevant information (Underwood et al., 2003).

The SEEV model (Horrey et al., 2006; Wickens, 2007, 2008; Wickens and McCarley, 2008) establishes four components that predict allocation of visual attention (e.g., in a scan path) across any large scale environment such as the road: saliency, effort, expectancy, and value. Saliency expresses the bottom-up attention capturing property of a stimulus (e.g., the size of a roadside advertisement and the contrast of its graphical design). Effort expresses the inhibition of attention shifts due to the costs of large saccadic movements, or head movements (e.g., the lower visual attention to roadside advertisements that are placed with higher clearance from the road). Expectancy and value are mainly top-down processes. The first expresses the likelihood of seeing an event, and the second the importance and the relevance of the event (e.g., enhanced attention to gas price signs because the fuel light is on). In our study, we focused only on bottom-up processes, saliency and effort, as predictors of visual attention to roadside advertisements differing for their category, size, design, and content.

The main dependent variable in our study was the proportion of advertising signs that were fixated by drivers for each of the six categories. Additional dependent variables were fixation duration, driver's speed at the time of first fixation, and fixation distance (i.e., the distance at which the sign was first fixated). This last parameter was recently used by Costa et al. (2018a,b) considering road signs, finding that the distance or first glance was in mean 51 m. Fixation distance allows the identification of the specific road location in which the potential distraction effect can occur. Furthermore, fixation distance could be considered an indirect measure of sign conspicuity.

A further aim of the study was to relate the fixation patterns to a set of parameters that characterized advertising sign placement, size, and design. Specifically, for each of the 154 advertising signs included in this study we considered clearance from the carriageway, elevation from the road, sign height, width, surface area, text character count, and big, medium, and small text character counts.

We hypothesized that frequency of fixation to advertising signs would mirror the results of frequency for road signs as found in Costa et al. (2014), and that fixation frequency would be enhanced by a lower clearance from the carriageway, a higher elevation, size, and the amount of text. Furthermore, we hypothesized that billboards would receive higher visual attention than vendor signs, and that vendor signs would attract more attention than movable display boards and directional signs. Billboards and vendor signs, in fact, tend to be greater in size and they are specifically designed to attract attention in comparison to directional signs and movable display boards that tend to be characterized by a more informative content. We also hypothesized that the amount of text embedded in the roadside advertisement would strongly influence the amount of time in which the advertisement is glanced by the driver. Considering the urban/rural context, we suggested that advertisement in a rural context would receive higher and longer fixations due to lower density of roadside advertisements in rural environments.

2. Method

2.1. Participants

Fifteen participants were involved in the study, 10 males (M_{age} = 27.1, SD = 13.08) and 5 females (M_{age} = 24.53, SD = 0.89). All participants had normal vision and none of them wore eyeglasses or lenses, since this would have excluded eye-movement recording. Participation was on a voluntary basis. Participants, blinded to the aims of the study, were informed that the experimental purpose was to test the mobile eye recording equipment in a driving context. None of the participants had a previous experience of the experimental route used in this study. All participants were Italians and had a standard Category-B driving license. Average driving experience was 6.53 years (SD = 5.41), while mean kilometers per year were 6292 (SD = 3425).

The study was approved by the Ethics Committee of the University of Bologna and an informed consent was signed by each participant prior to the participation.

2.2. Materials and procedure

A BMW series 1 car was used for all participants. Eye movements were recorded with an ASL Mobile-Eye XG equipment. Sample rate was 30 Hz with an angular precision of 1°. Vehicle kinematic data (speed, GPS positioning and acceleration) were recorded with a Video-Vbox-Pro system.

The eye-tracking system was composed by two digital high-resolution cameras, both mounted on lightweight eyeglasses. One of the camera recorded the visual scene in front of the participant while the other camera recorded the participant's eye movements (right eye). The ASL software superimposed fixation spots to the driver's visual scene in the form of a red cross with a time resolution of 33 ms (Fig. 1).

A calibration procedure was carried out to map eye movements to the driver's visual scene. The calibration took place in a parking lot while the car was stationary. Participants were requested to fixate 20 specific points, vertexes and centers of small objects in the visual scene.

The Video-Vbox-Pro system consisted of two cameras positioned on the left and right frontal areas of the car roof, and a GPS antenna that was attached to the top of the car, in central position. The system recorded speed (accuracy: ± 1 km/h) and acceleration with a 20 Hz sample rate. The kinematic data were synchronized with the video recording of the driver's visual scene.

The experimental route was a round trip of 30 km (15 + 15 km) along the B-road SP 610 "Selice-Montanara", a single carriageway, two-lane-road. Lane width was ≈2.5 m. The experimental route intersected six small towns, allowing the recording of both rural and urban segments. The context was rural in 47% of the route and urban in the 53% of the route.

The annual average daily traffic (AADT) along the examined road segment was 198 vehicles/day with a design speed of 50 km/h (urban segments), and 90 km/h (rural segments). The distinction between urban and rural segments was based on the position of entry and exit town signs. The advertising signs considered in the analysis were located on both sides of the experimental route and belonged to six different categories:

a) Billboard: roadside advertising device, freestanding or attached to a building, advertising products via words, symbols and pictorial displays.
b) Single directional sign: single directional signs for commercial activities and tourist attractions.
c) Multiple directional sign: multiple directional signs for commercial activities or tourist attractions stacked on a same pole.
d) Vendor signs: signs and advertisements installed in the forecourts of business premises to draw attention to commercial services, goods for sale, or other services available at the premises. Only vendor signs that were not parallel to the road were included, since they were directly visible for the driver.
e) Gas price LED signs: signs placed at gas stations equipped with LED displays that showed the prices of the different fuels.
f) Movable display boards: vendor commercials of small size presented through temporary displays at ground level.

Examples of the roadside advertisement categories included in the study are shown in Fig. 2. A total of 154 signs were considered. Eighty-nine (57.79%) were located on the driving side, while 65 (42.21%) were located on the opposite side. The distribution and side allocation of the different categories are reported in Table 1.

Overall mean distance between two consecutive advertising signs was 169 m (SD = 241 m). Density of advertising signs was higher in the urban sections (M = 124 m, SD = 188 m) than in the rural sections (M = 302 m, SD = 321 m): $F(1, 153) = 19.8$, $p < .001$. Distance for consecutive advertising signs was lower at the driving side (M = 236 m, SD = 302 m), than at the opposite side (M = 598 m, SD = 897 m): $F(1, 153) = 15.75$, $p < .001$, and was significantly affected by the type of roadside advertisement: $F(4, 149) = 31.26$, $p < .001$. Mean distance was 488 m (SD = 698) for billboards, 573 m (SD = 726) for vendor signs, 573 m (SD = 660) for directional signs (single and multiple directional signs), 2390 (SD = 2049) for mobile vendor signs, and 4099 m (SD = 3460) for gas price LED signs. Fig. 3 shows the positioning of all the roadside advertisements along the route differentiating for category, context (urban vs. rural) and side (driver side vs. opposite side).

For each sign the following design parameters were considered: a) clearance from road shoulder; b) elevation from road surface; c) width; d) height; e) surface area; f) character count differentiating between small, medium, and big characters; g) total character count; h) side (driver's side – opposite side); i) context (rural – urban). The main dimensional properties for the six categories are presented in Table 2.

All measures were obtained with a Bosch GLM 40 laser telemeter. Elevation was computed as the difference between the lower sign

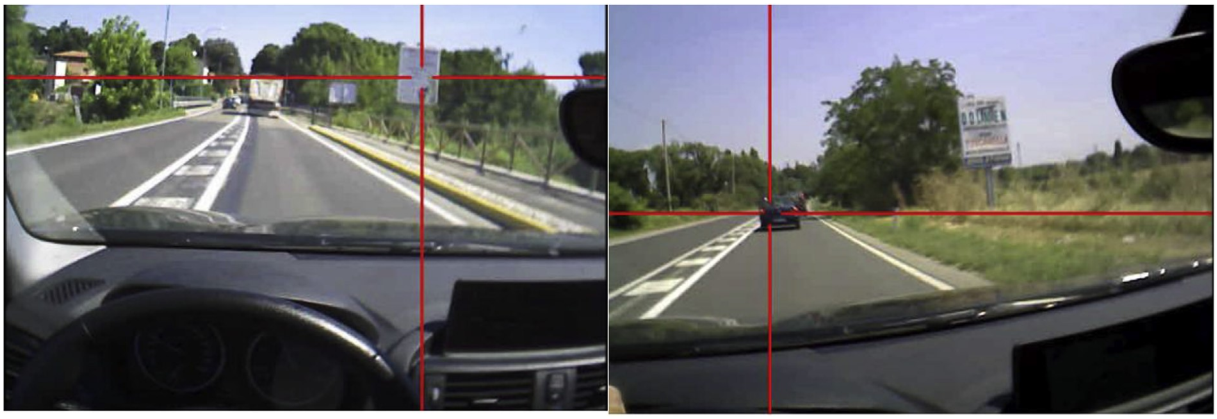

Fig. 1. Superposition of the eye tracking output to the driver's visual scene. On the left example of fixation to a roadside billboard. On the right example of billboard not fixated by the driver.

Fig. 2. Advertising sign categories included in the study.

Table 1
Distribution for side of the six advertising sign categories.

Category	Driving side	Opposite side	Total
Vendor signs	38 (42.7%)	26 (40%)	64 (41.6%)
Billboards	26 (29.3%)	24 (36.9%)	50 (32.5%)
Multiple directional signs	8 (9%)	6 (9.2%)	14 (9.1%)
Single directional signs	7 (7.8%)	4 (6.2%)	11 (7.1%)
Movable display board	7 (7.8%)	3 (4.6%)	10 (6.5%)
Gas price LED signs	3 (3.4%)	2 (3.1%)	5 (3.2%)
Total	89	65	154

margin and the carriageway level. Background colours of the roadside advertisements were: white (53.62%), black (12.31%), blue (10.14%), green (10.14%), brown (8.69%), red (2.89%), and yellow (2.21%). Single and multiple directional signs had a black background (68%) or white background (32%). Text in the advertisement area was divided in 3 categories according to its height: $\leq$ 10 cm (small), $>$ 10 cm and $<$ 20 cm (medium), $\geq$20 cm (big).

Each participant drove 1 km before starting the experimental route to familiarize with the eye recording equipment and the car. Participants had to drive following the SP 610 road for 15 km until they were given instructions to return to the starting point along the same road. They did not have to plan the route or follow specific directional signs. Data collection was carried out from 10 to 12 a.m. and from 2 to 5 p.m. to avoid high density traffic conditions. All experimental sessions were run in weather conditions that guaranteed homogeneous visibility conditions for all participants, without rain, fog, or mist.

2.3. Data analysis

Fixations to road side advertising signs were assessed by a frame-by-frame analysis of the ASL eye-tracking-output video. A traffic sign was considered as fixated if the fixation point (intersection point of the cross shown in Fig. 1) was positioned within the road sign surface area (AOI: area of interest) for a minimum duration of two frames (66 ms), to avoid the inclusion of saccadic movements. The cases of one-frame glances to the sign AOI that were discarded were 18 (3.91%). The threshold of 66 ms, which is fairly low in comparison to a common high-pass filtering of 100 ms or higher usually found in eye-tracking studies (Holmqvist et al., 2015) was dictated by the specific dynamic setting of this study that involved the recording of eye movements while driving. Velichkovsky et al. (2000), for example, reported that fixations around 60 ms made up around 7% of all fixations during a simulated driving task. In a real driving task, as the one considered in this study, car movements and the high-speed of the dynamical visual scene could result in very rapid fixations and saccades. Fixations recorded for two frames (66 ms) accounted for 5.9% of all fixations, and three-frame fixations (100 ms) accounted for 7.9% of all fixations. Previous studies clearly showed that visual stimuli exposed for intervals shorter than 100 ms could lead to a correct identification and categorization of a visual stimulus. Specifically, Costa et al. (2018a,b) showed that a 66 ms presentation lead to an accuracy rate of 80% in road sign recognition. Since road signs tend to be smaller in size than road advertising signs, *a fortiori*, such a duration could be enough for the driver to have a gist of the advertising sign content, and therefore the exclusion of these data could have resulted in a loss of significant events.

Fixation duration was assessed as the number of adjacent frames in which the fixation point was included in the sign area. Synchronization the eye tracking output video with the Video-Vbox-Pro output video allowed the assessment of the driver's speed at the time of each fixation. The distance at which an advertising sign was first fixated was computed as the difference between the position of the first fixation to the sign, and the position in which the car was perpendicular to the roadside sign.

Fixation rate for each of the six advertising sign categories and for each participant were analyzed with a repeated-measure ANOVA that tested if the fixation rate varied significantly within the six categories and if fixation rate was significantly affected by side (driver's side, opposite side), and context (urban, rural). A preliminary analysis showed that the distributions for fixation duration and fixation distance were highly asymmetrical and positively skewed. Skewness persisted also after a log-transformation. For this reason, we applied non-parametric tests (Kruskall-Wallis and Mann-Whitney) to analyze the effects of advertising sign category, context (urban vs. rural), and side (driver's side, opposite side) on both variables.

The effects of clearance from road shoulder, width, height, surface area, elevation, number of small, medium, big text characters, and total number of text characters on fixation duration and fixation distance were analyzed with multiple linear regressions. The same parameters were entered in a binomial logistic regression to test the differences between fixated and not fixated advertising signs.

Greenhouse-Geisser correction was applied in repeated-measure ANOVAs. Pairwise comparisons with Bonferroni correction was used when comparing factors with more than two levels. Effect size was reported as Cohen partial eta square when reporting ANOVA results, and as phi when reporting Chi-square results. SPSS ver. 23 was used for all statistical computations.

3. Results

3.1. Fixation rate

Advertising signs were fixated on average with a proportion of .24. As shown by the ANOVA results, fixation rate was significantly different between the 6 advertising sign categories: $F(5, 60) = 6.47$, $p = .002$, $\eta_p^2 = .35$. Mean proportions for the six categories, in decreasing order, were: 0.31 (± 0.14) for billboards, 0.27 (± 0.17) for gas price LED signs, 0.23 (± 0.08) for single directional signs, 0.23 (± 0.08) for

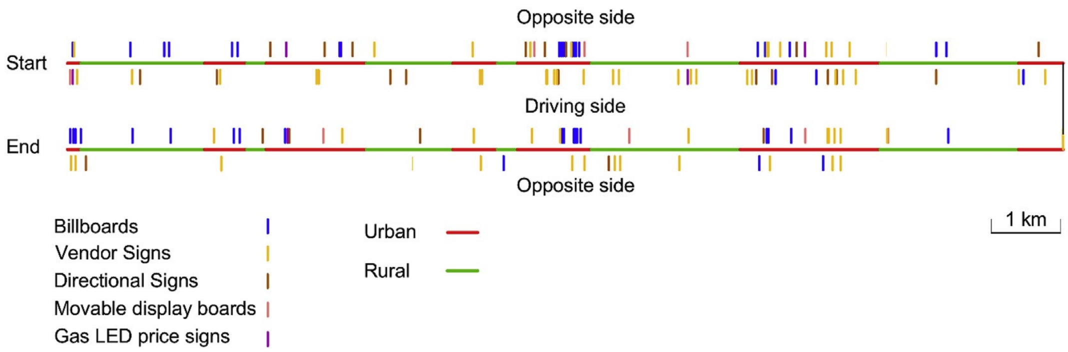

Fig. 3. Distribution of roadside advertising signs along the experimental route, differentiating for category, context (urban vs. rural), and side (driving side vs. opposite site). All the distances are in scale. The vertical black line shows the turning point between the outbound and the return section.

Table 2
Dimensional properties (*M* and *SD*) of the six advertising sign categories.

	Clearance (cm)	Elevation from the road (cm)	Surface area (m^2)	Characters #
Vendor signs	289 (138)	259 (120)	1.33 (1.48)	34 (60)
Billboards	298 (82)	152 (48)	2.31 (0.9)	93 (159)
Multiple directional signs	283 (100)	146 (34)	1.04 (0.79)	63 (61)
Single directional signs	235 (113)	178 (18)	0.29 (0.12)	13 (5)
Movable display boards	324 (277)	24 (9)	0.43 (0.12)	53 (16)
Gas price LED signs	370 (115)	450 (72)	6.3 (0.6)	64 (13)
Mean	290 (125)	198 (111)	1.57 (1.41)	53 (101)

vendor signs, 0.16 (± 0.09) for multiple directional signs, and 0.12 (± 0.12) for movable display boards.

Pairwise comparisons between the single categories showed a significant difference between billboards and vendor signs (p = .02), billboards and movable display boards (p < .001), vendor signs and movable display boards (p = .001). Advertising signs on the driver's side were looked at in a significantly higher proportion (M = 0.27) in comparison to signs located on the opposite side (M = 0.21): χ^2 = 8.71, p = .003, ϕ = 0.41.

A binomial logistic regression tested the differences between fixated and not fixated signs on these variables: clearance from the road, elevation from the road, width, height, surface, big, medium, and small character counts, and total number of characters. The results are reported in Table 3. The binomial logistic regression model was significant: $\chi^2(8)$ = 26.82, p = .001. The model explained 7.2% of the variance (Nagelkerke R^2) of fixation rate, and correctly classified 76.1% of cases. The six advertising signs that received the highest fixation rates (0.67–0.58) are shown in Fig. 4. With the exclusion of the gas price LED sign, they are all characterized by a rich text content and medium-large size. Furthermore, they are all positioned near the carriageway shoulder.

3.2. Fixation duration

The distribution of fixation duration is reported in Fig. 5. The distribution was highly asymmetrical and positively skewed with a kurtosis of 4.63 and an asymmetry of 2.09. Shapiro-Wilk normality test was significant (p < .001), showing that the distribution was not normal. Mode value was 99 ms, median value was 297 ms. The effect of advertisement sign category on fixation duration was tested with a Kruskal-Wallis test that was not significant (p < .45). Median fixation time for the different categories were: 330 ms (SD = 502) for billboards, 264 ms (SD = 414) for vendor signs, 264 ms (SD = 102) for gas price LED signs, 231 ms (SD = 414) for movable display board, 297 ms (SD = 357) for single directional signs, and 280 ms (SD = 338) for multiple directional signs.

A Mann-Whitney test checked the difference between fixations in urban versus rural context and the difference was significant (*U = 17318, p* = .03). Fixation duration was significantly higher in the rural context (Mdn = 363 ms) than in the urban context (Mdn = 264 ms).

The side of the road in which the sign was positioned was not critical in modulating fixation duration (p = .37).

In a multiple linear regression, we tested the influence of speed, clearance from road shoulder, elevation, height, width, surface area, number of characters on fixation duration. The regression model was significant: F(11, 388) = 3.07, p = .001, Adjusted R^2 = 0.05. The significant variables were speed (β = -0.15, t = −2.74, p = .006), elevation (β = -0.17, t = −2.86, p = .004), and medium size character count (β = *.13*, t = 1.99, p = .04).

Fig. 6 shows the six advertisement signs that received the longest fixation times (> 2 s). Their design was mainly textural, with a remarkable number of words, and their size tended to be high.

Mean speed at the time of first fixation was 55.54 km/h (SD = 11.59). Speed was significantly different between urban-rural segments: (F(1, 434) = 50.82, p < .001, η_p^2 = .11). Mean speed in urban areas was 54.77 km/h (SD= 10.81), while in rural areas it was 64.47 km/h (SD = 11.71).

3.3. Fixation distance

Fixation distance density plot is shown in Fig. 7. The distribution was strongly positively skewed with a kurtosis of 4.29 and an asymmetry of 1.46. Shapiro-Wilk normality test was significant (p < .001), showing that the distribution was not normal. Median fixation distance was 58.10 m. A Kruskal-Wallis test was used for testing the influence of roadside advertising category on fixation distance. The test was significant: χ^2 = 55.91, p < .001. Median distances for the different categories, in decreasing order, were: gas price LED signs 73.16 m (SD = 31.50), billboards 69.9 m (SD = 38.76), vendor signs 57.25 m (SD = 35.72), multiple directional signs 42.15 m (SD = 22.17), movable display board 37.3 m (SD = 22.09), single directional signs 34.25 m (SD = 20.81). Median values and distribution of fixation distance for the six categories is shown in Fig. 8. Side of the road was also significant (Mann-Withney U(1) = 26.39, p < .001). Signs positioned on the driver's side were fixated at a shorter distance (Mdn = 51.99 m, SD = 35.72) than signs that were positioned in the opposite direction (Mdn = 67.75 m, SD = 37.13). Fixation distance was not significantly different in urban and rural settings.

Table 3
Descriptive statistics (M and SD) and results of the binary logistic regression that tested advertising sign size and text differences between fixated and not fixated signs.

	Fixated signs	Not-fixated signs	χ^2	p
Clearance from the road (cm)	270.02 (144.77)	293.28 (169.83)	14.38	< .001
Elevation (cm)	172.09 (115.87)	174.96 (122.87)	1.05	ns
Width (cm)	124.54 (60.60)	105.72 (66.09)	11.34	= .001
Height (cm)	123.39 (88.61)	100.17 (82.77)	3.19	ns
Surface(m^2)	1.77 (1.55)	1.31 (1.42)	.37	ns
Characters ≥ 20 cm (#)	3.15 (5.34)	2.30 (4.78)	4.90	= .02
Characters > 10 cm and < 20 cm (#)	13.62 (22.69)	9.55 (16.91)	6.82	= .009
Characters ≤ 10 cm (#)	34.82 (46.41)	43.23 (37.35)	5.99	= .01
Total characters (#)	51.37 (47.95)	55.09 (39.80)	3.21	= .01

Fig. 4. The six advertising signs that received the highest fixation rate. In brackets the mean proportion of fixations for each sign.

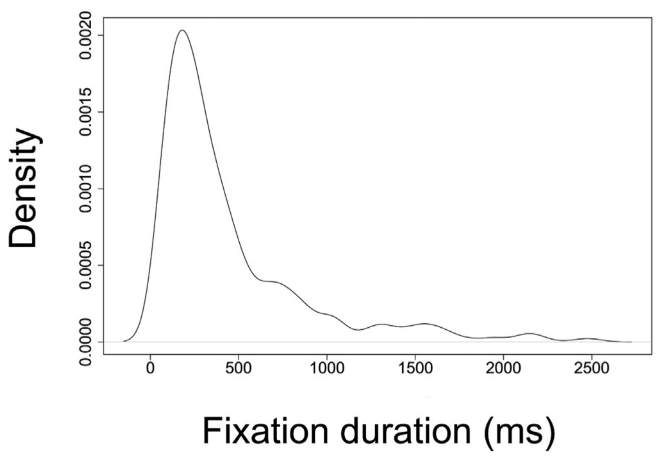

Fig. 5. Density histogram for fixation duration for all the advertising sign categories.

The influence of speed, clearance from the road, elevation, height, width, surface, character count on fixation distance was tested with a multiple linear regression. The model was significant: $F(11, 388) = 14.78$, $p < .001$, $R^2 = 0.27$. The significant variables in the model were speed ($\beta = 0.25$, $t = 5.44$, $p < .001$), width ($\beta = -0.20$, $t = -2.01$, $p = .04$), surface ($\beta = 0.54$, $t = 2.68$, $p = .008$), total number of characters ($\beta = 0.16$, $t = 2.05$, $p = .04$), and medium size character count ($\beta = 0.18$, $t = 3.19$, $p = .002$).

4. Discussion

The visual complexity of roadsides could contribute to attract the driver's attention, diverting it from the driving and resulting in safety issues that should be assessed and controlled. This study assessed the impact of different categories of advertising signs on the driver's visual behavior. Considering the whole sample of 154 advertising signs, the mean fixation rate was 24%, a proportion that closely matched the mean fixation rate for vertical traffic signs of 25.06% found by Costa et al. (2014). While previous research focused almost exclusively on billboards, in our study we assessed visual exploratory parameters to

Fig. 6. The six advertising signs that received the longest fixations. In brackets the mean fixation time (± *SD*).

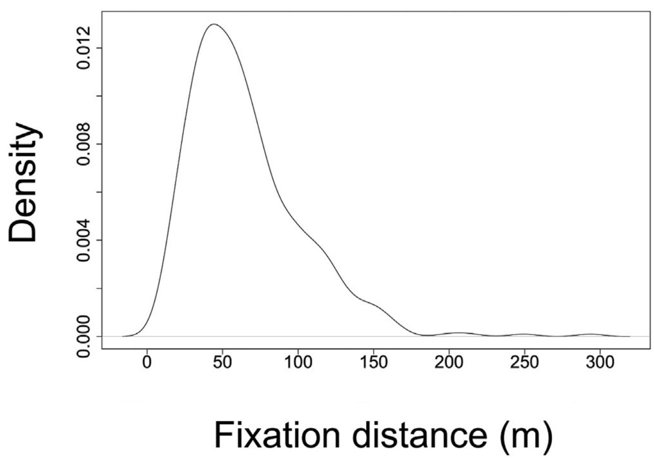

Fig. 7. Fixation-distance density plot.

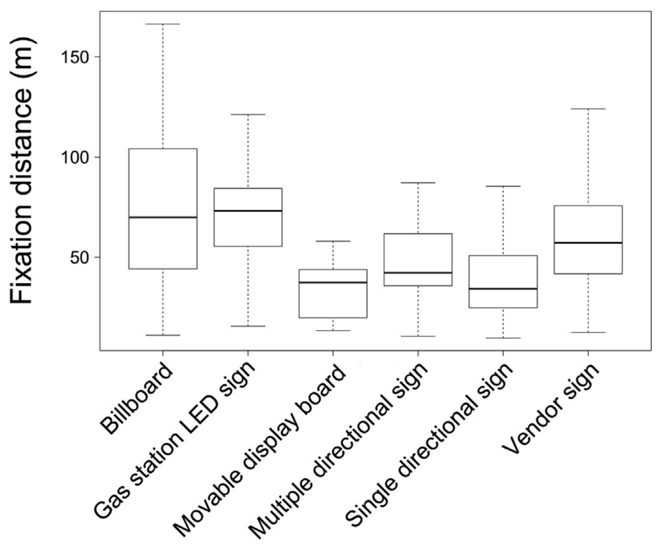

Fig. 8. Fixation-distance boxplots for the six roadside advertisement categories.

other typologies of road advertisements such as vendor signs, gas price LED signs, movable display boards, and commercial directional signs (both single and multiple). Fixation rate was highest for billboards (0.31) and was progressively lower for gas price LED signs (0.27), single directional signs (0.23), vendor signs (0.23), multiple directional signs (0.16), and movable display boards (0.12). This effect could be explained by the higher conspicuity and saliency of billboards in comparison to vendor signs, probably due to their higher size, and the higher conspicuity and saliency of vendor signs in comparison to movable display boards and directional signs. In addition, billboards and vendor signs tended to have a higher textual content with large-size characters, and the amount of large-size text included in a roadside advertisement was one of the best predictors of both fixation rate and fixation duration.

Advertising signs that were placed on the driving side were fixated more (0.27) than the signs placed on the opposite side (0.21), and the lower was the lateral offset of the advertisement from the road the higher was the proportion of fixations. A higher lateral offset of the roadside advertisement implied that the driver had to make a large saccadic movement to foveate the sign, and that the sign was firstly processed in a more peripheral visual field, where vision accuracy is lower (Costa et al., 2018a,b; Crundall et al., 2002).

This result was also confirmed by the multiple linear regression in which clearance from the road was a significant predictor of fixation rate. Other significant predictors of advertising sign fixation were related to the amount of text included in the sign. The amount of text with characters greater than 10 cm induced a higher fixation rate, whereas the amount of small characters (≤ 10 cm) and in general long texts were related to a lower fixation rate. Advertising signs with long textual display probably fail to attract the driver's visual attention due to the poor readability in a dynamic context.

The distribution of fixation duration for advertising signs was strongly asymmetrical and positively skewed, with a median value of 297 ms. This mean duration was higher in comparison to the mean fixation duration of 154 ms for traffic signs found in Costa et al. (2014). This might be explained by the higher visual complexity of advertising signs in comparison to traffic signs. The distribution had a long tail for long fixation intervals. For example, 24.8% of fixations on a single sign had a total duration of more than 500 ms, 16.1% exceeded 750 ms, 9.8% exceeded 1 s, and 1.5% exceeded 2 s. This is in line with the results of the review by Decker et al. (2015) who found that about 10–20% of all glances towards billboards lasted more than 750 ms. Considering that 750 ms has been suggested as the minimum perception-reaction time for a vehicle slowing ahead of the driver (Smiley et al., 2004), and that 2 s or longer eyes-off-the-road interval is strongly associated with motor vehicle collisions and other traffic incidents (Klauer et al., 2006), we can conclude that there is a discrete amount of cases in which distraction induced by roadside advertisement could adversely impact traffic safety.

Fixation duration was not affected by the advertising sign category and was inversely related to speed and elevation from the road and positively related to the number of medium size characters. Longer fixations to advertising signs positioned near the road level is in line with the results of Crundall et al. (2006), who compared in a simulator study visual patterns for street-level advertisements to raised-level advertisements, finding that street-level advertisements received longer fixations.

Fixation distance exhibited also a positively skewed distribution with a predominance of short distances and a long tail for long distances. Median value for fixation distance was 58.10 m, a value that strongly mirrored the 51 m distance found for traffic signs (Costa et al., 2018a,b). This result should be weighted considering that the average speed was 55.54 km/h. Interestingly, fixation distance differed according to the advertising sign category, being highest for gas price LED displays (73.16 m), and progressively lower for billboards (69.9 m), vendor signs (57.27 m), multiple directional signs (42.10 m), movable display boards (37.3 m), and single directional signs (34.25 m), mirroring the size of the signs in the different categories. The distance at which advertising signs were fixated increased linearly with speed, sign size, and text length. These results could also help designers to optimize advertising sign layout and graphical content since distance and speed significantly affect the level of details and the amount of information that can be retained by a driver.

It is important to state that an eye fixation to an advertising sign does not necessarily implies a driver's explicit awareness of the sign content. Foveal fixation is necessary for a detailed vision, but in many cases is not sufficient for awareness (Luoma, 1984; Schütz et al., 2011). This aspect is further supported considering that the median value for fixation duration was quite short (297 ms), and that fixations were recorded in a highly dynamic context. Furthermore, in this study the probability to attract eye fixations was assessed on a macro level, considering the whole sign area. Future studies with the use of eye recording devices with higher spatial and temporal resolution could investigate how the advertisement image features (i.e., luminance, contrast, color, text positioning and style) could affect the driver's scan path when exploring roadside advertisements. These studies could complement those that have modeled the influence of image features on fixation selection in naturalistic scenes (Nuthmann and Einhäuser, 2015; Nuthmann et al., 2017).

External distraction accounts for approximately 6–9% of all motor-vehicle collisions, and approximately 4% of all motor-vehicle collisions are caused by driver's errors (NHTSA, 2016). In a review of laboratory experiments, Wallace (2003) suggested that distraction accounts for roughly between 10% and 30% of all accidents. Therefore, the study of driver's visual behavior in response to sources of external distraction as advertising signs is of primary importance. In our study we showed that, although the proportion of advertising signs that are actually fixated was relatively low (0.24), and that the median of glance duration was rather short (297 ms), there was a significant amount of "long" fixations that in the circumstance of an immediate reaction required by the driver, could pose serious problems for traffic safety. Analyses of the 100 Car Study of high-mileage drivers clearly indicated that longer eye glances off the road were associated with a greater risk of accidents (Dingus et al., 2006). Specifically, eye glances away from the forward roadway of 2 s duration doubled the risk of a crash or near crash (Klauer et al., 2006; Klauer et al., 2010). In our case, we had 1.1% of advertising signs with a mean glance duration higher than 2 s, and 10% of the advertising signs received a glance higher than 1 s, a distraction interval that could adversely impact traffic safety, especially in urban contexts, where the visual scene tends to be more complex. For example, among the six advertising signs that attracted the higher fixation duration we recorded two vendor signs and a gas price led sign.

In this research, we have highlighted that advertising signs other than billboards could have a significant distracting potential. Vendor signs, in particular, tend to be more frequent than billboards, and in many cases their size, visual complexity, and textual content is higher, determining a serious distraction source for drivers. They are mainly distributed in urban areas and the glance duration to vendor signs (264 ms) was similar to that of billboards (330 ms), and therefore they are a concrete source of distraction for drivers.

Advertising signs are the result of a difficult balancing between the opposite requirements of traffic safety and advertising effectiveness. As far as road safety is concerned, advertising signs are a form of external distraction with potential detrimental effects on driving-relevant tasks. To the contrary, advertisement efficacy is proportional to the amount of attention that is captured from the driver. When assessing the distracting potential of an advertising sign (i.e., fixation rate and fixation duration) the critical factors that emerged from this study were the clearance from the road and the amount of text included in the advertising sign. Distraction increases when the advertising sign is placed near the road. The text included in the advertisement induce the driver to read, an activity that requires much more time than capturing the graphical content. Examining the advertising signs that received the highest proportions of fixation rate and fixation duration, as shown in Figs. 5 and 6, it is clear that the majority were characterized by a conspicuous textual content. National highway regulations should incorporate the results of this study establishing limits for size, placement and character content that could help limiting the potential distracting effect of the various roadside advertising signs. Regulations should focus not only on billboards but also on vendor signs, commercial directional signs, gas price LED displays, and movable display boards since all compete with the driver's attention undermining traffic safety.

Acknowledgement

The research was supported by the Alma Idea 2017 Costa grant from the University of Bologna.

References

Ady, R., 1967. An investigation of the relationship between illuminated advertising signs and expressway accidents. Traffic Saf. Res. Rev. 3, 9–11.
Beijer, D., Smiley, A., Eizenman, M., 2004. Observed driver glance behavior at roadside advertising signs. Transport. Res. Rec.: J. Transportat. Res. Board 1899, 96–103. https://doi.org/10.3141/1899-13.
Belyusar, D., Reimer, B., Mehler, B., Coughlin, J.F., 2016. A field study on the effects of digital billboards on glance behavior during highway driving. Accid. Anal. Prev. 88, 88–96. https://doi.org/10.1016/j.aap.2015.12.014.
Bendak, S., Al-Saleh, K., 2010. The role of roadside advertising signs in distracting drivers. Int. J. Ind. Ergon. 40 (3), 233–236. https://doi.org/10.1016/j.ergon.2009.12.001.
Chan, M., Singhal, A., 2013. The emotional side of cognitive distraction: implications for road safety. Accid. Anal. Prev. 50, 147–154. https://doi.org/10.1016/j.aap.2012.04.004.
Costa, M., Bonetti, L., Bellelli, M., Lantieri, C., Vignali, V., Simone, A., 2016. Reflective tape applied to bicycle frame and conspicuity enhancement at night. Hum. Factors 59 (3), 485–500. https://doi.org/10.1177/0018720816677145.
Costa, M., Bonetti, L., Vignali, V., Lantieri, C., Simone, A., 2018a. The role of peripheral vision in vertical road sign identification and discrimination. Ergonomics. https://doi.org/https://doi.org/10.1080/00140139.2018.1508756.
Costa, M., Simone, A., Vignali, V., Lantieri, C., Bucchi, A., Dondi, G., 2014. Looking behavior for vertical road signs. Transport. Res. F Traffic Psychol. Behav. 23, 147–155.
Costa, M., Simone, A., Vignali, V., Lantieri, C., Palena, N., 2018b. Fixation distance and fixation duration to vertical road signs. Appl. Ergon. 69, 48–57. https://doi.org/10.1016/J.APERGO.2017.12.017.
Crundall, D., Underwood, G., 1998. The effects of experience and processing demands on visual information acquisition in drivers. Ergonomics 41, 448–458.
Crundall, D., Underwood, G., Chapman, P., 2002. Attending to the peripheral world while driving. Appl. Cognit. Psychol. 16, 459–475.
Crundall, D., Van Loon, E., Underwood, G., 2006. Attraction and distraction of attention with roadside advertisements. Accid. Anal. Prev. 38 (4), 671–677. https://doi.org/10.1016/j.aap.2005.12.012.
Decker, J.S., Stannard, S.J., McManus, B., Wittig, S.M.O., Sisiopiku, V.P., Stavrinos, D., 2015. The impact of billboards on driver visual behavior: a systematic literature review. Traffic Inj. Prev. 16, 234–239. https://doi.org/10.1080/15389588.2014.936407.
Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., 2006. The 100-car Naturalistic Driving Study, Phase II - Results of the 100-car Field Experiment. National Highway Traffic Safety Administration, Washington, DC.
Edquist, J., Horberry, T., Hosking, S., Johnston, I., 2011. Effects of advertising billboards during simulated driving. Appl. Ergon. 42 (4), 619–626. https://doi.org/10.1016/j.apergo.2010.08.013.
Elliot, A.J., Maier, M.A., 2014. Color psychology: effects of perceiving color on psychological functioning in humans. Annu. Rev. Psychol. 65, 95–120. https://doi.org/10.1146/annurev-psych-010213-115035.
Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., van de Weijer, J., 2015. Eye Tracking. Oxford University Press, Oxford.
Horrey, W.J., Wickens, C.D., Consalus, K.P., 2006. Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. J. Exp. Psychol. Appl. 12 (2), 67–78. https://doi.org/10.1037/1076-898X.12.2.67.
Italian Highway Code, 1992. D.L. 285, April 30th 1992.
Kaber, D., Pankok, C., Corbett, B., Ma, W., Hummer, J., Rasdorf, W., 2015. Driver behavior in use of guide and logo signs under distraction and complex roadway conditions. Appl. Ergon. 47, 99–106. https://doi.org/10.1016/j.apergo.2014.09.005.
Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J., Ramsey, D.J., 2006. The Impact of Driver Inattention on Near-Crash/crash Risk: an Analysis Using the 100-car Naturalistic Driving Study Data. NHTSA, Washington, DC.
Klauer, S.G., Guo, F., Sudweeks, J.D., Dingus, T.A., 2010. An Analysis of Driver Inattention Using a Case-Crossover Approach on 100-car Data: Final Report. U.S. Department of Transportation: National Highway Traffic Safety Administration.
Land, M.F., Lee, D.N., 1994. Where we look when we steer. Nature 369 (6483), 742–744. https://doi.org/10.1038/369742a0.
Land, M.F., Tatler, B.W., 2009. Looking and Acting: Vision and Eye Movements in Natural Behaviour. Oxford University Press, Oxford.
Lansdown, T.C., Stephens, A.N., Walker, G.H., 2015. Multiple driver distractions: a systemic transport problem. Accid. Anal. Prev. 74, 360–367. https://doi.org/10.1016/j.aap.2014.07.006.
Lantieri, C., Lamperti, R., Simone, A., Costa, M., Vignali, V., Sangiorgi, C., Dondi, G., 2015. Gateway design assessment in the transition from high to low speed areas. Transport. Res. F Traffic Psychol. Behav. 34, 41–53. https://doi.org/10.1016/j.trf.2015.07.017.
Lesley, G., 1995. Enhancing the daytime conspicuity of pedestrians through the usage of fluorescent materials. Color Res. Appl. 20 (2), 117–123. https://doi.org/10.1002/col.5080200207.
Luoma, J., 1984. Perception and eye movements in simulated traffic situations. Acta Ophthalmol. 161, 128–134.
Marciano, H., Yeshurun, Y., 2012. Perceptual load in central and peripheral regions and its effects on driving performance: advertising billboards. Work 41 (Suppl. 1), 3181–3188. https://doi.org/10.3233/WOR-2012-0580-3181.
Megías, A., Maldonado, A., Catena, A., Di Stasi, L.L., Serrano, J., Cándido, A., 2011. Modulation of attention and urgent decisions by affect-laden roadside advertisement in risky driving scenarios. Saf. Sci. 49 (10), 1388–1393. https://doi.org/10.1016/j.ssci.2011.06.001.
Mourant, R.R., Rockwell, T.H., 1972. Strategies of visual search by novice and experienced drivers. Hum. Factors 14, 325–335.
NHTSA, 2016. Traffic Safety Facts: Research Note. NHTSA, Washington, DC.
Nuthmann, A., Einhäuser, W., 2015. A new approach to modeling the influence of image features on fixation selection in scenes. Ann. N. Y. Acad. Sci. https://doi.org/10.1111/nyas.12705.
Nuthmann, A., Einhäuser, W., Schütz, I., 2017. How well can saliency models predict fixation selection in scenes beyond central bias? A new approach to model evaluation using generalized linear mixed models. Frontiers in Human Neuroscience. https://doi.org/10.3389/fnhum.2017.00491.
Rayner, K., Rotello, C.M., Stewart, A.J., Keir, J., Duffy, S.A., 2001. Integrating text and pictorial information: eye movements when looking at print advertisements. J. Exp. Psychol. Appl. 7 (3), 219–226. https://doi.org/10.1037/1076-898X.7.3.219.
Schütz, A.C., Braun, D.I., Gegenfurtner, K.R., 2011. Eye movements and perception: a selective review. J. Vis. 11 (5), 1–30. https://doi.org/10.1167/11.5.9.
Smiley, A., Smahel, T., Eizenman, M., 2004. Impact of video advertising on driver fixation patterns. Transport. Res. Rec. 1899 (1), 76–83. https://doi.org/10.3141/1899-10.
Topolšek, D., Areh, I., Cvahte, T., 2016. Examination of driver detection of roadside traffic signs and advertisements using eye tracking. Transport. Res. F Traffic Psychol. Behav. 43, 212–224. https://doi.org/10.1016/j.trf.2016.10.002.
Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., Crundall, D., 2003. Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. Ergonomics 46 (6), 629–646. https://doi.org/10.1080/0014013031000090116.
Velichkovsky, B.M., Dornhoefer, S.M., Pannasch, S., Unema, P.J.A., 2000. Visual fixations and level of attentional processing. In: Proceedings of the Symposium on Eye Tracking Research & Applications - ETRA '00, pp. 79–85. https://doi.org/10.1145/355017.355029.
Wallace, B., 2003. Driver distraction by advertising: genuine risk or urban mith? Municip. Eng. 156 (3), 185–190.
Wickens, C.D., 2007. Attention to attention and its applications: a concluding view. In: Kramer, A.F., Wiegmann, D.A., Kirlik, A. (Eds.), Attention: from Theory to Practice. Oxford University Press, Oxford.
Wickens, C.D., 2008. Multiple resources and mental workload. Hum. Factors: J. Hum. Factors Ergonom. Soc. 50 (3), 449–455. https://doi.org/10.1518/001872008X288394.
Wickens, C.D., McCarley, J.S., 2008. Applied Attention Theory. CRC Press, Boca Raton.
Young, M.S., Mahfoud, J.M., Stanton, N.A., Salmon, P.M., Jenkins, D.P., Walker, G.H., 2009. Conflicts of interest: the implications of roadside advertising for driver attention. Transport. Res. F Traffic Psychol. Behav. 12 (5), 381–388. https://doi.org/10.1016/j.trf.2009.05.004.

Transportation Research Part F 60 (2019) 770–782

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

T-junction priority scheme and road user's yielding behavior

Marco Costa [a,*], Arianna Bichicchi [b], Mattia Nese [a], Claudio Lantieri [b], Valeria Vignali [b], Andrea Simone [b]

[a] Department of Psychology, University of Bologna, Italy
[b] Department of Civil, Chemical, Environmental, and Material Engineering, University of Bologna, Italy

ARTICLE INFO

Article history:
Received 2 July 2018
Received in revised form 5 December 2018
Accepted 11 December 2018
Available online 20 December 2018

Keywords:
Yielding behavior
T-junction
Affordance
Intersection design
Priority
Road security

ABSTRACT

Four studies investigated yielding behavior in yield-controlled T-junctions that differed for two priority schemes. In one case road users in the intersecting arm had to give way to road users in the straight arm (priority to straight arm). In the other case road users in the straight arm had to give way to road users approaching from the intersecting arm (priority to intersecting arm). In two studies, yielding behavior was assessed with approaching speed and gaze behavior to the critical areas of the intersection. Two additional studies monitored road users' speed and eye movements approaching the intersection. The results of the two behavioral studies showed a significant speed reduction and an increase of driver's visual inspection to the intersection area in the priority-to-straight-arm condition in comparison to the priority-to-intersecting-arm condition. The eye movement analysis showed that total fixation time towards the intersection critical area and horizontal eye movements were significantly higher in the priority-to-straight-arm condition. The results emphasize the importance of considering perceptual affordances and expectations for priority in intersection design to increase drivers' compliance to yielding rules.

1. Introduction

Failing to yield at intersection is one of the most frequent cause of road accident. According to Chai, Wong, and Wang (2017) about one in four (23.4%) road traffic accidents in Singapore occurred for failure to yield at signalized intersections. Schepers, Kroeze, Sweers, and Wüst (2011) reported that 95% of the bicycle–motor vehicle crashes in the Netherlands that occurred at unsignalized priority intersections were caused by failure-to-yield.

Many accidents caused by the driver not looking in the appropriate direction ("failed to look") could have a primary cause in a driver's erroneous perception of the priority scheme of a specific intersection due to the violation of priority affordances in the intersection design. Hence, it appears particularly relevant for improving road safety to investigate how yielding behaviour can be influenced by intersection design.

Failure to yield being a behaviour is to be intended as a secondary cause of an accident, since the cause of this behaviour could be traced in perceptual and cognitive processes of the driver. Successful driving in intersections requires a large amount of visual and mental resources (Hills, 1980), and failure to yield the right of way could be caused by distraction, inattention, temporary or permanent sight obstruction, insufficient knowledge of the priority rules, and faulty diagnosis – information failure, between driver and environment. The intersection design and the priority scheme assignment to the arms of

* Corresponding author at: Department of Psychology, University of Bologna, Viale Berti Pichat, 5, 40127 Bologna, Italy.
E-mail address: marco.costa@unibo.it (M. Costa).

https://doi.org/10.1016/j.trf.2018.12.009
1369-8478/

the intersection could facilitate or hinder the acquisition and processing of information about priority, and the study of this interaction between intersection geometry and priority assignment is the main aim of this study.

Priority is regulated in most countries by different rules. In addition to a general priority-to-the-right (or to-the-left) rule in uncontrolled intersections, priority is regulated by the stop/yield signs, road markings, or by traffic lights. Before the intersection, warning and priority signs usually inform the driver about the priority scheme that will regulate the intersection. These vertical signs, however, due to their lateral placement and the automaticity of driving behavior in expert drivers are very often unattended (Costa et al., 2014; Costa, Bonetti, Vignali, Lantieri, & Simone, 2018).

Priority perception in this paper is investigated according to the self-explaining road approach. In this framework road design and traffic control is based on perceptual affordances that could facilitate per se correct expectations and behaviors of road users, to create a safe and user-friendly road network (Rothengatter, 1999; Theeuwes & Godthelp, 1995; Wegman, Zhanga, & Dijkstra, 2012; Weller, Schlag, Friedel, & Rammin, 2008). Self-explanation is enhanced when a design is strongly based on an affordance analysis of its components. Affordance, a concept firstly introduced by Gibson (1979), and further developed by Norman (2004, 2013), refers to action and function possibilities that are readily perceivable in an object (Jones, 2003). Road elements can function as affordances that serve as built-in instructions that can guide driving behaviour, either implicitly or explicitly (Walker, Stanton, & Chowdhury, 2013; Weller et al., 2008). They include road markings, delineated lane width, roadside objects, pavement evenness, but also more intrinsic properties such as the geometry of road intersections (Charlton, 2007; Elliot, Mccoll, & Kennedy, 2003; Weller et al., 2008). For example, narrow lanes and street vegetation close to road shoulders have been shown to reduce vehicle speed by reducing the perceived road width (Ewing & Dumbaugh, 2009). The impact of these perceptual countermeasures could be affected by the context. Specifically, in this study we tested an affordance that associate road linearity to the perception of priority in intersections. We suggested that having to yield along a straight road would lead to unsafe behaviors, whereas having to yield at an intersecting road would prompt a much safer behavior.

When roads design violates self-explaining rules road users' behavior could often conflict with formal traffic rules (Björklund & Åberg, 2005). For example, if the transition from a high-speed to low-speed area is not emphasized by specific markings a hysteresis emerges in which drivers tend to extend driving speed and expectations appropriate for the high-speed segment when they have already entered an urban area with low-speed requirements. In this context Lantieri et al. (2015) tested the efficacy of gateways in the transition from rural to urban road segments. The use of chicane, central island, dragon teeth, and an extended town sign helped the drivers to reduce speed when entering the urban area, resulting in a more intuitive transition from a high-speed to a low-speed zone.

In this study we investigated yielding behavior in both drivers and bicyclists in T-junctions differing in their priority assignment. A T-junction is an at-grade three-way intersection between three road segments (arms) where two arms belong to a straight road. We decided to focus on T-junctions because, when they are controlled by a yield or stop sign, priority could be assigned according to two very different and asymmetric schemes: in one case priority could be assigned to the vehicles approaching from the straight through road (priority-to-straight-arm condition) and in the other case priority is assigned to vehicles approaching from the intersecting arms (priority-to-intersecting-arm condition). We suggested that a straight road could prompt an affordance of having priority, whereas a road that intersects another road could elicit an affordance of having to yield. Following this hypothesis the priority-to-straight-arm condition matches our suggestions inducing a perception of having to yield that would result in a safe driver's behavior (i.e., speed reduction, enhanced visual attention to the intersection). To the contrary, the priority-to intersecting-arm condition mismatches our suggestions inducing a perception of having priority that would clash with the actual need to yield. We hypothesized that this mismatch would result in a driver's unsafe behavior when approaching the intersection (i.e., higher speed and lower visual inspection of the intersection area).

An uncontrolled T-junction with priority to right is particularly problematic (Helmers & Åberg, 1978) when the major road is attributed to the intersecting arm. In this case, a road user driving through the straight road tends to have a high perception of priority that results in a higher rate of failures to yield to the road users coming from the intersecting arm. Symmetrically, road users approaching from the connecting arm and turning to right tend to yield to road users coming from the left arm also if they have priority (Helmers & Åberg, 1978).

When a road user has a perception of priority, horizontal head movements tend to be small in amplitude, and the driver's speed tends to be relatively high in comparison to conditions where the road user has the perception of not having priority. The same issue was investigated by Johannessen (1984) who found that in uncontrolled T-junctions a vehicle traveling in the straight arms yielded to an upcoming vehicle from the right connecting arm only in 56% of the situations.

Especially in uncontrolled intersection the behaviour of other road users can prime expectations on priority perception. Slowing down or accelerating, for example, and the vehicle position are primary cues for interpreting a driver's intention to yield or not to yield, as showed by Janssen, van der Horst, Bakker, and ten Broeke (1988) in car-car and car-bicycle interactions.

The mismatch between affordance expectations and formal rules was investigated also by Björklund and Åberg (2005). In their research, a sample of 1276 Swedish drivers responded to questions about how often they would yield to another driver in ten hypothetical crossing situations. In all crossing situations priority was to the right. The results showed that a wider road was associated to a higher perception of having priority. In T-junctions where a driver was approaching along the straight road, 40% of the respondents reported they would never give way to a vehicle approaching from the right (connecting arm). Many respondents (49%) also reported that they would frequently or always yield to a vehicle approaching the intersection from the left along the straight road, even if they had the right of way.

As in Björklund and Åberg (2005), we investigated driver's yielding behavior in intersections but instead of using virtual scenarios, we observed driver's in real conditions, and instead of focusing on uncontrolled intersections we focused on controlled intersection with the yield sign and the pavement markings because controlled intersections with priority signs are much more common. We compared two priority arrangements in T-junctions. In one case road users approaching the intersection from the connecting arm had to yield to road users approaching from the straight road (priority to the straight arm). We considered this condition as safer because yielding the right of way matched with the physical stop of the road at the intersection and the need for the driver to turn left or right. In the other condition, the road users in the straight road had to yield to the road users approaching from the connecting arm (priority to intersecting arm). This condition was considered as unsafe and problematic because yielding the right of way did not match with the straight geometry of the road that continued through the intersection. Having to yield is associated to a slowing down behavior, and the intersection geometry could facilitate this behavior with curves and irregularities in the road geometry, and, to the contrary, inhibit this behavior with a straight and regular road geometry.

Drivers' and bicyclists' behavior approaching the intersection when it was free from other vehicles was assessed using both behavioural parameters (speed and head turning), and eye movements (total fixation time and mean angular eye movements on the horizontal plane). We expected that in the priority-to-straight-arm condition road users would approach the intersection with a lower speed and would exhibit a higher frequency of head turning, eye glances, and horizontal eye movements toward the intersection area where potential road users could come.

In the first two studies we examined two T-junctions in an urban context with a mixed flow of cars and bicyclists while in the last two studies we examined two T-junctions located in a more peripheral context that included only a vehicular flow. The two T-junctions were matched for traffic volume.

2. Study 1

2.1. Method

2.1.1. Participants

The sample resulted from six hours of road-user observations in two separate T-junctions, one with priority assigned to the straight arm and one with priority assigned to the intersecting arm. Observations were divided in two three-hour sessions distributed in two separate days. One session was in the interval 9–12 a.m., and the other in the interval 3–6 p.m., to avoid peak hours. Both T-junctions were included in an urban zone with 30-km/h speed limit. Observations were focused on vehicles (cars and bicycles) that had to yield at the intersection. Criteria for the inclusion were that the intersection was free from other users and that the vehicle was not preceded or followed by other vehicles. These criteria were aimed to better isolate the effect of the priority scheme and intersection layout on the road user's behaviour, excluding contingent behaviours due to the need to stop or slowing down for the presence of other vehicles in the intersection. To have a more homogeneous sample taxi and buses were also excluded since they included professional drivers with an extensive past experience of the intersections under scrutiny. The resulting sample included 518 road users, 287 in the priority-to-straight-arm condition (76 women and 211 men, 223 cars and 64 bicycles), and 231 in the priority-to-intersecting-arm condition (116 women and 115 men, 44 cars and 187 bicycles). The study was approved by the Ethics Committee of the University of Bologna.

2.1.2. Procedure and data analysis

The study compared two T-junctions in the city center of Bologna (Fig. 1). In the priority-to-straight-arm condition, the road user approached the T-junction from the connecting road and had to yield to the vehicles approaching from the straight through road (Fig. 1A). In the priority-to-intersecting-arm condition road users approaching the intersection from the straight through road had to yield to the vehicles coming from the connecting arm (Fig. 1B).

Two ELP high-speed (120 fps) high-definition digital camera connected via USB to a laptop were used to record the road user's behavior and speed when approaching the yielding point. One camera (camera A in Fig. 1) was placed 10 m before the transverse yielding marking, framing the 20-m segment that preceded the transverse marking in both conditions. The second camera (camera B in Fig. 1B) was positioned 10 m after the transverse yielding marking in the case of priority-to-intersecting-arm T-junction, and at the opposite side of the road for the priority-to-straight-arm T-junction (Fig. 1A). Both camera bodies and laptops were concealed in a gray metal box positioned in the exterior margin of the sidewalk.

The recordings of the camera A in each T-junction were used to compute road users' average speed in the 20-m interval that preceded the transverse yielding marking. Average speed was computed dividing the 20-m distance by the number of frames employed by the road user to cover this distance. Specifically, to compute the start and end time of the 20-m segment we targeted the center of the left-anterior wheel in case of cars, and the center of the anterior wheel in case of bicycles. Number of frames were converted in seconds considered that each frame had a duration of 8.33 ms. Speed in m/s was then converted to km/h.

In addition, the combined analysis of the recordings of cameras A and B (Fig. 1) were used to extrapolate three categorical variables. In the first variable we evaluated if the road user slowed down approaching the intersection. Slowing down was computed comparing speed in the interval $-20/-10$ m to speed in the interval $-10/0$ m, with reference to

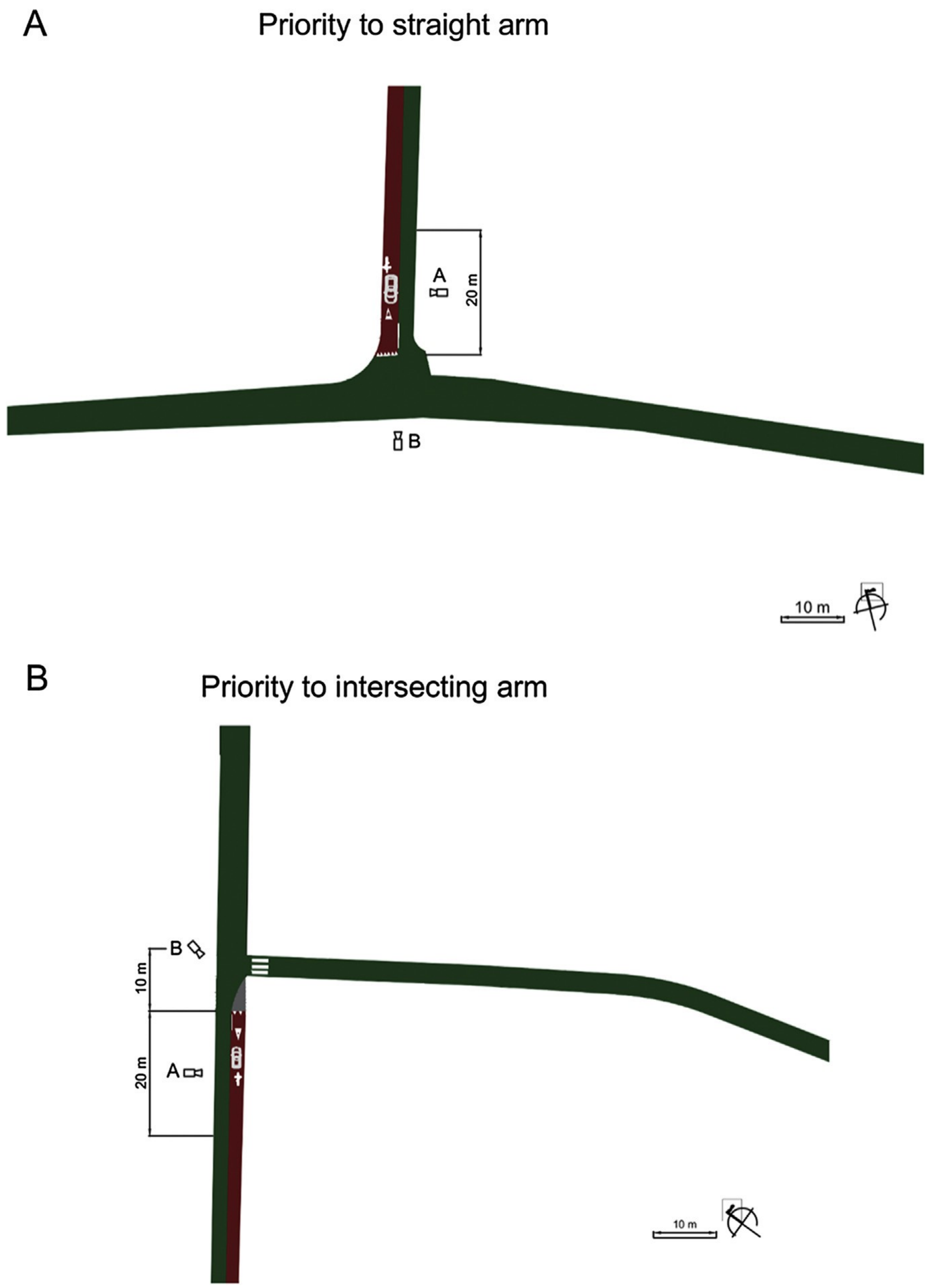

Fig. 1. Topological layout of the two T-junctions examined in Study 1. In the priority-to-straight-arm condition (A) drivers from the connecting arm (red), had to yield to road users in the straight through road (green). In the priority-to-intersecting-arm condition (B) drivers approaching from the straight through road (red) had to give way to vehicles approaching from the connecting arm (green). Plans are in scale. Points A and B show the positions of the recording cameras. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the transverse yielding marking. The road user was considered to have slowed down if the speed in the interval −10/0 m was inferior of at least 5% to the speed in the interval −10/−20 m. In the second variable, we coded if the road user stopped at the intersection. Although we analyzed the behavior of road users in a context of free intersection, without other drivers, and the traffic was Yield controlled and not Stop controlled, we considered in a categorical variable if the user stopped at the intersection or not. This behavior was considered has an additional and strengthened cue of lack-of-priority perception. The third categorical variable coded if the road user turned the head to the intersection arm that had priority, checking for incoming vehicles. Head turning was assessed in the 20-m segment that preceded the transverse yielding marking, using a threshold of 20°.

The differences between the two priority scheme conditions were tested with Chi-square for the three categorical variables, and with a one-way ANOVA for the speed data.

2.2. Results

2.2.1. Speed

Approaching speed to the intersection was significantly lower for both cars and bicycles in the priority-to-straight-road condition than in the priority-to-intersecting-arm condition: $F(1, 418) = 61.67$, $p < .001$, $\eta_p^2 = 0.13$. Mean speed for cars was 15.08 km/h ($SD = 4.02$) in the first condition and 20.82 km/h in ($SD = 9.14$) in the second condition. Mean speed for bicycles was 11.5 km/h ($SD = 2.67$) in the first condition and 17.47 km/h ($SD = 7.25$) in the second condition.

2.2.2. Slowing down

Considering car drivers, 96.4% slowed down when approaching to the priority-to-straight-arm T-junction, whereas 31.8% slowed down in the priority-to-intersecting-arm condition: χ^2 (1) = 125.61, $p < .001$, $\phi = 0.69$. Considering bicyclists, 31.2% slowed down in the first condition *versus* 0% in the second condition: χ^2 (1) = 63.49, $p < .001$, $\phi = 0.50$.

2.2.3. Stopping

Considering car drivers, 47.1% stopped at the intersection in the priority-to-straight-arm condition, whereas only 2.3% stopped in the priority-to-intersecting-arm condition: χ^2 (1) = 30.82, $p < .001$ $\phi = 0.34$. None of the bicyclists stopped in both the T-junctions.

2.2.4. Head turning

One-hundred percent of car drivers turned their head to check for incoming vehicles in the priority-to-straight-arm condition, whereas 59.1% of car drivers in the priority-to-intersecting-arm condition turned their head χ^2 (1) = 97.82, $p < .001$, $\phi = 0.60$. Considering bicyclists, 93.8% checked for incoming vehicles turning their head in the first condition, whereas 29.4% turned their head in the second condition: χ^2 (1) = 79.50, $p < .001$, $\phi = 0.56$.

2.3. Discussion study 1

In a yield-controlled T-junction the assignment of priority to road users approaching from the straight through arm or approaching from the connecting arm had a remarkable effect on the drivers' behavior. When the priority scheme followed the perceptual affordance, giving priority to the vehicles in the straight through road, then the driver's approach to the intersection was much safer, with a significant reduction in speed, an increase in visual inspection to the arm with priority, and an increase of stopping behavior. These results can be explained by the change of the road layout when the priority scheme was changed. In the priority-to-straight-arm condition the requirement to yield matched with a road layout that imposed a physical constraint to the driver that had to turn either to left or right, whereas in the priority-to-intersecting-arm condition the road layout had no physical constraint, continuing along a straight trajectory after the intersection.

3. Study 2

Monitoring approaching vehicles from the road with priority is an important and critical behavior for a safe crossing of road intersections. In the first study, the visual inspection was evaluated as a categorical variable, considering the driver's head turning. The aim of the second study was to better investigate the intersection visual monitoring using an eye-movement recording methodology, considering the same two T-junctions included in Study 1. Since both intersections were placed in the city center, in a restricted traffic zone with access reserved to only residents and public transportation, in which the bicycle flow was rather high, we decided to focus on bicyclists only. We analyzed the total fixation duration to the critical area in which vehicles with priority could approach the intersection and the angular eye movements on the horizontal plane suggesting that the priority-to-straight-arm scheme would significantly increase the inspection time and the horizontal eye movements in comparison to a priority-to-intersecting-arm scheme. Contrary to Study 1 the sample was not the result of observations of the spontaneous vehicular traffic, but it was composed by bicyclists that volunteered to cover with their bike a specific route that included the two T-junctions under exam. As in Study 1, we included only the cases of crossing in a free condition, in which the bicyclist approached the intersection alone, without other road users nearby. The reason was to isolate the effect of the priority scheme between the two conditions, excluding intervening variable due to the presence of other road users in the intersection.

3.1. Method

3.1.1. Participants

Nine cyclists, 6 men (M_{age}: 24.5, *SD*: 2.07) and 3 women (M_{age}: 22.33 *SD*: 2.89) volunteered in this study. None of them wore glasses or contact lenses that prevented eye movement recording. None of them had previous experience with both intersections. Participants were psychology and engineering undergraduate students. They were blind about the aim of the study and were told that the study aimed to test a mobile eye-movement-recording device in bicyclists. All the participants had a class B driving license (for cars). The study was approved by the Ethics Committee of the University of Bologna.

3.1.2. Procedure

Each participant performed both conditions in a within-subject design, and was asked, for each condition, to ride an experimental bicycle equipped with the mobile eye-movement recording system, along a predefined route that included the two T-junctions under examination. For five participants the experimental route included the priority-to-straight-arm T-junction first, whereas for the remaining four participants the route met the priority-to-intersecting-arm T-junction first. The route, including both T-junctions and the return to the starting point, was 1278 m. The distance between the two T-junctions was 428 m. The participant was instructed to ride the experimental route once. Experimental sessions were run in the intervals 9–12 a.m., and 3–6 p.m., to avoid peak hours.

An ASL Mobile Eye-XG system was used to record eye movements. The system is equipped with two digital cameras attached to lightweight eyeglasses. One camera recorded the scene image while the other recorded the participant's eye and pupil. Eye tracking was performed on the right eye with a sampling rate of 30 Hz and an accuracy of 0.5–1°. Each data collection was preceded by a calibration procedure to get a good accuracy of the eye-movement recording system. The calibration consisted in looking at 15 points (vertexes and centres of small objects) encompassing the whole visual scene, and recording their position in the acquisition software.

3.1.3. Data analysis

One region of interest (ROI) was defined for each condition (Fig. 2). In the priority-to-intersecting-arm condition, the ROI included the connecting arm, and in the priority-to-straight-arm condition the ROI included the left arm (bicyclist's point of view) of the straight through road (Fig. 2). These ROIs outlined the area of potential approaching vehicles with priority. ASL Results software was used to compute the total fixation time to the ROI in the two conditions. Given the 30 Hz sample rate of the eye-movement recording equipment the temporal resolution was 33 ms. Total fixation time to the ROIs was computed by the ASL Results software with a frame-by-frame tracking of the ROI areas. Fixations were defined by the permanence of the pupil center in an area of 1° visual angle for at least two frames ($\geq$66 ms).

This threshold takes into account the results of previous studies (Costa, Simone, Vignali, Lantieri, & Palena, 2018; Sodhi, Reimer, & Llamazares, 2002; Velichkovsky, Dornhoefer, Pannasch, & Unema, 2000) that showed a significant occurrence of short fixations (i.e., <100 ms) in the dynamical contexts of driving. Total fixation time to the priority area was assessed in the -20/+5 m segment with reference to the yield line. The difference between the two priority schemes was tested with a repeated-measure ANOVA.

Angular eye movements on the horizontal plane were computed considering that the exterior scene camera had a horizontal resolution of 640 pixels covering an angle of 60° (i.e., $10.\bar{6}$ px for 1°). The angular eye movements on the horizontal plane (M_x) was computed for each sample frame with the Formula 1, where x was a x-coordinate of the eye position on the driver's visual scene, i was the sample frame number considering a 30 Hz sampling rate, and n was the total number of frames acquired in the interval -50 – 0 m approaching the yield line at the intersection. Each data was multiplied by 30 to refer the value to a 1-s base instead of the time base of the single frame ($33.\bar{3}$ ms). A repeated-measure ANOVA was applied for testing the difference between the two conditions.

$$M_x = \frac{\sum_{i=1}^{n}(x_i - x_{i-1})/10.\bar{6} * 30}{n} \tag{1}$$

Fig. 2. Region of interest (in red) for the evaluation of intersection visual inspection in the two priority scheme conditions. In blue it is showed the incoming direction of the experimental bicyclist. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

A third parameter analyzed if the bicyclist glanced to the "give way" vertical sign at the intersection. The ROI was defined by the inverted triangle area of the road sign, and the sign was considered as "fixated" if the pupil center target was stationary inside the ROI area for at least two samples ($\geq$66 ms). The difference between the two conditions was tested with a Chi-square analysis.

3.2. Results

3.2.1. Fixation time to the priority area

The total fixation time to the ROI area for bicyclists approaching to the intersection was higher in the priority-to-straight-arm condition (M = 1336 ms, SD = 941) than in the priority-to-intersecting-arm condition (M = 355 ms, SD = 637). The difference was significant: $F(1, 8) = 9.68$, $p = .01$, $\eta_p^2 = 0.55$. Fixation time to the ROI area before the transverse yielding marking was 977 ms (SD = 537) in the first condition and 141 ms (SD = 249) in the second condition: $F(1, 8) = 16.93$, $p = .003$, $\eta_p^2 = 0.68$. After the transverse yield marking the difference between the two conditions was not significant ($p = .34$), and was in mean 286 ms (SD = 463).

3.2.2. Angular eye movements on the horizontal plane

Mean angular eye movements on the horizontal plane, computed on the interval −20/+5 m with reference to the yield line, was 43.11 deg/s (SD = 7.38) in the priority-to-straight-arm condition and 29.32 deg/s (SD = 13.17) in the priority-to-intersecting-arm condition. The difference, tested with an ANOVA, was significant: $F(1, 12) = 6.26$, $p = .02$, $\eta_p^2 = 0.34$. Fig. 3 shows the grand-averaging of the angular horizontal eye movements in the five seconds of approach to the two T-junctions with different priority scheme.

3.2.3. Fixation to the 'Give way' road sign

One participant (11%) glanced at the 'Give way' sign in the priority-to-intersecting-arm condition for 231 ms, whereas in the priority-to-straight-arm condition the sign was glanced by 7 participants (77%), for a median value of 133 ms (SD = 130). The difference in frequency between the two conditions was significant: $\chi^2 = 4.5$, $p < .03$.

3.3. Discussion study 2

The results of this study confirmed those of Study 1, using a more accurate and objective methodology for the assessment of bicyclists' visual monitoring behavior when approaching T-junctions differing for their priority scheme. A traffic control and intersection design that support the affordance for speed reduction and promote a perception of no priority resulted in an increase of the monitoring time for potential incoming vehicles by a factor of 3.7. Furthermore, horizontal eye movements when approaching the intersection were 47.03% higher in the priority-to-straight-arm condition. Although the number of participants was rather low, the effect size of the difference was high, particularly considering the segment before the yielding marking. A straight road is strongly connected to an affordance of having priority, and as a consequence a driver tends to decrease visual attention to lateral connecting roads. Inserting a "give way" traffic control in this context strongly weaken the intrinsic safety of the intersection.

The driver's priority expectations promoted by the T-junction priority layout affected also the proportion of the "give way" road sign detection. The glances to the road sign were significantly higher when the bicyclist approached the intersection from the connecting arm in comparison to the condition in which the bicyclist approached from the straight through road.

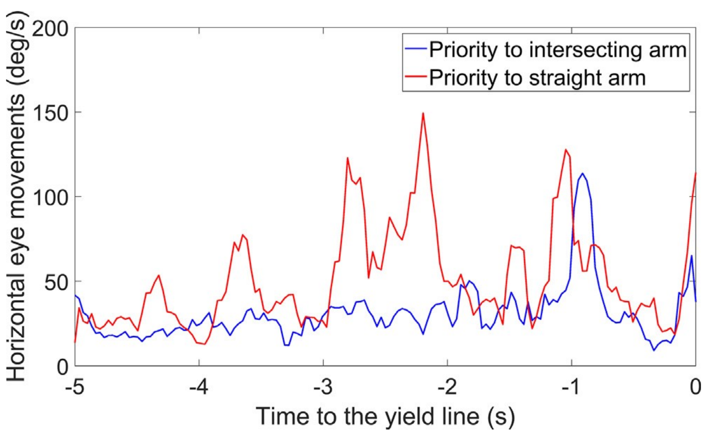

Fig. 3. Grand-averaging of the angular eye movements on the horizontal plane for the 5 s before the yield line in the two T-junctions differing for the priority scheme.

The intersections examined in Studies 1 and 2 were placed in a historical urban context, in a zone in which vehicular traffic was restricted to residents, and with a medium-high proportion of bicycles. Therefore, for a better generalization of the results, we decided to extend the study to another context in which both T-junction conditions were placed in a peripheral zone with higher vehicular flow and operational speed.

4. Study 3

Study 3 mirrored the methodology of Study 1, changing the context of the two T-junctions that in this study were placed in a suburban zone. We investigated mean speed, driver's slowing down behavior, stopping, and head turning to the intersection leg with priority, comparing two T-junctions with opposite priority schemes.

4.1. Method

4.1.1. Participants

Six hours of video-recording of the vehicular flow in one priority-to-straight-arm and one priority-to-intersecting-arm T-junction resulted in the observation of 302 car drivers in the first condition, and 219 in the second condition, for a total of 521 participants. Observations were divided in two three-hour sessions in two consecutive days. The sample included 72 women and 147 men. As in Study 1 we included only the vehicles that approached in a context of free intersection, without other road users in the T-junction area. Differently from Study 1 we considered only car drivers since the bicycle flow was very low. The study was approved by the Ethics Committee of the University of Bologna.

4.1.2. Procedure and data analysis

The layout for the priority-to-straight-arm and the priority-to-intersecting-arm priority schemes used in this study is shown in Fig. 4. Both intersections connected two-lane roads, with one lane for each direction. In the priority-to-straight-arm condition vehicles approaching from the connecting arm had to yield to the vehicles approaching through the straight road (Fig. 4A), establishing a match between the appropriate behavior for giving the right-of-way (i.e., slowing down, monitoring incoming vehicles) and the spontaneous behavior induced by the road layout (i.e. the road ends at the intersection and the driver is forced to turn). In the priority-to-intersecting-arm condition, vehicles approaching from the straight road had to yield to road users approaching from the right (connecting arm), and to vehicles approaching from the opposite direction that had to turn left in the connecting arm (Fig. 4B). In this case there was a mismatch between the road geometry and the prompting of an appropriate yielding behavior. In the priority-to-intersecting-arm condition, the connecting arm was not perpendicular but met the straight through road with a large curve (37-m radius). The procedure for the video-recording and data analysis mirrored exactly those described for Study 1.

4.2. Results

4.2.1. Speed

Mean speed in the last 20-m approaching the intersection was 30.48 km/h (SD = 11.18) in the priority-to-straight-road condition and 35.18 km/h (SD = 11.36) in the priority-to-intersecting-arm condition. The difference was significant: $F(1, 519) = 22.11$, $p < .001$, $\eta_p^2 = 0.04$.

4.2.2. Slowing down

The drivers that slowed down approaching the intersection were 83.1% in the priority-to-straight-arm condition and 46.8% in the priority-to-intersecting-arm condition: $\chi^2(1) = 90.16$, $p < .001$, $\phi = 0.38$.

4.2.3. Stopping

The drivers that stopped approaching the intersection were 7.6% in the priority-to-straight-arm condition and 4.3% in the priority-to-intersecting-arm condition: $\chi^2(1) = 3.79$, $p = .05$, $\phi = 0.07$.

4.2.4. Head-turning

Head turning toward the intersection arm with priority was 98.3% in the priority-to-straight-arm condition and 90.8% in the priority-to-intersecting-arm condition. The difference was significant: $\chi^2(1) = 15.53$, $p < .001$, $\phi = 0.17$.

4.3. Discussion study 3

The results confirmed those of Study 1, showing that the driver's perception of having priority is much enhanced when the traffic control matched the affordances for priority. Drivers approaching the intersection from the connecting arm and having to yield to the straight through road exhibited a significant increase in safer behaviors as approaching at a lower speed, stopping at the yielding line, and increasing visual monitoring of the intersection area.

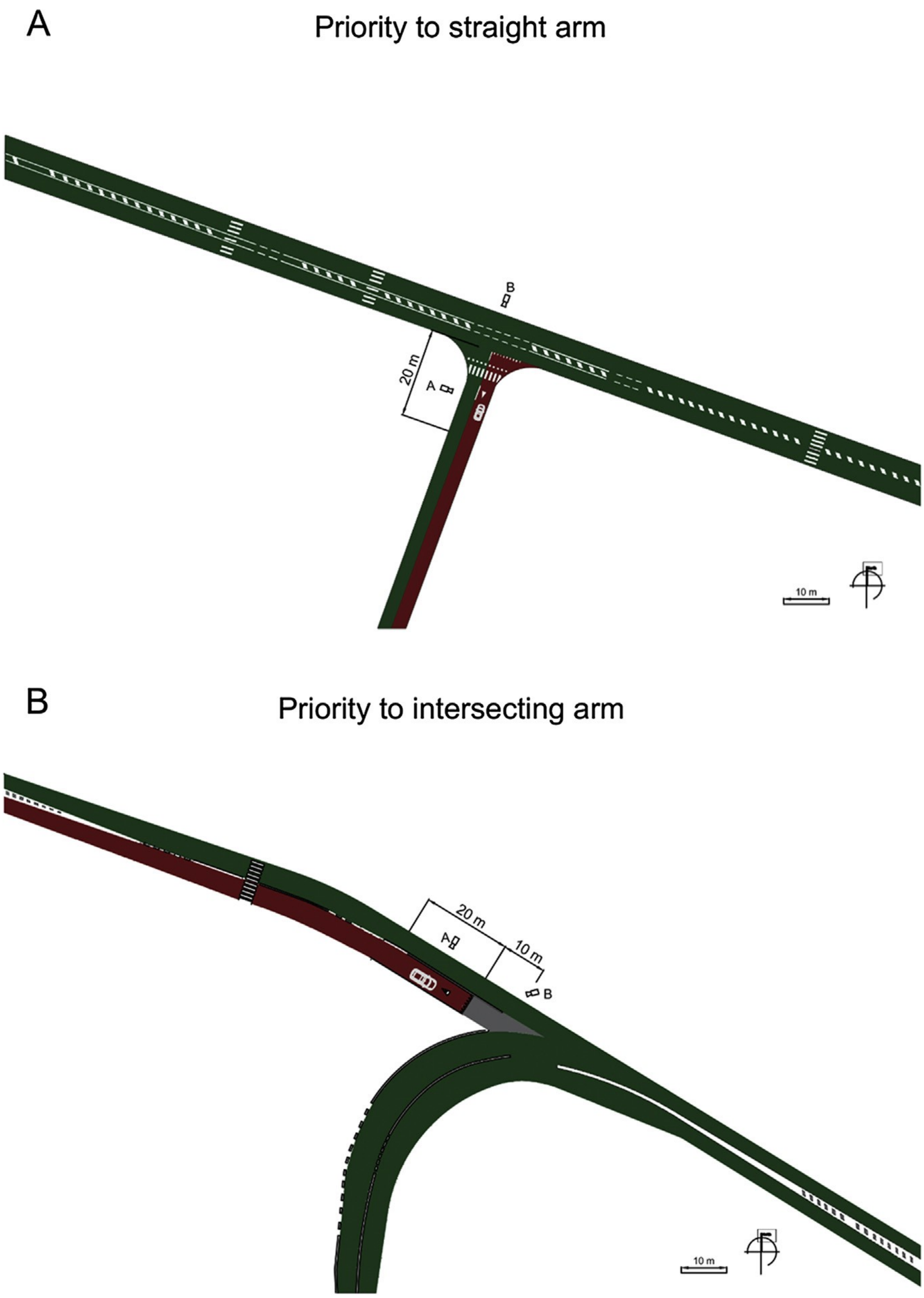

Fig. 4. Layout for the priority-to-straight-arm (A) and priority-to-intersecting-arm (B) T-junctions considered in Study 3. Vehicles approaching from the red lane had to yield to vehicles approaching from the green lanes. Plans are in scale. Points A and B show the location of the recording cameras. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Study 4

In study 4 the methodology of study 2 was applied to the intersections investigated in study 3. A mobile eye-movement recording system was used to better investigate the visual behavior of drivers approaching to the T-junction examined in study 3 differing for their priority scheme. As in study 2, we applied a repeated-measure design in which each driver met along a defined route both type of intersections. To avoid contextual factors, we assessed the total duration of fixations to critical areas of the intersection when the driver approached the junction area and the mean angular eye movements on the horizontal plane in a condition of free intersection. In addition, the kinematic parameters of the drivers were recorded (speed, acceleration), with the aim to compare speed when approaching the two types of T-junctions. We expected that participants would spend more time looking at the critical areas of the intersection in the priority-to-straight-arm scheme than in the priority-to-intersecting-arm scheme, and that approaching speed would be higher in the priority-to-intersecting-arm condition.

5.1. Method

5.1.1. Participants

Ten car drivers, eight men (M_{age} 26.5 years, SD = 3.12) and two women (M_{age} 24 ± 1.12) participated in this study. None of them wore glasses or contact lenses, since this would have prevented the recording of eye-movements. Participants were psychology or engineering undergraduate students. They were blind about the aim of the study and were told that its aim was to test a mobile eye-recording device during driving. None reported a previous knowledge of the experimental route. All the participants had a class B driving license (for cars). The study was approved by the Ethics Committee of the University of Bologna.

5.1.2. Procedure

All participants drove a BMW 1 Series car. Each driver was given a two kilometres trial run to get used to the car before starting the test route. The test route was a two kilometres long itinerary which included both priority-scheme conditions. The order of the two intersections was counter-balanced between participants. The priority-to-intersecting-arm T-junction was preceded by a 394 m straight segment while the priority-to-straight-arm T-junction was preceded by a 115 m straight segment (Fig. 4). Speed and acceleration data were recorded through the Video V-Box Pro data logger. Eye movements were recorded from the right eye using the ASL Mobile Eye-XG system, as in Study 2. The eye-tracking equipment along with a computer and the Video V-Box Pro equipment were kept on the back seat where one experimenter monitored the devices during the experiment. He was instructed not to talk to the driver except for giving instructions, directions, or assistance. As in Study 2 the eye-tracking device was calibrated for each participant for spatial accuracy.

5.1.3. Data analysis

Mean speed approaching the intersection was computed from the Video Vbox Pro data considering the −20/+5-m segment with reference to the transverse yielding marking. Mean speed in the two conditions was compared with a repeated-measure ANOVA design.

Two regions of interest (ROI) were defined in each condition as showed in Fig. 5. The ROIs included segments of both the left and right arms, excluding the area in which the driver looked straight ahead. The aim was to explore peripheral visual inspection of the intersection area by the driver approaching the intersection. Peripheral area in fact included potential incoming vehicles that could conflict with the driver's trajectory. Total fixation time to the ROIs was computed by the ASL Results software tracking frame-by-frame the ROI areas. Fixations were defined by the permanence of the pupil center in an area of 1° visual angle for at least two frames (≥66 ms). Fixation analysis was performed on the −50/+5 m road segment from the yielding transverse marking. The segment was more extended in comparison to Study 2 because the intersections were placed in a suburban area with lower building density and increased visibility, with the driver that could fixate to the priority road in advance to the −20 m interval considered in Study 2. The data were statistically tested with a repeated-measure ANOVA.

Angular eye movements on the horizontal plane were assessed with the same methodology described for Study 2, with the only exception that in this study we considered the segment −50/+5 m. A repeated-measure ANOVA was applied for testing the difference between the two conditions.

Fig. 5. Regions of interest (in red) for the evaluation of intersection visual inspection in the priority-to-straight-arm and priority-to-intersecting-arm conditions. In blue it is showed the incoming direction of the experimental car. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.2. Results

5.2.1. Total fixation time to ROIs

Visual inspection for incoming vehicles in the segment −50/+5 m was significantly different between the two T-junctions: $F(1, 9) = 10.14$, $p = .01$, $\eta_p^2 = 0.53$. Total fixation time in the priority-to-straight-road T-junction was 2036 ms ($SD = 252$), whereas in the priority-to-intersecting-arm T-junction it was 1253 ms ($SD = 247$).

5.2.2. Angular eye movements on the horizontal plane

Mean horizontal eye movements (deg/s) in the interval −50/0 m approaching the intersections were 34.31 deg/s ($SD = 12.44$) in the priority-to-intersecting-arm condition and 54.12 deg/s ($SD = 21.46$) in the priority-to-straight-arm condition. The difference was significant: $F(1, 9) = 13.53$, $p = .005$, η_p^2=0.60. Fig. 6 shows the grand averaging of angular eye movements on the horizontal plane for all the participants in the 5 s before the yield line, distinguishing between the two priority schemes.

5.2.3. Speed

Driver's approaching speed in the segment −20/+5 between the two intersections was significantly different: $F(1, 9) = 70.53$, $p < .001$, η_p^2=0.88. Mean speed in the priority-to-straight-arm condition was significantly lower ($M = 20.25$ km/h, $SD = 4.33$) compared to the priority-to-intersecting-arm condition ($M = 29.83$ km/h, $SD = 6.61$).

5.3. Discussion study 4

Although the sample was limited to 10 participants, we got significant differences in the approaching speed and the visual monitoring of the priority road in the intersection area contrasting two opposite priority schemes. Specifically, the priority-to-intersecting-arm scheme resulted in a 47.31% increase in speed approaching the intersection, and in a 38.45% reduction of total fixation time to the lateral areas of the intersection, where potential conflicting road users could approach. Horizontal eye movements approaching the intersection were 57.73% higher in the priority-to-straigh-arm T-junction. These results confirmed those found in Study 3, adding further evidence that a match between affordances for priority and the geometrical layout of the intersection leads to a safer crossing.

6. General discussion

We examined different priority arrangements applied to T-junctions, and their effects on approaching speed and visual inspection of the intersection area, two measures that strictly concur to the safety of an intersection. We considered signalized T-junctions, distinguishing between a priority scheme in which road users in the intersecting arm had to yield to road users in the straight through arm (priority-to-straight-arm condition), and a priority scheme in which road users in the straight arm had to yield to those approaching from the lateral arm (priority-to-intersecting-arm condition).

We also considered two pairs of T-junctions differing for the context in which they were inserted. The first pair was composed by T-junctions located in a city center with a mixed flow of bicycles and cars, whereas the other pair was composed by T-junctions located in an urban peripheral context, with an almost exclusive flow of cars. To ensure a direct comparability between the observations we assessed the road-user's behaviour and eye movements when approaching the intersection in a condition of isolation, without other road users approaching or leaving the intersection.

The results in both contexts clearly showed that a priority-to-straight-arm priority scheme led to a safer and cautious approach to the intersection. In the T-junctions of studies 1 and 2 speed was reduced by 27.56% for car drivers, and by

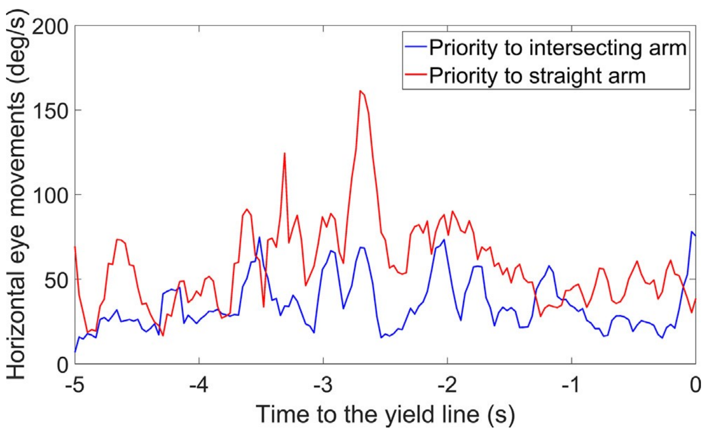

Fig. 6. Grand-averaging of the angular eye movements on the horizontal plane for all participants in the five seconds before the yield line for the two priority scheme conditions.

34.17% in case of bicyclists in the priority-to-straight-arm condition, and the difference was considerable considering behaviors as slowing down, stopping, head turning, and checking for approaching vehicles. The difference between the two conditions was very remarkable for bicyclists since none slowed down when approaching to the intersection with priority assigned to the intersecting arm, whereas 31.2% slowed down in the condition with priority assigned to the straight arm. The city center urban context in which the two intersections examined in studies 1 and 2 were placed, with a low vehicular flow, resulted in a self-confident driving style by the bicyclists, which was however significantly affected by the different priority scheme enforced in the two T-junctions.

The second study that examined eye movements of bicyclists approaching the two intersections confirmed that in a priority-to-straight-arm design, road users played more attention and explored more carefully and for a longer time the intersection area where potential conflicting road users could approach. The total fixation time in the priority-to-straight-arm scheme was higher with a magnitude of 3.7 in comparison to the priority-to-intersecting-arm scheme. Eye-movements on the horizontal plane, a measure of inspection of the visual field, were 47.03% higher in the priority-to-straight-arm condition, showing that bicyclists explored more the surround environment checking for potential other road users approaching the intersection.

In the contexts considered in studies 3 and 4 the results confirmed our hypothesis, with lower effect sizes than in studies 1 and 2. Mean speed was 4.7 km/h lower, and slowing behaviour was 36.3% more frequent approaching the intersection with priority-to-straight-arm scheme than in the priority-to-intersecting-arm condition. Also head turning behaviour for monitoring incoming road users from the right arm of the intersection was more frequent (+7.5%) in the condition with priority-to-straight-arm. Study 4, with the assessment of driver's eye movements approaching the intersection, further confirmed these results showing that total fixation time to the critical intersection area was 0.87 s higher in the priority-to-straight-arm condition. Furthermore, horizontal eye movements were 57.73% higher when priority assignment matched the affordance and perception for priority. Since mean speed, slowing down and visual inspection when approaching an intersection are important factor for crash prevention and safety in intersection (Choi, 2010; Walton, Buchanan, & Murray, 2013) it is possible to conclude that a priority-to-straight-arm scheme would significantly contribute to increase the safety level and reduce the crash rate in an intersection.

The affordance that was investigated in our studies was the perception of having priority when travelling along a straight road, and, conversely, the perception of having to yield when a road is discontinued in an intersection. This is only one of the potential affordances that could influence the perception of priority, and future investigations should focus on other affordances. For example, a road with a larger cross-section could be perceived as having priority over a road with narrow cross-section, or, in case the intersection arms are not planar, a road that is more elevated could be perceived as having priority over an intersecting road that is less elevated. The approaching curvature and geometry of a road to the intersection could also play a significant role for the affordance of priority. For example, if a road is connected to the intersection by a straight line the perception of having priority could be significantly higher than a condition in which a road is connected to an intersection with a curve. In roundabouts, for example, road users enter with a curvilinear trajectory and this contributes to a speed reduction and hence to an increased safety in comparison to standard signalized intersections (Elvik, 2003; Gross, Lyon, Persaud, & Srinivasan, 2013; Jensen, 2013). Also Stephens et al. (2017), proposed two intersection designs that succeeded in speed reduction eliminating the possibility for a road user to cross the intersection along a straight trajectory, introducing islands that induced curvilinear trajectories.

Apart from the intersection geometry, other elements could increase the legibility of the priority scheme in an intersection. For example, the use of different pavement colors in the arms with priority and in the arms without priority could facilitate road users in understanding priority assignments between the intersection arms. This methodology, for example, has been applied with success in crosswalks in which adding a red color increased the safety for pedestrians (Iasmin, Kojima, & Kubota, 2016). Another possibility would be to modify the road pavement texture of the intersection part in which road users have to yield, using bricks or a more irregular surface texture that would increase friction and noise, following the positive results obtained by transverse rumble strips (Liu, Huang, Wang, & Xu, 2011; Yang, Zhou, Zhu, & Qu, 2016), and longitudinal rumble strips in centreline two-lane roads (Vadeby & Anund, 2017), and in road shoulders (Wu, Donnell, & Aguero-Valverde, 2014).

Usually the relative importance of each road and the amount of traffic flows are generally the main factors governing intersection priority assignment. In this paper we highlighted that perceptual affordances emerging from the intersection geometry could play a significant role in increasing an intersection safety. The mere analysis of traffic flow could result in intersections that violate drivers' expectations, increasing the frequency of unsafe behaviors. Hence, intersection design could strongly benefit from including criteria connected to perceptual affordances and drivers' expectations for priority.

Acknowledgements

This research was supported by the Alma Idea COSTA_2017 Grant from the University of Bologna.

References

Björklund, G. M., & Åberg, L. (2005). Driver behaviour in intersections: Formal and informal traffic rules. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(3), 239–253. https://doi.org/10.1016/j.trf.2005.04.006.

Chai, C., Wong, Y. D., & Wang, X. (2017). Safety evaluation of driver cognitive failures and driving errors on right-turn filtering movement at signalized road intersections based on Fuzzy Cellular Automata (FCA) model. *Accident Analysis and Prevention, 104*, 156–164. https://doi.org/10.1016/j.aap.2017.04.026.

Charlton, S. G. (2007). Delineation effects in overtaking lane design. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(2), 153–163. https://doi.org/10.1016/J.TRF.2006.09.003.

Choi, E. H. (2010). Crash factors in intersection-related crashes: An on-scene perspective (DOT No. HS-811 366). *National Highway Traffic Safety Administration.*.

Costa, M., Bonetti, L., Vignali, V., Lantieri, C., & Simone, A. (2018). The role of peripheral vision in vertical road sign identification and discrimination. *Ergonomics*. https://doi.org/10.1080/00140139.2018.1508756.

Costa, M., Simone, A., Vignali, V., Lantieri, C., Bucchi, A., & Dondi, G. (2014). Looking behavior for vertical road signs. *Transportation Research Part F: Traffic Psychology and Behaviour, 23*, 147–155. https://doi.org/10.1016/j.trf.2014.01.003.

Costa, M., Simone, A., Vignali, V., Lantieri, C., & Palena, N. (2018). Fixation distance and fixation duration to vertical road signs. *Applied Ergonomics, 69*, 48–57. https://doi.org/10.1016/J.APERGO.2017.12.017.

Elliot, M. A., Mccoll, V.A., & Kennedy, J. V. (2003). Road design measures to reduce drivers' speed via 'psychological' processes: A literature review (No. TRL564).

Elvik, R. (2003). Effects on road safety of converting intersections to roundabouts: Review of evidence from non-U.S. studies. *Transportation Research Record: Journal of the Transportation Research Board, 1847*, 1–10. https://doi.org/10.3141/1847-01.

Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety. *Journal of Planning Literature, 23*(4), 347–367. https://doi.org/10.1177/0885412209335553.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gross, F., Lyon, C., Persaud, B., & Srinivasan, R. (2013). Safety effectiveness of converting signalized intersections to roundabouts. *Accident Analysis and Prevention, 50*, 234–241. https://doi.org/10.1016/j.aap.2012.04.012.

Helmers, G., & Åberg, L. (1978). *Förarbeteende i gatukorsningar i relation till företrädesregler och vägutformning*. Linköping, Sweden: En explorativ studie.

Hills, B. L. (1980). Vision, visibility, and perception in driving. *Perception*. https://doi.org/10.1068/p090183.

Iasmin, H., Kojima, A., & Kubota, H. (2016). Safety effectiveness of pavement design treatment at intersections: Left turning vehicles and pedestrians on crosswalks. *IATSS Research, 40*(1), 47–55. https://doi.org/10.1016/j.iatssr.2016.04.001.

Janssen, W., van der Horst, R., Bakker, P., & ten Broeke, W. (1988). Auto-auto and auto-bicycle interactions in priority situations. In T. Rothengatter & R. de Bruin (Eds.), *Road user behaviour: Theory and research* (pp. 639–644). Assen/Maastricht: van Gorcum.

Jensen, S. (2013). Safety effects of converting intersections to roundabouts. *Transportation Research Record: Journal of the Transportation Research Board, 2389*, 22–29. https://doi.org/10.3141/2389-03.

Johannessen, S. (1984). Kjöreatferd i uregulerte t-kryss: Höyreregel eller vikepliktregulering? [Driving behaviour in unregulated T-junctions. The right-hand rule or duty to give way?]. Trondheim, Norway.

Jones, K. S. (2003). What is an affordance? *Ecological Psychology, 15*, 107–114. https://doi.org/10.1207/S15326969ECO1502_1.

Lantieri, C., Lamperti, R., Simone, A., Costa, M., Vignali, V., Sangiorgi, C., & Dondi, G. (2015). Gateway design assessment in the transition from high to low speed areas. *Transportation Research Part F: Traffic Psychology and Behaviour, 34*, 41–53. https://doi.org/10.1016/j.trf.2015.07.017.

Liu, P., Huang, J., Wang, W., & Xu, C. (2011). Effects of transverse rumble strips on safety of pedestrian crosswalks on rural roads in China. *Accident Analysis and Prevention, 43*(6), 1947–1954. https://doi.org/10.1016/j.aap.2011.05.012.

Norman, D. (2004). *Emotional design - Why we love (or hate) everyday things*. New York: Basic Books.

Norman, D. (2013). *The design of everyday things* (revised ed). New York: Basic Books. https://doi.org/620.8'2—dc20.

Rothengatter, J. A. (1999). Road users and road design. In G. B. Grayson (Ed.), *Behavioral research in road safety* (pp. 86–93). Crowthorne: Transport Research Laboratory.

Schepers, J. P., Kroeze, P. A., Sweers, W., & Wüst, J. C. (2011). Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections. *Accident Analysis and Prevention, 43*(3), 853–861. https://doi.org/10.1016/j.aap.2010.11.005.

Sodhi, M., Reimer, B., & Llamazares, I. (2002). Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods, Instruments, and Computers, 34*(4), 529–538. https://doi.org/10.3758/BF03195482.

Stephens, A. N., Beanland, V., Candappa, N., Mitsopoulos-Rubens, E., Corben, B., & Lenné, M. G. (2017). A driving simulator evaluation of potential speed reductions using two innovative designs for signalised urban intersections. *Accident Analysis and Prevention, 98*, 25–36. https://doi.org/10.1016/j.aap.2016.09.022.

Theeuwes, J., & Godthelp, H. (1995). Self-explaining roads. *Safety Science, 19*(2–3), 217–225. https://doi.org/10.1016/0925-7535(94)00022-U.

Vadeby, A., & Anund, A. (2017). Effectiveness and acceptability of milled rumble strips on rural two-lane roads in Sweden. *European Transport Research Review, 9*(2). https://doi.org/10.1007/s12544-017-0244-x.

Velichkovsky, B. M., Dornhoefer, S. M., Pannasch, S., & Unema, P. J. A. (2000). Visual fixations and level of attentional processing. Proceedings of the Symposium on Eye Tracking Research & Applications - ETRA '00 (pp. 79–85). 10.1145/355017.355029.

Walker, G. H., Stanton, N. A., & Chowdhury, I. (2013). Self-explaining roads and situation awareness. *Safety Science, 56*, 18–28. https://doi.org/10.1016/j.ssci.2012.06.018.

Walton, D., Buchanan, J., & Murray, S. J. (2013). Exploring factors distinguishing car-versus-car from car-versus-motorcycle in intersection crashes. *Transportation Research Part F: Traffic Psychology and Behaviour, 17*, 145–153. https://doi.org/10.1016/j.trf.2012.11.001.

Wegman, F., Zhanga, F., & Dijkstra, A. (2012). How to make more cycling good for road safety? *Accident Analysis & Prevention, 44*(1), 19–29. https://doi.org/10.1016/J.AAP.2010.11.010.

Weller, G., Schlag, B., Friedel, T., & Rammin, C. (2008). Behaviourally relevant road categorisation: A step towards self-explaining rural roads. *Accident Analysis and Prevention, 40*(4), 1581–1588. https://doi.org/10.1016/j.aap.2008.04.009.

Wu, K. F., Donnell, E. T., & Aguero-Valverde, J. (2014). Relating crash frequency and severity: Evaluating the effectiveness of shoulder rumble strips on reducing fatal and major injury crashes. *Accident Analysis and Prevention, 67*, 86–95. https://doi.org/10.1016/j.aap.2014.02.004.

Yang, L., Zhou, H., Zhu, L., & Qu, H. (2016). Operational effects of transverse rumble strips on approaches to high-speed intersections. *Transportation Research Record, 2602*, 78–87. https://doi.org/10.3141/2602-10.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 18 December 2018
doi: 10.3389/fnhum.2018.00509

EEG-Based Mental Workload Neurometric to Evaluate the Impact of Different Traffic and Road Conditions in Real Driving Settings

Gianluca Di Flumeri[1,2,3], Gianluca Borghini[1,2,3], Pietro Aricò[1,2,3], Nicolina Sciaraffa[1,2,4], Paola Lanzi[5], Simone Pozzi[5], Valeria Vignali[6], Claudio Lantieri[6], Arianna Bichicchi[6], Andrea Simone[6] and Fabio Babiloni[1,3,7]*

[1] BrainSigns srl, Rome, Italy, [2] IRCCS Fondazione Santa Lucia, Neuroelectrical Imaging and BCI Lab, Rome, Italy, [3] Department of Molecular Medicine, Sapienza University of Rome, Rome, Italy, [4] Department of Anatomical, Histological, Forensic and Orthopedic Sciences, Sapienza University of Rome, Rome, Italy, [5] Deep Blue srl, Rome, Italy, [6] Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), School of Engineering and Architecture, University of Bologna, Bologna, Italy, [7] Department of Computer Science, Hangzhou Dianzi University, Hangzhou, China

OPEN ACCESS

Edited by:
Muthuraman Muthuraman, University Medical Center of the Johannes Gutenberg University Mainz, Germany

Reviewed by:
Edmund Wascher, Leibniz-Institut für Arbeitsforschung an der TU Dortmund (IfADo), Germany
Bahamn Nasseroleslami, Trinity College Dublin, Ireland

***Correspondence:**
Gianluca Di Flumeri
gianluca.diflumeri@uniroma1.it; gluca.diflumeri@gmail.com

Received: *16 July 2018*
Accepted: *05 December 2018*
Published: *18 December 2018*

Citation:
Di Flumeri G, Borghini G, Aricò P, Sciaraffa N, Lanzi P, Pozzi S, Vignali V, Lantieri C, Bichicchi A, Simone A and Babiloni F (2018) EEG-Based Mental Workload Neurometric to Evaluate the Impact of Different Traffic and Road Conditions in Real Driving Settings. Front. Hum. Neurosci. 12:509. doi: 10.3389/fnhum.2018.00509

Car driving is considered a very complex activity, consisting of different concomitant tasks and subtasks, thus it is crucial to understand the impact of different factors, such as road complexity, traffic, dashboard devices, and external events on the driver's behavior and performance. For this reason, in particular situations the cognitive demand experienced by the driver could be very high, inducing an excessive experienced mental workload and consequently an increasing of error commission probability. In this regard, it has been demonstrated that human error is the main cause of the 57% of road accidents and a contributing factor in most of them. In this study, 20 young subjects have been involved in a real driving experiment, performed under different traffic conditions (rush hour and not) and along different road types (main and secondary streets). Moreover, during the driving tasks different specific events, in particular a pedestrian crossing the road and a car entering the traffic flow just ahead of the experimental subject, have been acted. A Workload Index based on the Electroencephalographic (EEG), i.e., brain activity, of the drivers has been employed to investigate the impact of the different factors on the driver's workload. Eye-Tracking (ET) technology and subjective measures have also been employed in order to have a comprehensive overview of the driver's perceived workload and to investigate the different insights obtainable from the employed methodologies. The employment of such EEG-based Workload index confirmed the significant impact of both traffic and road types on the drivers' behavior (increasing their workload), with the advantage of being under real settings. Also, it allowed to highlight the increased workload related to external events while driving, in particular with a significant effect during those situations when the traffic was low. Finally, the comparison between methodologies revealed the higher sensitivity of neurophysiological measures with respect to ET and subjective ones.

In conclusion, such an EEG-based Workload index would allow to assess objectively the mental workload experienced by the driver, standing out as a powerful tool for research aimed to investigate drivers' behavior and providing additional and complementary insights with respect to traditional methodologies employed within road safety research.

Keywords: electroencephalography, mental workload, human factor, machine-learning, asSWLDA, neuroergonomics, car driving, road safety

INTRODUCTION

According to the reports of World Health Organization (WHO) (World Health Organization, 2015), every year traffic accidents cause the death of 1.3 million people around the world, and moreover about 50 million people suffer from a disability caused by accidents related to cars. By 2020, it is estimated that traffic accidents will be the fifth leading cause of death in the world, reaching 2.4 million deaths per year (World Health Organization, 2013). Among the principal causes of the car accidents and related mortality there is the human factor (Hansen, 2007; Subramanian, 2012). In particular, it has been demonstrated that human error is the main cause of the 57% of road accidents and a contributing factor in over 90% of them (Treat et al., 1979). Driver's common errors are largely correlated to overload, distractions, tiredness, or the simultaneous realization of other activities during driving (Allnutt, 1987; Horowitz and Dingus, 1992; Summala and Mikkola, 1994; Petridou and Moustaki, 2000). In fact, the human performance decrease, and consequently the errors commission, are directly attributable to aberrant mental states, in particular the mental workload while degrading in overload, which is considered one of the most important human factor constructs in influencing performance (Reason, 2000; Parasuraman et al., 2008; Paxion et al., 2014). The model theorized by De Waard (1996), widely used in automotive psychological research, establishes the relation between task demands and performance depending on the driver workload. This model describes the driving activity with a hierarchy of tasks on three levels, the *strategical*, the *tactical* and the *operational*, each of them divided into different subtasks, describing the driving as a very complex and often high-demanding activity. Therefore, the cognitive resources required in very complex situations can exceed the available resources, leading to an increase of workload and to performance impairments (Robert, 1997; Paxion et al., 2014).

The aforesaid statistics and findings justify the increasing attention received by the Human Factor within the road safety research during the last decades. As well as in other human-centered domains such as aviation and industry (Vicente, 2013; Toppi et al., 2016; Vecchiato et al., 2016; Borghini et al., 2017a), psychological disciplines have been taken on a considerable scientific importance receiving more and more attention. They have become a fundamental instrument for understanding and interpreting the behavior of the driver (Bucchi et al., 2012), trying to provide cognitive models in order to predict and avoid unsafe actions as well as to understand the relationship between such unsafe behaviors and different factors related to traffic, road complexity, car equipment and external events. The most frequently adopted techniques in this research field are those based on questionnaires and interviews after large-scale experiments in naturalistic (i.e., real driving) and simulated (i.e., by using simulator) settings. They make it possible to acquire useful information for personality tests and profiles, they help to highlight and correct behavioral difficulties and, therefore, they shape the driver to have a safe relationship with driving in different conditions, and in particular in emergency situations, as well as to improve road and car design and adapt safety education with respect to the driver background (Cestac et al., 2014; Kaplan et al., 2015).

In order to increase the strength of such psychological research applied to road safety, this discipline could now benefit from recent advancements and outcomes coming from Neuroscience and Neuroergonomics. The field of the Neuroergonomics aims to study the relationship between the human behavior and the brain at work (Parasuraman and Rizzo, 2008). It provides a multidisciplinary translational approach that merges elements of neuroscience, cognitive psychology, human factors and ergonomics to study brain structure and function in everyday environments. Applied to the driving safety domain, a Neuroergonomic approach should allow to investigate the relationship between human mental behavior, performance and road safety, taking advantage from neurophysiological measures and providing a deeper understanding of human cognition and its role in decision making and possible error commission at the wheel (Lees et al., 2010). In fact, it is widely accepted in scientific literature the limit of using subjective measures alone, such as questionnaires and interview, because of their intrinsic subjective nature and the impossibility to catch the "unconscious" phenomena behind human behaviors (Gopher and Braune, 1984; Dienes, 2004; Wall et al., 2004; Aricò et al., 2017b). In this context, technological advancements enable the use of neurophysiological measures, for example the measure of brain activity, heart activity, eye movements, to obtain objective measures of specific mental states with low invasiveness (Aricò et al., 2017c). Among the several neuroimaging techniques, such as functional Magnetic Resonance and Magnetoencephalography, Electroencephalographic technique (EEG) has been demonstrated to be one of the best techniques to infer, even in real time, objective assessment of mental states and in particular the mental workload experienced by the user, since other than being a direct measure of brain activations, it is characterized by high temporal resolution, limited cost and invasiveness (Prinzel et al., 2000; Aricò et al., 2016b). EEG-based measures of drivers' mental states have been already investigated during the recent decades in order to determine brain cues of incoming risky psychophysical states, e.g., fatigue, drowsiness, inattention, overload (Lin et al., 2005; Michail et al., 2008; Brookhuis and de Waard, 2010; Borghini et al., 2012, 2014; Maglione et al., 2014; Wang et al., 2015; Zhang et al., 2015; Kong et al., 2015, 2017), and to develop futuristic Human-Machine interaction solutions and automation (Kohlmorgen et al., 2007; Lin et al., 2009; Göhring et al., 2013; Aricò et al., 2015). Nevertheless, two important gaps are still present in this domain:

(1) the majority of neurophysiological studies about drivers' behaviors have been conducted in simulated environments or in poor realistic settings, but it has been proven that same experimental tasks are perceived differently, in terms of mental workload, if performed in a simulator or in real environment (de Winter et al., 2014); also, not only the task perception but the driver behavior itself related to a specific condition could change if the same condition is reproduced in simulators or in a real scenario (Philip et al., 2005);

(2) in scientific literature there is still the lack of a synthetic EEG-based workload index to adopt in a systematic way within the road safety research, in order to integrate results coming from traditional techniques, such as subjective measures and car parameters analysis, with additional insights arising directly from drivers' brain (Paxion et al., 2014). Several studies about EEG correlates of driver's mental workload have been carried on, however experimental examples in real settings of a multimodal approach integrating neurophysiological with traditional measures are still lacking (Xing et al., 2018).

In this study, it has been investigated the possibility to adopt the approach recently developed and patented by the authors of this work (Aricò et al., 2016b, 2017a), to evaluate the mental workload experienced by car drivers by means of their EEG activity. More specifically, such an approach is based on a machine-learning method able to assess, even online and in high-realistic environments, the user's mental workload through a synthetic index. The authors successfully employed and validated such approach in different aviation-related applications, such as adaptive automation (Aricò et al., 2016a), personnel training (Borghini et al., 2017c), personnel expertise evaluation (Borghini et al., 2017b), moreover highlighting the higher sensitivity of such measures compared with subjective ones (Di Flumeri et al., 2015; Aricò et al., 2016b). Furthermore, the feasibility of obtaining EEG-based measures of driver's workload has already been validated through a pilot study of the present work conducted with eight subjects while performing a simplified version of the real driving task employed within the present work (Di Flumeri et al., 2018).

For the present work, 20 young subjects have been involved in a real driving task along urban roads, performed under different traffic conditions (rush hour and not) and going through different road types (main and secondary streets). Also, during the driving tasks specific events, in particular a pedestrian crossing the road and a car entering the traffic flow just ahead of the experimental subject, have been acted. During the experiments the drivers' brain activity, through EEG technique, and eye movements, through Eye-Tracking (ET) devices, have been collected. In addition, subjective measures, car parameters (e.g., position, speed, etc.) and videos around the car have been gathered. Thanks to this multimodal approach, the present study aimed at:

- Validating the machine-learning approach developed by the authors also in automotive domain, through an experiment in high-realistic settings, i.e., real driving;
- Employing the EEG-based Workload index obtained from the hence validated approach to evaluate the impact of different factors, specifically the road complexity, the traffic intensity (depending on the hour of the day), and two specific events (a pedestrian crossing the road and a car entering in the traffic flow), on the drivers' mental workload;
- Comparing the neurophysiological measures with eye movements and subjective ones, in order to provide evidence of the complementarity of the obtained insights.

In conclusion, the present work will explore the potential of integrating these new methodologies, i.e., neurophysiological measures, with traditional approaches in order to enhance and extent research on drivers' behaviors and road safety.

MATERIALS AND METHODS

The Experimental Protocol

Twenty male students (24.9 ± 1.8 years old, licensed from 5.9 ± 1 years, with a mean annual mileage of 10350 km/year) from the University of Bologna (Italy) have been recruited and involved on a voluntary basis in this study. They were selected in order to have a homogeneous experimental group in terms of age, sex, and driving expertise. The experiment was conducted following the principles outlined in the Declaration of Helsinki of 1975, as revised in 2000. Informed consent and authorization to use the video graphical material were obtained from each subject on paper, after the explanation of the study.

Two equal cars have been used for the experiments, i.e., Fiat 500L 1.3 Mjt, with diesel engine and manual transmission. The subjects had to drive the car along a route going through urban roads at the periphery of Bologna (Italy). In particular, the route consisted in three laps of a "circuit" about 2500 m long to be covered with the daylight (**Figure 1**).

The circuit was designed with the aim to include two segments of interest, both about 1000 m long but different in term of road complexity and so supposed different also in terms of cognitive demand, thus named hereafter "*Easy*" and "*Hard*": (i) Easy was a secondary road, mainly straight, with an intersection halfway with the right-of-way, one lane and low traffic capacity, serving a residential area; (ii) Hard was a main road, mainly straight, with two roundabouts halfway, three lanes and high traffic capacity, serving a commercial area. This factor will be hereafter named "*ROAD*." This assumption has been made on the basis of several evidences coming from scientific literature about road safety and behavior (Harms, 1991; Verwey, 2000; Paxion et al., 2014).

Furthermore, each subject had to repeat the task two times within the same day, one time during rush and one during normal hour: this factor will be hereafter named "*HOUR,*" while the

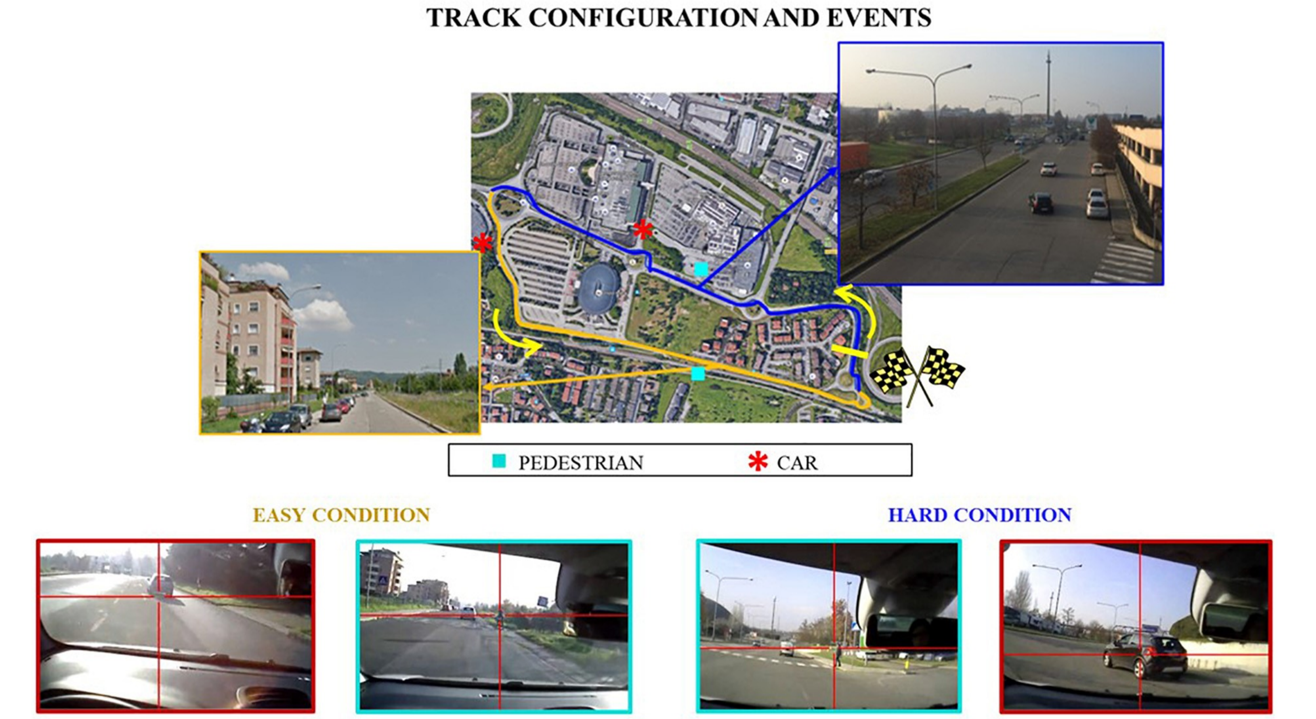

FIGURE 1 | The experimental circuit about 2500 m long along Bologna roads. The blue line indicates the circuit segment labeled as "Hard" in terms of road complexity, while the yellow one the "Easy" segment. The cyan squares and the red asterisks represent the points were the events, respectively the pedestrian and the car, have been acted along the 3rd lap of both the task repetitions.

TABLE 1 | Data extracted from the General Plan of Urban Traffic of Bologna (Italy) referred to the traffic flow intensities in the experimental area during the day.

Transits	Total	RUSH hour		NORMAL
	14 h (6 20)	Morning (12:30–13:30)	Afternoon (16:30–17:30)	12 h
Total	19385	2024	2066	15295
Frequency (Transits/hour)	–	2024	2066	1274,6

These data have been used to design two experimental conditions different in terms of traffic: the RUSH hours are characterized by traffic higher than during NORMAL hours.

two conditions "*Rush*" and "*Normal.*" The rush hours of that specific area have been determined according to the General Plan of Urban Traffic of Bologna (PGTU, please see **Table 1**): the two "Rush hour" time-windows were from 12:30 to 13:30 (lunchtime) and from 16:30 to 17:30 (work closing time), with the experiments performed from 9.30 to 17.30, in order to ensure a homogeneous daylight condition.

Finally, during the last lap (i.e., the 3rd one) of each task repetition (i.e., Rush and Normal hour) two different events have been simulated, by involving actors, twice (i.e., along the Hard and the Easy circuit segment) along the route: a pedestrian crossing the road, and a car entering the traffic flow just ahead of the experimental subject, hereafter labeled respectively "*Pedestrian*" and "*Car.*" The event types have been selected as the most probable events coherently with the urban context, as well as the safest to act, i.e., without introducing any risk for the actors, for the experimental subjects and for the traffic in general.

The **Figure 1** shows the experimental circuit along Bologna roads, highlighting the "ROAD complexity" distribution as well as the occurred events.

To summarize, each subject, after a proper experimental briefing, performed a driving task of three laps along a circuit through urban roads two times, during Rush and Normal hours. The order of Rush and Normal conditions has been randomized among the subjects, in order to avoid any order effect (Kirk, 2015). Each lap consisted in a Hard and an Easy segment, where hard and easy are referred to the road complexity and thus task difficulty. Also, despite the initial briefing, the first lap of both the tasks has been considered an "adaptation lap," while the data recorded during the second and third laps have been taken into account for the analysis. Finally, during the third lap two equal events have been simulated both along the Easy and the Hard segment (i.e., four events in total for each subject for each task, Rush and Normal).

The **Figure 2** shows a graphical representation of the experimental protocol.

During the whole protocol physiological data, in terms of brain activity through Electroencephalographic (EEG) technique and eye gazes through ET devices, and data about driving

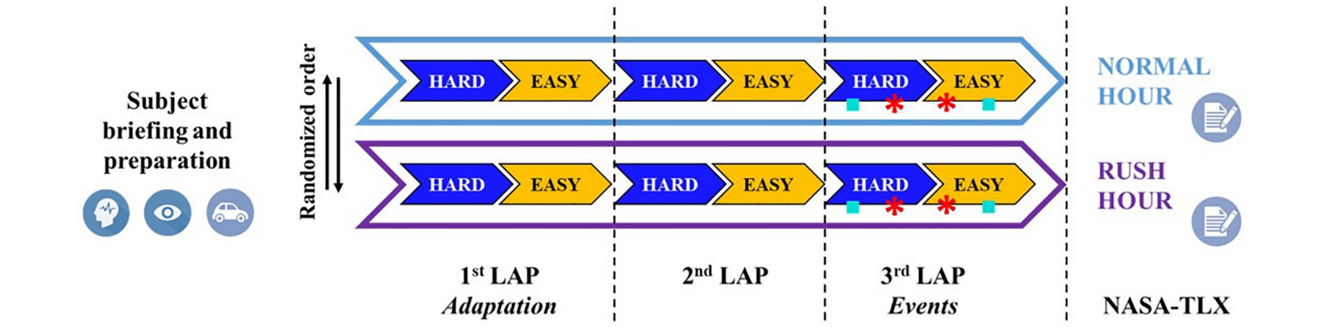

FIGURE 2 | Overview of the experimental protocol, consisting of two main driving tasks different in terms of traffic (normal and rush hour) and performed in a randomized order. Each task consisted of three laps along the circuit in **Figure 1**: the first lap aimed to allow the driver to take confidence with the circuit, while the second and third lap have been used for analysis. In particular, during the third lap four events have been acted as indicated in **Figure 1**. Before the experiment the participant received a briefing and was equipped by EEG and Eye Tracking devices, while his car with Video VBOX system. At the end of each task the participant had to fill a questionnaire (NASA-TLX) about the experienced mental workload.

behavior, through a professional device mounted on the car (i.e., a VBOX Pro), have been recorded. In addition, subjective measures of perceived Mental Workload have been collected from the subjects after both the tasks through the NASA Task Load Index (NASA-TLX) questionnaire (Hart and Staveland, 1988). It was possible to use Eye Tracker just with half of the subjects' sample (i.e., 10 subjects) because of device availability, so eye tracker–related data have been analyzed for 10 subjects. The following paragraphs will describe in detail the collection and processing of the aforementioned data, while the **Figure 3** shows the subject preparation and the recording setup within the car.

The Data Collection

Electroencephalographic Signal Recording and Processing

The EEG signals have been recorded using the digital monitoring BEmicro system (EBNeuro, Italy). Twelve EEG channels (FPz, AF3, AF4, F3, Fz, F4, P3, P7, Pz, P4, P8, and POz), placed according to the 10–20 International System, were collected with a sampling frequency of 256 Hz, all referenced to both the earlobes, grounded to the Cz site, and with the impedances kept below 20 kΩ. During the experiments the EEG data have been recorded without any signal conditioning, the whole processing chain has been applied offline. In particular, EEG signal has been firstly band-pass filtered with a fourth-order Butterworth filter (high-pass filter cut-off frequency: 1 Hz, low-pass filter cut-off frequency: 30 Hz). The Fpz channel has been used to remove eyes-blink contributions from each channel of the EEG signal by using the REBLINCA algorithm (Di Flumeri et al., 2016). This step is necessary because the eyes-blink contribution could affect the frequency bands correlated to the mental workload, in particular the theta EEG band. This method allows to correct EEG signal without losing data.

For other sources of artifacts (i.e., environmental noise, drivers' movements, etc.), specific procedures of the EEGLAB toolbox (Delorme and Makeig, 2004) have been employed. Firstly, the EEG signal is segmented into epochs of 2 s (Epoch length), through moving windows shifted of 0.125 s (Shift), thus with an overlap of 0.875 s between two contiguous epochs. This windowing has been chosen with the compromise to have both a high number of observations, in comparison with the number

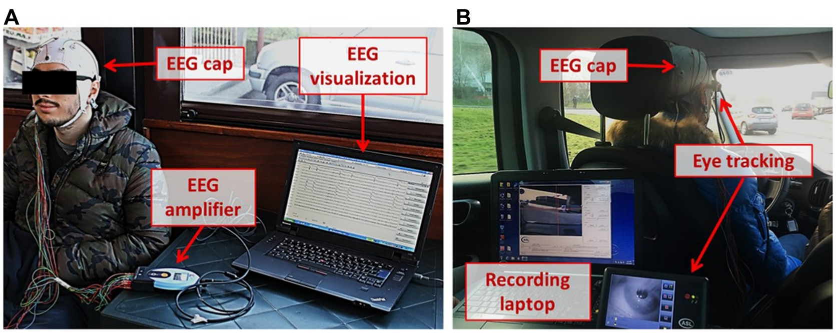

FIGURE 3 | On the left **(A)**, the participant preparation phase. In particular, the EEG signal has been acquired through the EEG amplifier in *holter* modality: the EEG signal and the electrodes impedances were checked on a computer before starting the experiments. On the right **(B)**, a picture representing the experimental setup within the car: in particular, other than the EEG cap, also the Eye Tracking device and its recording laptop are shown. The subject in picture, as all the participants, gave their signed authorization to use the video graphical material for dissemination purposes.

of variables, and to respect the condition of stationarity of the EEG signal (Elul, 1969). In fact, this is a necessary assumption in order to proceed with the spectral analysis of the signal. The EEG epochs with the signal amplitude exceeding ± 100 µV (*Threshold criterion*) are marked as "artifact." Then, each EEG epoch has been interpolated in order to check the slope of the trend within the considered epoch (*Trend estimation*). If such a slope is higher than 10 µV/s, the considered epoch is marked as "artifact." Finally, the signal sample-to-sample difference (*Sample-to-sample criterion*) has been analyzed: if such a difference, in terms of absolute amplitude, is higher than 25 µV, i.e., an abrupt variation (no-physiological) happened, the EEG epoch is marked as "artifact." At the end, the EEG epochs marked as "artifact" have been removed from the EEG dataset with the aim to have a clean EEG signal to perform the analyses.

From the clean EEG dataset, the Power Spectral Density (PSD) has been calculated for each EEG channel for each epoch using a Hanning window of the same length of the considered epoch (2 s length, that means 0.5 Hz of frequency resolution). Then, the EEG frequency bands of interest has been defined for each subject by the estimation of the *Individual Alpha Frequency* (IAF) value (Klimesch, 1999). In order to have a precise estimation of the alpha peak and, hence of the IAF, the subjects were been asked to keep the eyes closed for a minute before starting the experimental tasks. Finally, a spectral features matrix (EEG channels × Frequency bins) has been obtained in the frequency bands directly correlated to the mental workload. In particular, only the theta band [IAF – 6 ÷ IAF – 2], over the EEG frontal channels, and the alpha band [IAF – 2 ÷ IAF + 2], over the EEG parietal channels, were considered as variables for the mental workload evaluation (Gevins and Smith, 2003; Aricò et al., 2016b; Borghini et al., 2017a).

At this point the automatic-stop-StepWise Linear Discriminant Analysis (asSWLDA), a specific Machine-Learning algorithm (basically an upgrade version of the well-known StepWise Linear Discriminant Analysis) previously developed (Aricò et al., 2016b), patented (Aricò et al., 2017a) and applied in different applications (Aricò et al., 2016a; Borghini et al., 2017b,c) by the authors has been employed. On the basis of the calibration dataset, the asSWLDA is able to find the most relevant spectral features to discriminate the Mental Workload of the subjects during the different experimental conditions (i.e., EASY = 0 and HARD = 1). Once identified such spectral features, the asSWLDA assigns to each feature specific weights ($w_{i\ train}$), plus a bias (b_{train}), such that an eventual discriminant function computed on the training dataset [$y_{train}(t)$] would take the value *1* in the hardest condition and *0* in the easiest one. This step represents the calibration, or "*Training phase*" of the classifier. Later on, the weights and the bias determined during the training phase are used to calculate the Linear Discriminant function [$y_{test}(t)$] over the testing dataset (*Testing phase*), that should be comprised between 0 (if the condition is Easy) and *1* (if the condition is Hard). Finally, a moving average of 8 s (8MA) is applied to the $y_{test}(t)$ function in order to smooth it out by reducing the variance of the measure: its output is defined as the *EEG-based Workload index* (WL_{SCORE}). For the present work, the training data consisted in the *Easy* segment of the 2nd lap during the *Normal* condition and the *Hard* segment of the 2nd lap during the *Rush* condition (they have been hypothesized the two conditions characterized by respectively the lowest and highest mental workload demand), while the testing data consisted of the data of the 3rd lap of both the conditions.

Here below the training asSWLDA discriminant function (Equation 1, where $f_{i\ train}(t)$ represents the PSD matrix of the training dataset for the data window of the time sample *t*, and of the i^{th} feature), the testing one (Equation 2, where $f_{i\ test}(t)$ is as $f_{i\ train}(t)$ but related to the testing dataset) and the equation of the *EEG-based workload index* computed with a time-resolution of 8 s (WL_{SCORE}, Equation 3), are reported.

$$y_{train}(t) = \sum\nolimits_i w_{i\,train} \cdot f_{i\,train}(t) + b_{train} \tag{1}$$

$$y_{test}(t) = \sum\nolimits_i w_{i\,train} \cdot f_{i\,test}(t) + b_{train} \tag{2}$$

$$WL_{SCORE} = 8MA(y_{test}(t)) \tag{3}$$

Eye-Tracking Data and Its Processing

Eye movements of the participants have been recorded through an ASL Mobile Eye-XG device (EST GmbH, Germany), a system based on lightweight eyeglasses equipped with two digital high-resolution cameras. One camera recorded the scene image and the other the participant's eye, that is monitored through infrared rays. The data were recorded with a sampling rate of 30 Hz (i.e., 33 ms time resolution), and a spatial resolution of 0.5 ÷ 1°. ASL software was used to analyze the data, obtaining information about the drivers' fixation points frame by frame (33 ms). A preliminary calibration procedure was carried out for each subject inside the car before starting driving, asking them to fix their gaze on thirty fixed visual points spread across the whole scene, in order to get a good accuracy of the eye-movement recorder. The gazes recorded during the driving task were manually analyzed, in order to group them into three different categories: road infrastructure, traffic vehicles, and external environment. For each subject, each lap (second and third), and each condition (Easy and Hard ROAD, Rush and Normal HOUR) the distribution of eye fixations between the three categories was calculated in terms of percentage of the total.

Additional Measures

Each car has been equipped with a Video VBOX Pro (Racelogic Ltd., United Kingdom), a system able to continuously monitor the cinematic parameters of the car, integrated with GPS data and videos coming from up to four high-resolution cameras. The system has been fixed within the car, at the center of the floor of the back seats, in order to put it as close as possible to the car barycenter, while two cameras have been fixed over the top of the car. The system recorded car parameters (e.g., speed, acceleration, position, etc.) with a sampling rate of 10 Hz. For the purpose of the present study, the average speed for each task has been computed. Also, the cameras' videos have been used to count the number of vehicles encountered by the driver during each task.

Also, at the end of each task (thus only the HOUR condition, i.e., Rush vs. Normal, can be compared) the subjects had to evaluate the experienced workload by filling the NASA-TLX questionnaire (Hart and Staveland, 1988). In particular, the subject had (i) to assess, on a scale from 0 to 100, the impact of six different factors (i.e., Mental demand, Physical demand, Temporal demand, Performance, Effort, Frustration), and (ii) to assess the more impacting factor through 15 comparisons between couple of the previously evaluated factors. The result of this questionnaire is a score from 0 to 100 corresponding to the driver's mental workload perception.

Performed Analyses

Validation of Experimental Design Assumptions

The first analysis aimed to validate the assumptions in terms of experimental design, that is:

(i) The subjects drove during two conditions different in terms of traffic, i.e., Rush and Normal hour;
(ii) The circuit was constituted by two segments different in terms of road complexity, thus in terms of difficulty, i.e., Hard and Easy.

In order to validate the first assumption, the number of vehicles encountered by the experimental subjects and the average driving speed during the two conditions have been computed and statistically compared. It is expected that the number of vehicles is significantly higher and the average speed significantly lower during rush hours (Bucchi et al., 2012).

The second assumption has been validated by investigating the percentage of fixations over the external environment, since such indicator has been proven to be inversely correlated with mental workload: the more the experienced workload is, the less the number of fixations over the external environment is, since the driver gaze will mostly focus on infrastructure and vehicles (Costa et al., 2014; de Winter et al., 2014; Lantieri et al., 2015). Also, we verified the difference in terms of mental workload from a neurophysiological point of view: we computed the ratio between Theta rhythms over frontal sites ("*ThetaF*") and Alpha rhythms over parietal sites ("*AlphaP*"), since it is considered a well-established metric of mental workload (Borghini et al., 2014). In particular, The *ThetaF/AlphaP* has been proven to increase if the mental workload experienced by the user is increasing as well (Gevins and Smith, 2003; Holm et al., 2009; Borghini et al., 2015). The metric has been computed as the ratio between the averaged PSD values in theta band over the frontal electrodes (AF3, AF4, F3, Fz, F4) and the averaged PSD values in alpha band over the parietal electrodes (P3, P7, Pz, P4, P8, POz). Both the analysis have been performed comparing the two conditions employed to train the classifier (please see Electroencephalographic Signal Recording and Processing), i.e., the *Easy* segment of the 2nd lap during the *Normal* condition and the *Hard* segment of the 2nd lap during the *Rush* condition, assumed as the two conditions characterized by respectively the lowest and highest mental workload demand.

All the statistical comparisons have been performed through two-sided Wilcoxon signed rank tests. In fact, data come from multiple observations on the same subjects, but it is not possible to assume or robustly assess (the number of observations is always equal or less than 16) that the observations distribution is Gaussian, therefore paired non-parametric tests have been used (Siegel, 1956).

Classification Performance

Firstly, a synthetic analysis of the brain features selected by the algorithm has been performed in order to evaluate any eventual recurrence of a specific feature. The initial features domain for each subject consisted in a matrix of 187 features (11 EEG channels * 17 bins of frequency – from IAF-6 Hz to IAF+2 Hz with a resolution of 0.5 Hz –). Actually, only 99 of these features can be selected by the algorithm because of the Regions of Interest defined *a priori*: 45 features related to frontal Theta and 54 related to parietal Alpha.

Then, in order to investigate the algorithm (i.e., the asSWLDA) classification accuracy, the analysis of the Area Under Curve (AUC) of the Receiver Operator Characteristic (ROC) curve of the classifier has been performed (Bamber, 1975). In particular, AUC represents a widely used methodology to test the performance of a binary classifier: the classification performance can be considered good with an AUC higher than at least 0.7 (Fawcett, 2006). In this case there are actually two classes in terms of mental workload, i.e., Easy and Hard, related to the two different difficulty levels characterizing the circuit. As previously described, for each subject the training dataset consisted in the Easy segment of the 2nd lap during the Normal condition and the Hard segment of the 2nd lap during the Rush condition (they have been hypothesized the two conditions characterized by respectively the lowest and highest cognitive demand), while the testing dataset consisted of the data of the 3rd lap of both the conditions (*Real data*). Therefore, the classifier has been tested shuffling the testing dataset related labels (*Random*), in order to verify that classifier performance on measured data (*Real data*) was significantly higher than that one obtained on random data (*Random*), independently from the traffic intensity (i.e., both in Rush and Normal hour conditions). In both the cases (Real and Random), the time resolution of WL_{scores} is equal to 8 s, obtained as the best compromise between a high time resolution and good classification performance. Three two-sided Wilcoxon signed rank tests have been performed between *Real* and *Random data*, one for each HOUR condition (i.e., comparison Real vs. Random in *Normal* and *Rush* hour) and one comparing the *Normal* and *Rush* conditions only in terms of real data. The results of these multiple comparisons have been validated by applying the False Discovery Rate (FDR) correction (Benjamini and Hochberg, 1995).

Workload Assessment

Once demonstrated the reliability of the classification algorithm to obtain the EEG-based index of mental workload in the specific driving scenarios, the workload scores (*WL score*) have been used to evaluate the impact of different factors, that is the road complexity and the traffic as well as specific events along the driving experience. Depending on the analysis, the EEG-based WL scores have been analyzed in relation to ET and subjective data.

Evaluation of traffic and road complexity impact

The WL indexes obtained with a time resolution of 8 s from the testing dataset (i.e., the third lap) were averaged for each subject and for each condition (i.e., HOUR and ROAD). A Friedman test, the non-parametric version of the repeated measures ANOVA (Analysis of Variance), has been performed in order to investigate any possible effect due to traffic and road complexity on the workload perceived by the subject. Furthermore, since *post hoc* tests specifically designed for Friedman test do not exist but both the factors have been measured on the same subjects, two Wilcoxon signed rank tests have been performed in order to investigate potential within effects among the two factors, i.e., HOUR and ROAD.

Also, the results in terms of workload indexes have been compared with those obtained from ET in order to evaluate the different sensitivity to the phenomenon (i.e., mental workload variations) of the two technologies. In terms of ET measures, it has been investigated the percentage of fixations on the road infrastructure and vehicles, since such indicator has been proven to be directly correlated with mental workload while driving: the more the experienced workload is, the more the number of fixations over the road will be, since the driver gaze will mostly focus on infrastructure and vehicles (Costa et al., 2014; de Winter et al., 2014). Multiple two-sided Wilcoxon signed rank tests have been performed in order to reveal any difference with respect to the two investigated factors.

Furthermore, a two-sided Wilcoxon signed rank test has been performed on the NASA-TLX measures. Please note that for the continuity of the experiment the questionnaires were filled by the subjects only after the tasks end, therefore only the comparison between Normal and Rush hour has been possible (please refer to Section "Additional Measures").

Evaluation of single events impact

On the basis of the average duration of the events among the subjects during the driving experience, and to homogenize the measures with respect of this parameter (i.e., event duration), a fixed window of 20 s for the car event (from the first fixation of the car to its overtaking) and of 10 s for the pedestrian event (from the first fixation of the pedestrian to the acceleration after its road crossing) has been defined, independently from the traffic and the road complexity. Remembering that the events were acted only during the third lap of each task repetition, similar windows corresponding to the same circuit position were defined during the second lap in order to compare the event's happening vs. no-happening. The WL indexes were averaged for each subject, for each condition (i.e., HOUR and ROAD) and for each event. Multiple two-sided Wilcoxon signed rank tests have been performed in order to reveal any difference (i) with respect to the events' happening, and (ii) among the events types.

RESULTS

The following results are referred to a sample of 16 subjects (8 with Eye Tracking), since one subject has been discarded because of technical issues on the EEG data, while three subjects have been discarded because of no objective difference in terms of encountered vehicles (measured through the VBOX cameras) between the two tasks, i.e., during Rush and Normal hours.

Experimental Design Validation

Figure 4 shows the results of the comparisons between (a) the number of vehicles encountered by the experimental subjects and (b) the average driving speed during the two different traffic conditions, i.e., during *Normal* and *Rush* hours. The performed statistical analysis revealed a significant increasing (p = 0.001) of vehicles encountered by the experimental subjects and a significant decreasing (p = 0.039) of driving average speed from

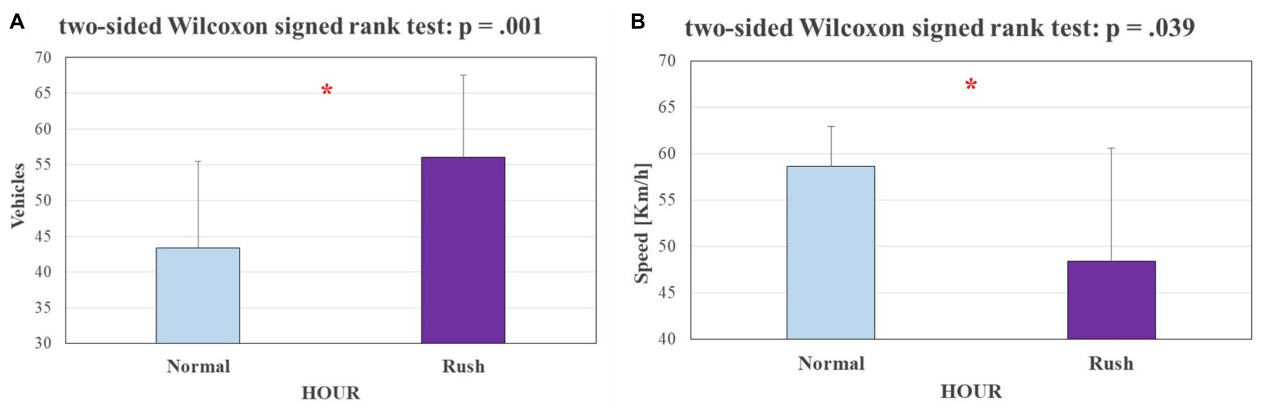

FIGURE 4 | On the left **(A)**, a bar graph representing the mean and the standard deviation of vehicles encountered by the participants during the experiments. The Wilcoxon test showed a significantly higher (p = 0.001) number of vehicles during rush hour. On the right **(B)**, a bar graph representing the mean and the standard deviation of participants driving speed during the experiments. The Wilcoxon test showed a significantly lower (p = 0.039) speed during rush hour. The statistical tests showing a significant effect.

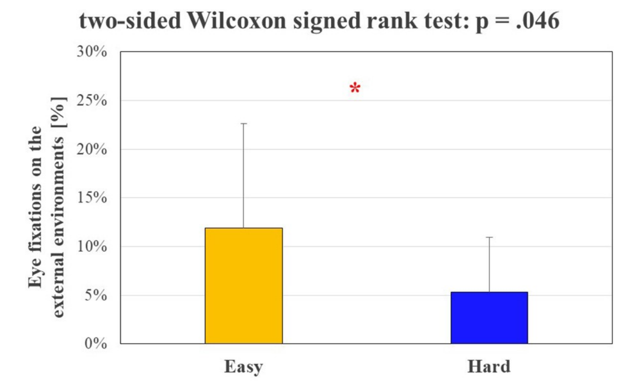

FIGURE 5 | The bar graph represents the mean and the standard deviation of the percentage of drivers' eye fixations along the two different segments of the circuit. The Wilcoxon test showed a significant reduction ($p = 0.046$) of such percentage during the circuit segment characterized by hard complexity. The statistical tests showing a significant effect.

Normal to *Rush* hours, validating the experimental hypothesis about the two different conditions of traffic made *a priori* on the basis of the General Plan of Urban Traffic of Bologna (see The Experimental Protocol).

Figure 5 shows the results in terms of percentage of fixations over the external environment between the *Easy* and *Hard* segments of the circuit, since such indicator has been proven to be inversely correlated with mental workload. The performed statistical analysis revealed a significant decreasing ($p = 0.046$) of driver gazes over the external environment, validating the experimental hypothesis about the two different conditions of difficulty made *a priori* on the basis of scientific literature (see The Experimental Protocol).

Figure 6 shows the results in terms of *ThetaF/AlphaP* value between the *Easy* and *Hard* segments of the circuit, since such ratio has been proven to be a physiological indicator directly correlated to mental workload. The performed statistical analysis revealed a significant increasing ($p = 0.009$) of the proposed index, validating the assumption about the different cognitive demand related to the two conditions, made *a priori* on the basis of scientific literature (see The experimental Protocol).

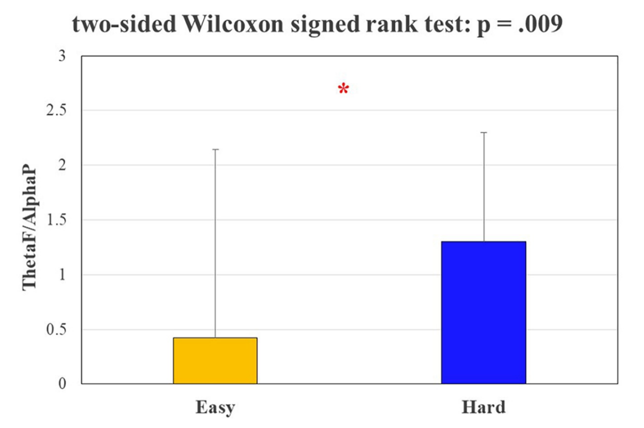

FIGURE 6 | The bar graph represents the mean and the standard deviation of the EEG-based ThetaF/AlphaP indicator along the two different segments of the circuit. The Wilcoxon test showed a significant increasing ($p = 0.009$) of such indicator during the circuit segment characterized by hard complexity. The statistical tests showing a significant effect.

Classification Performance

Figure 7 shows the distribution of the features, and the relative frequency of selection, chosen by the asSWLDA during the training phase. The analysis of features selected by the algorithm revealed that the asSWLDA selected on average 4 features per subject, coming from 3 of the 11 channels available. The frequency bins, actually equal to 17 because included between IAF-6 Hz and IAF+2 Hz with a resolution of 0.5 Hz, have been grouped into four areas of interest: Lower Theta [IAF – 6 ÷ IAF – 4], Upper Theta [IAF – 4 ÷ IAF – 2], Lower Alpha [IAF – 2 ÷ IAF] and Upper Alpha [IAF ÷ IAF + 2]. The results show that Lower Theta over F4 and Upper Alpha over POz have been used for more than the 50% of subjects.

The AUC analysis (**Figure 8**) revealed that, by using such approach, it has been possible to achieve mean AUC values of 0.744 ± 0.13 for the *Normal* hour and of 0.727 ± 0.06 for the *Rush* hour. In particular, the two Wilcoxon tests demonstrated that the classifier performance on the Real data was significantly higher than on Random data in both the conditions (respectively $p = 0.01$ and $p = 0.0005$). Also, there were no significant differences ($p = 0.64$) in terms of AUC values on Real data between *Normal* and *Rush* hours, in other words the classification performance was not dependent on the traffic condition. Because of the three repeated tests, the False Discovery Rate correction has been performed: with respect to the p-values obtained and ordered (0.0005, 0.01, and 0.64), the three corrected q-values are respectively 0.0015, 0.015, and 0.64, thus the first two results are still significant.

Workload Assessment

Evaluation of Traffic and Road Complexity Impact

Figure 9 shows the results of the non-parametrical statistical analysis in terms of effects of the two investigated factors, i.e., the traffic (HOUR) and the road complexity (ROAD), on the mental workload experienced by the drivers. In particular, the Friedman test at the top of **Figure 9A** highlights a significant main effect ($p = 0.00001$) among the different factors: the mental workload significantly increased because of the higher road complexity (i.e., from Easy to Hard), and even more because of the higher traffic intensity (i.e., from Normal to Rush hours). The Wilcoxon tests performed in order to investigate any *within* effect showed two significant main effects in term of workload increasing if both complexity [bottom left (**Figure 9B**), ROAD, $p = 0.0038$] and traffic [bottom right (**Figure 9C**), HOUR, $p = 0.0032$] increase.

Figures 10, **11** show the results of the Wilcoxon tests comparing the sensitivity of ET measures with respect to EEG-based ones. For these analyses the EEG-based WL scores of only the subjects wearing also the Eye Tracker (eight of sixteen) have been considered, in order to make the results comparable (i.e.,

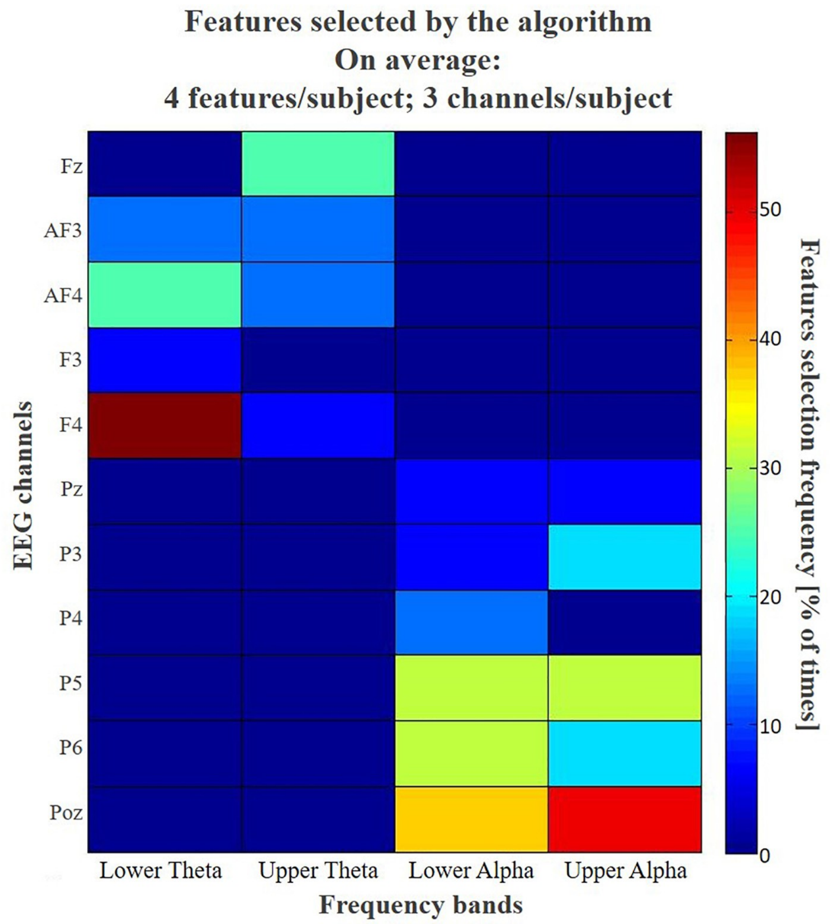

FIGURE 7 | The colormap summarizes the frequency of the selection of each feature for the whole subjects' sample. The initial features domain for each subject consisted of a matrix of 187 features (11 EEG channels * 17 bins of frequency – from IAF – 6 Hz to IAF+2 Hz with a resolution of 0.5 Hz –). Actually, only 99 of these features can be selected by the algorithm, because of the Regions of Interest defined *a priori*: 45 features related to frontal Theta and 54 related to parietal Alpha. For a synthetic and effective representation, the frequency bins, actually equal to 17 because included between IAF-6 Hz and IAF+2 Hz with a resolution of 0.5 Hz, have been grouped into four areas of interest: Lower Theta [IAF – 6 ÷ IAF – 4], Upper Theta [IAF – 4 ÷ IAF – 2], Lower Alpha [IAF – 2 ÷ IAF] and Upper Alpha [IAF ÷ IAF + 2]. The results show that Lower Theta over F4 and Upper Alpha over POz have been used for more than the 50% of subjects.

both the measures have been collected during same experience). In particular:

- **Figure 10**: in terms of ROAD complexity, while the EEG-based measures have been able to significantly discriminate ($p = 0.008$) the two conditions at least during Normal hour, the ET-based ones have not been able to show any significant difference both during Normal and Rush hours;
- **Figure 11**: in terms of traffic HOUR, while the EEG-based measures have been able to significantly discriminate the two conditions both along Easy ($p = 0.019$) and Hard ($p = 0.039$) segments, the ET-based ones have been able to significantly discriminate Normal and Rush hours only along the Hard segment ($p = 0.0192$).

Finally, **Figure 12** shows the results in terms of NASA-TLX scores, revealing that there is not any significant difference in terms of workload subjectively assessed between the *Normal* and *Rush* hour conditions.

Evaluation of Single Events Impact

Figure 13 shows the results in terms of EEG-based WL scores about how the presence of a specific event impacts the mental workload of the driver, with respect to the different experimental conditions. In terms of external events (the condition *EVENT* is referred to the event actually happened during the 3rd lap, the condition *NO EVENT* is referred to the same circuit portion during the 2nd lap when no events were acted), the pedestrian crossing the road induced a significantly higher workload only during the *Normal* hour along the *Hard* circuit segment (Wilcoxon test's $p = 0.037$), while the car induced a significantly higher workload along both the *Easy* and *Hard* circuit segments but only during *Normal* hour (respectively Wilcoxon test's $p = 0.007$ and $p = 0.008$).

Considering only the condition "EVENT," despite a decreasing trend from *Easy* to *Hard* segments, no significant differences ($p > 0.05$) have been found for each event during the same traffic

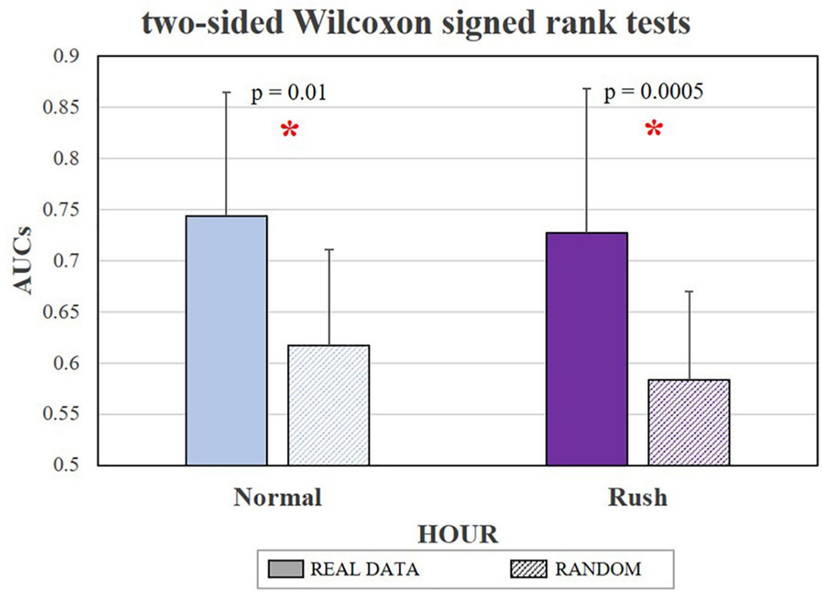

FIGURE 8 | The bar graph represents the mean and the standard deviation of AUC values obtained in discriminating the Easy and Hard circuit segments. Both in Normal and Rush hour conditions, the classification performance obtained by training the classifier with real data (solid color) have been significantly higher (respectively $p = 0.01$ and $p = 0.0005$) than by using random data (lines pattern), achieving mean AUC values of respectively 0.74 and 0.73. The statistical tests showing a significant effect.

condition (HOUR). However, if considering the same difficulty level (ROAD), all the events induced a significant workload increasing during *Rush* hours, except the pedestrian along the Easy segment (Pedestrian *Hard*: $p = 0.009$; Car *Easy*: $p = 0.023$; Car *Hard*: $p = 0.002$).

DISCUSSION

Since the impact of drivers' errors in terms of human lives and costs is very high and the next future previsions are even worse (World Health Organization, 2015), the relationship between human errors and driving performance impairment due to a high mental workload has been deeply investigated in the automotive domain. Recent technological advancements as well as the growth of disciplines such as Neuroscience and Neuroergonomics now allow to record human neurophysiological signals, such as in this study brain activity through Electroencephalographic technique, in a robust way also outside the laboratory, and to obtain from them objective neurometrics of human mental states (i.e., workload) (Aricò et al., 2017c, 2018). The present work aimed to validate a machine-learning approach, i.e., the asSWLDA (Aricò et al., 2016a), for the objective assessment of human mental workload while driving in real settings, as well as its integration with traditional tools (e.g., questionnaires, car parameters, eye tracking) in order to evaluate the impact of different factors (road complexity, traffic intensity, external events), thus suggesting new innovative tools for enhancing research in road safety. In order to achieve these objectives, 20 young subjects have been involved in a real driving task along urban roads, performed under different traffic conditions (rush hour and not), driving through different road types (main and secondary streets) and facing to external events.

Firstly, the experiments have been designed making two *a priori* assumptions:

(1) the experiments have been conducted in two different conditions of traffic intensity, depending on the hours (i.e., normal and rush hour) of the day; the experimental design initially referred to the General Plan of Urban Traffic of Bologna;
(2) the circuit consisted of two segments of different difficulty, i.e., Easy and Hard, because of the related road complexity (Harms, 1991; Verwey, 2000; Paxion et al., 2014).

The statistical analysis performed on the average speed of the experimental subjects and the number of vehicles encountered during the experiments (**Figure 4**) validated the Assumption 1: in fact, the subjects encountered a significantly higher ($p = 0.001$) number of vehicles and they drove at a significantly lower ($p = 0.039$) speed during the rush hours, as expected from scientific literature (Bucchi et al., 2012). Statistical analysis of driver's eye fixations over the external environment (**Figure 5**) and physiological brain patterns (**Figure 6**) validated the Assumption 2: in fact the drivers' gazes over the external environment (such index inversely correlates with mental workload; de Winter et al., 2014; Lantieri et al., 2015) have been significantly lower ($p = 0.046$) along the circuit segment that was hypothesized as *Hard*, while the ratio between frontal theta and parietal alpha rhythms significantly increased ($p = 0.009$). These results confirmed the properness of the experimental design. Nevertheless, the analysis of encountered vehicles, determined

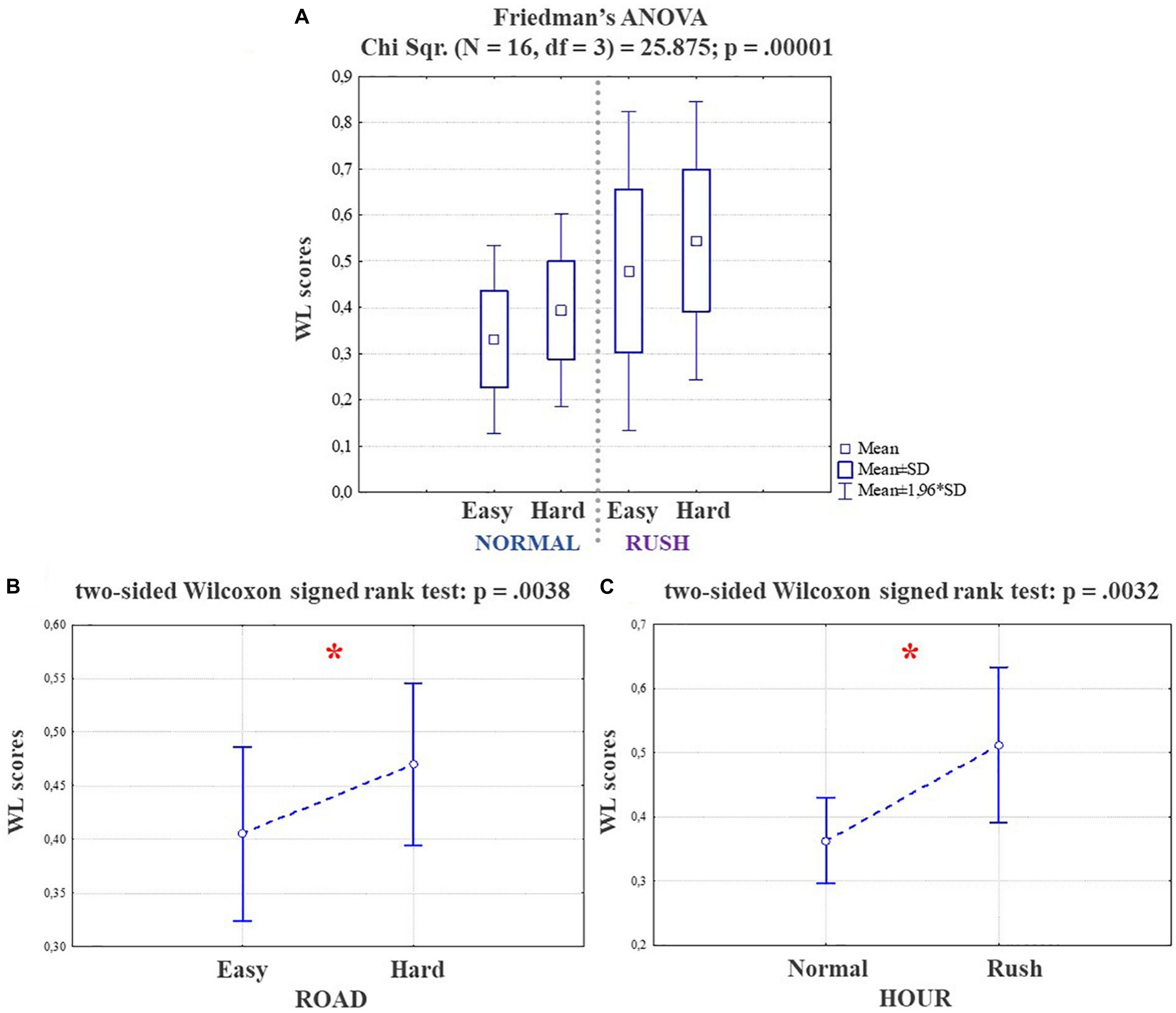

FIGURE 9 | At the top **(A)**, the Friedman test highlighting a significant main effect ($p = 0.00001$), in terms of mental workload increasing among the different factors. At the bottom, on the left **(B)** the Wilcoxon test on the factor ROAD and on the right **(C)** the same test on the factor HOUR, showing how both the factors produced a significant mental workload increasing (respectively $p = 0.004$ and $p = 0.003$). The statistical tests showing a significant effect.

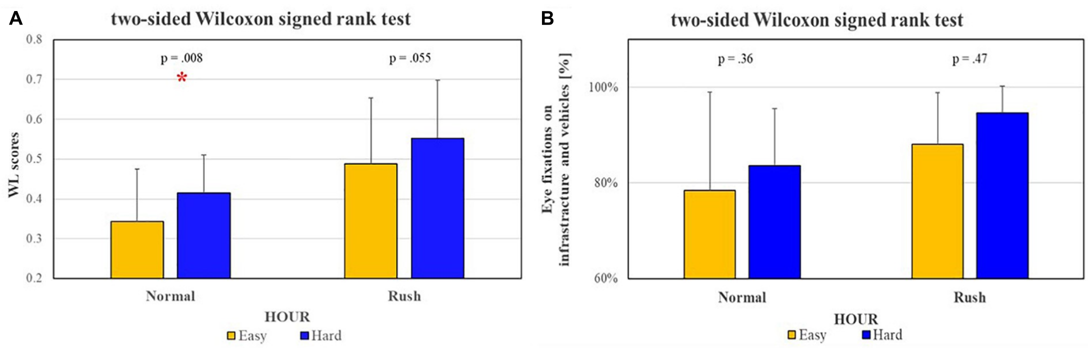

FIGURE 10 | The Wilcoxon tests performed to investigate eventual sensitivity differences between Eye-Tracking [left **(A)**] and EEG [right **(B)**] measures, considered on the same subjects, in relation to ROAD complexity showed that EEG-based measures have been able to significantly discriminate ($p = 0.008$) the two conditions at least during Normal hour, while the ET-based ones have not been able to show any significant difference both during Normal and Rush hours. The statistical tests showing a significant effect.

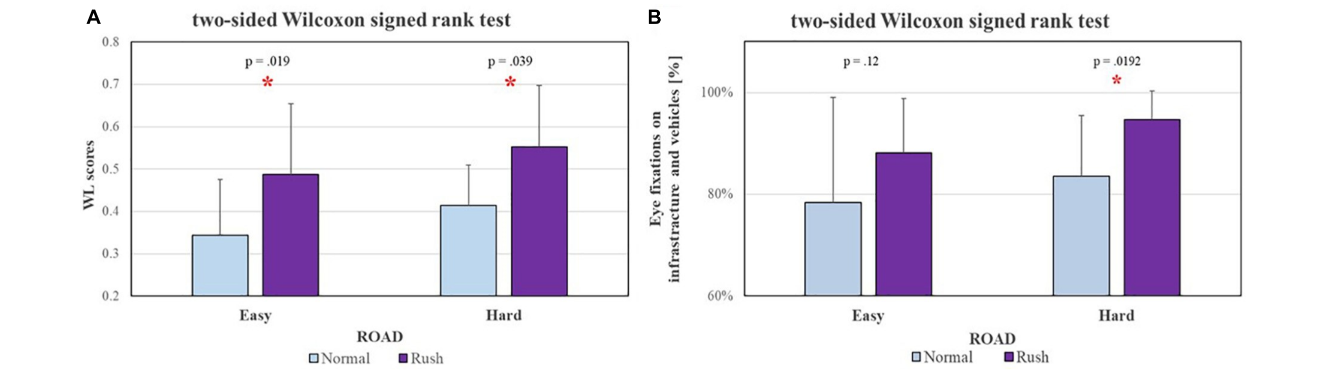

FIGURE 11 | The Wilcoxon tests performed to investigate eventual sensitivity differences between Eye-Tracking [left **(A)**] and EEG [right, **(B)**] measures, considered on the same subjects, in relation to traffic intensity (i.e., HOUR) showed that, while the EEG-based measures have been able to significantly discriminate the two conditions both along Easy (p = 0.019) and Hard (p = 0.04) segments, the ET-based ones have been able to significantly discriminate Normal and Rush hours only along the Hard segment (p = 0.02). The statistical tests showing a significant effect.

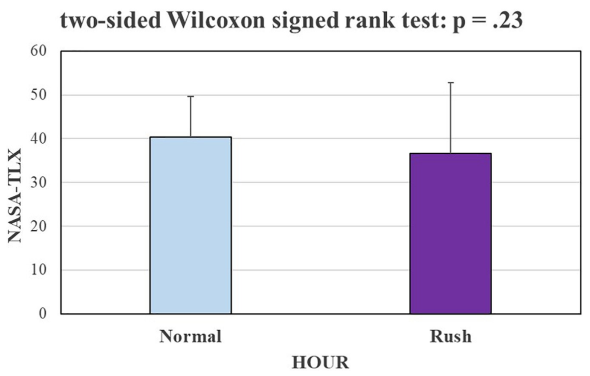

FIGURE 12 | The bar graph represents the mean and the standard deviation of NASA-TLX scores, i.e., the subjective assessment of the mental workload experienced by the participants of the circuit. The Wilcoxon test does not reveal any significant difference in terms of workload subjectively assessed by the subjects between the Normal and Rush hour conditions.

through videos from the VBOX videos, led to discard three subjects because of no differences between rush and normal hours (see Results). Therefore, this validation approach should be taken into account for future works in real driving conditions, where external conditions and events are less controllable, even unpredictable, if compared with laboratory experiments.

Once validated the experiment in terms of differences between the road and traffic conditions, the EEG-based Workload measures have been validated. In particular, the analysis of AUC related to the asSWLDA-based classifier demonstrated that the adopted approach achieves considerable performance, i.e., AUCs > 0.7 (Fawcett, 2006). More in detail, the AUC analysis (**Figure 8**) revealed that it has been possible to achieve mean AUC values of 0.74 for the Normal hour and of 0.73 for the Rush hour, significantly higher than a random classification in both the conditions (respectively p = 0.01 and p = 0.0005). Also, there were no significant differences (p = 0.64) in terms of AUC values on Real data between Normal and Rush hours. All the previous results have been also confirmed by the correction for multiple comparisons, in this case the False Discovery Rate. It is also true that, within the machine-learning theory, AUCs greater than 0.7 are considered remarkable if compared with a random distribution that is assumed to produce AUCs equal to 0.5. In the present study, the performance of the classifier on randomized data achieved AUCs values of about 0.6. A possible explanation could be that the random value would be closer and closer to 0.5 only if the number of repetitions tends to infinite, however, this result undoubtedly encourages research about improving the proposed method. Of course, classification performance of about 0.75 are anyway remarkable, in particular because of the novelty of such application (the EEG-based Workload index is provided with a time resolution equal to 8 s) and the real settings, where mental states assessment is more prone to misclassification: in fact, it is plausible to assume that outside the high controlled laboratory settings, the user experiences more complex mental states that consist of multiple different components having the potential to influence neurophysiological signals used to infer a specific state.

The analysis of the patterns of features selected by the algorithm during its training phase (**Figure 7**) provided interesting insights about its usability: in fact, the asSWLDA selected on average 4 discriminant features for each subject, and even more interesting, by involving 3 of the 11 available channels. It means that, once calibrated the system on a specific user, it would be able to work online during the driving experience involving only three EEG channels, in other words reducing significantly its invasiveness and increasing wearability, two critical aspects for applications outside the laboratory.

At this point, the asSWLDA output, in terms of EEG-based Workload index, has been used to evaluate the effects of road complexity, traffic intensity and external events on drivers' workload (**Figure 9**).

The Friedman ANOVA test (please see Figure) shows the results in terms of effects of the two investigated factors, i.e., the traffic (HOUR) and the road complexity (ROAD), on the mental workload experienced by the drivers: both the traffic and road complexity contributed to significantly increase (main effect: $p = 0.00001$; Wilcoxon tests respectively: HOUR, $p = 0.0032$; ROAD: $p = 0.0038$) the mental workload. In other words, the drivers' workload increased if traffic increased as well independently from the road complexity. At the same time, the drivers' workload increased while driving along more complex roads independently from the traffic intensity. These results have to be considered with respect to the experimental task: actually, the *Hard* segment was a three-lanes main street, that with respect to a one-lane main street (*Easy* segment) implies several additional decisions and actions, such as eventual car overtaking as well as looking at rear-view mirrors because of possible cars coming on lateral lanes. Of course, these actions increase with traffic increasing, because of the higher number of vehicles along the circuit (as demonstrated by Video analysis, please see **Figure 4**). Apparently, the Easy segment should not suffer traffic increasing, since being a one-lane segment the overtaking are very limited and drivers have not to frequently check rear-view mirrors since they cannot change lane. Nevertheless, because of the higher number of vehicles along the circuit during rush hours, the drivers had to continuously monitor eventual preceding cars, adapting safety distance and speed (in fact average speed during rush hour has been lower and drivers' gazes on infrastructure and vehicles higher also along Easy segment). These actions also induced a no-negligible workload increasing, giving a possible justification of the high accident rate along rural roads (Shankar et al., 1995), that are generally considered "*Easy to drive*" if compared with urban main roads (Harms, 1991; Paxion et al., 2014), thus mismatching the driver's expectations.

Very interestingly but not surprisingly, the neurophysiological measures showed a significantly higher sensitivity with respect to the ET ones (**Figures 10**, **11**) in discriminating the different impact of road complexity and traffic intensity on mental workload. It is important to consider that ET measures were available only for a reduced group of the experimental sample (8 of 16 subjects), therefore it could have affected the performance of such measures in discriminating the mental workload related to different factors. However, the paired statistical analysis highlighted that on the same subjects, EEG-based measures were more sensitive to workload fluctuations. Their high sensitivity has been pointed out also with respect to subjective measures (i.e., NASA-TLX questionnaires, **Figure 12**), that on the contrary were not able to discriminate ($p = 0.23$) normal from rush hours.

Finally, EEG-based workload measures revealed a significant workload increasing ($p < 0.05$) related to both the investigated events, that is the car and the pedestrian crossing the road, especially in normal hours independently from the road complexity. Instead, no significant workload increasing were associated to the event during rush hours, despite a significantly higher workload in comparison with the same events during normal hours (**Figure 13**). Although for this analysis neurophysiological measures are not integrated with additional ones (it was impossible to collect subjective data related to specific events, while from the ET point of view it was possible to assess only if the event was been perceived or not), it is possible to deduce that external events could lead to eventually risky situations especially with low traffic (normal hours). In fact, although a lower absolute workload if compared with high traffic

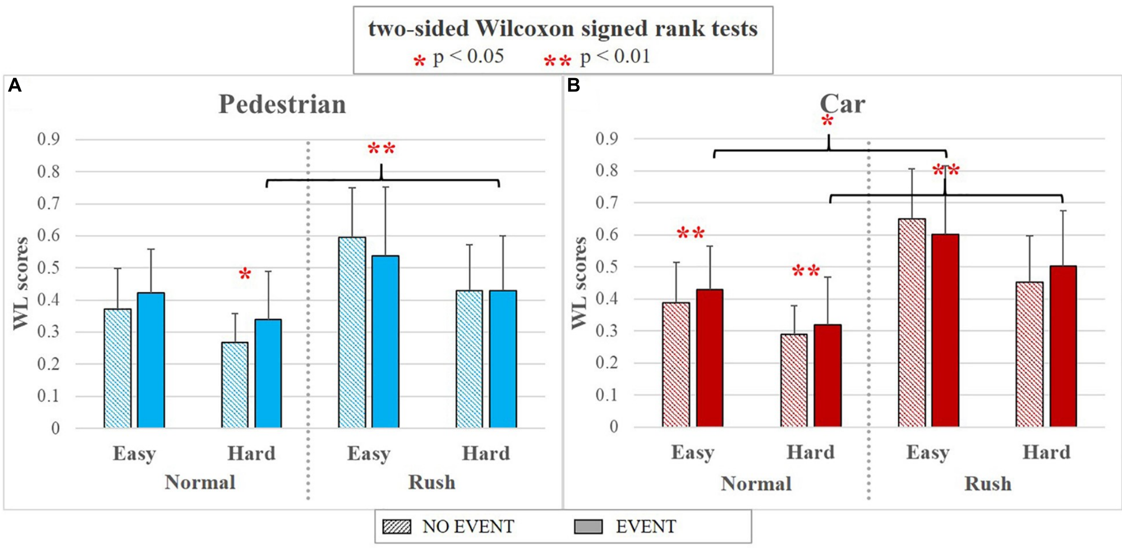

FIGURE 13 | The bar graphs show the mean values and the standard deviation of the EEG-based WL scores related to the different events along the various experimental conditions. In particular, the results are divided per events category, i.e., Pedestrian on the left **(A)** and Car on the right **(B)**. In both the cases, the condition EVENT (solid color) is referred to the event actually happened during the 3rd lap, the condition NO EVENT (lines pattern) is referred to the same circuit portion during the 2nd lap when no events were acted. The Wilcoxon tests revealed a significant workload increasing (one red asterisk stands for $p < 0.05$; two red asterisks stand for $p < 0.01$) related to both the investigated events, that is the car and the pedestrian crossing the road, especially in normal hours independently from the road complexity. Instead, no significant workload increasing were associated to the event during rush hours, despite a significantly higher workload in comparison with the same events during normal hours.

condition, they are characterized by an immediate cognitive demand increase, that could become dangerous if not expected by the driver.

Nevertheless, the main limit that affects the present study is the algorithm calibration with data coming from the task itself and recorded in very similar conditions. From one side, it could be argued that in everyday life context such a calibration would be unfeasible; from the other side it could be argued that the proposed algorithm is not classifying the targeted mental state, i.e., mental workload, but only two conditions that are very similar. Regarding the calibration, actually it is one of the main still open issues in transferring machine learning approaches from research to applied field: several solutions have been explored, such as cross-task calibration or employment of unsupervised algorithms, but the problem is still open and needs further investigation (Aricò et al., 2018). However, the present work did not aim at addressing such issue, but at investigating the possibility of applying a machine-learning algorithm for the mental workload evaluation, already validated in other domains, also in automotive applications. The highly challenging conditions of a "real driving experiment" with twenty subjects, jointly with the employment of high-quality instrumentation, already make the present work very innovative and of interest. Secondly, it is true that the algorithm has been calibrated on two conditions and employed in classifying two similar conditions, but it is also important to consider that calibration data for each subject came from two different repetitions (please refer to Section "Electroencephalographic Signal Recording and Processing" for more information): in fact data recorded during the *Easy* segment of *Normal* hour (2nd lap) have been used as EASY CLASS, while data recorded during the *Hard* segment of *Rush* hour (2nd lap) have been used as HARD CLASS. Even if assuming that *Easy* segment of *Normal* hour and *Hard* segment of *Rush* hour of 2nd and 3rd lap were intrinsically similar, no data from *Hard* segment of *Normal* hour and *Easy* segment of *Rush* hour have been used to train the classifier, therefore their coherent classification (e.g., *Hard* segment of *Normal* hour is not easier than the *Easy* segment during the same hour) is a mere and appreciable result of the proposed algorithm. Undoubtedly, mental workload is a Human Factor concept hard to define and even worse to measure (Moray, 2013), and confounds arising from different mental states are probably present, however, the results of the present study are already remarkable, especially if considering previous results obtained by the employment of the same algorithm in different applications (Aricò et al., 2016a,b, Borghini et al., 2017b,c).

It is important to remark how it is possible to achieve this kind of results only thanks to the proposed methodology: in fact, subjective measures cannot be gathered with high time resolution and without interfering with the main task, briefing and debriefing sessions can be performed only before and after the experience, while eye-tracker as well as other neurophysiological metrics (for example the *ThetaF/AlphaP* showed in **Figure 6**) are able to provide only an overall evaluation about a "long" condition. On the contrary, the proposed methodology is able to overcome these limitations, providing workload assessment with high time resolution (i.e., in this case 8 s) and thus allowing to evaluate also specific events.

In conclusion, the obtained results appear very interesting in terms of understanding driver's behaviors and its relationship with road environment, highlighting the added value of neurophysiological measures in providing insights about human mind that are not obtainable, or at least difficult to obtain, with traditional approaches. Certainly, further analyses are necessary in order to validate this multimodal approach with a larger sample of subjects, exploring the impact of other factors, such as different events, road signage and so on, and involving additional tools typical of road safety research, as well as exploring the possibility of calibrating the proposed algorithm without any task-related data.

CONCLUSION

The present study, through a real driving experiment, aimed to validate a methodology able to infer driver's mental workload on the basis of his/her brain activity through Electroencephalographic technique. Once validated, such methodology has been successfully employed to evaluate the impact of different factors, specifically the road complexity, the traffic intensity (depending on the hour of the day), and two specific events (a pedestrian crossing the road and a car entering in the traffic flow), on the drivers' experienced mental workload. The analyses have been supported by information coming from subjective measures, drivers' eye movements tracking and car parameters. The results demonstrated (i) the reliability and effectiveness of the proposed methodology based on human EEG signals to objectively measure driver's mental workload with respect to different road factors, and (ii) the added value of neurophysiological measures in providing insights about human mind while dealing with tasks that are difficult or even impossible to obtain by using traditional approaches. In conclusion, other than the specific obtained results, the present work breaks new ground for the integration of these new methodologies, i.e., neurophysiological measures, with traditional approaches in order to enhance and extend research on drivers' behaviors and road safety.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Good Clinical Practice (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the 'University of Bologna.'

AUTHOR CONTRIBUTIONS

GDF is the main author of the paper. Also, he was actively involved within the experiments as well as in the EEG data analysis and interpretation. PA, GB, and NS supported the experimental design, the data recording, the EEG data analysis, and the manuscript writing. PL and SP supported the experimental design and the results interpretation, providing their contribute in particular about the Human Factor concepts. VV, CL, AB, and AS were in charge of experiments planning, they contributed actively to the experiments execution, they analyzed Eye Tracking and subjective measures analysis, and supported the results interpretation. FB coordinated the research group, from the experimental design to the manuscript editing.

ACKNOWLEDGMENTS

The authors are grateful to the Unipol Group Spa and, in particular, ALFAEVOLUTION TECHNOLOGY, for the considerable help given in the research study. This work has been also co-financed by the European Commission through the Horizon2020 project H2020-MG-2016 *Simulator of behavioral aspects for safer transport*, "SimuSafe," (GA n. 723386), and the project "BrainSafeDrive: A Technology to Detect Mental States During Drive for Improving the Safety of the Road" (Italy-Sweden collaboration) with a grant of Ministero dell'Istruzione dell'Università e della Ricerca della Repubblica Italiana.

REFERENCES

Allnutt, M. F. (1987). Human factors in accidents. *Br. J. Anaesth.* 59, 856–864. doi: 10.1093/bja/59.7.856

Aricò, P., Borghini, G., Di Flumeri, G., and Babiloni, F. (2017a). *Method for Estimating a Mental State, in Particular a Workload, and Related Apparatus. EP3143933 A1, filed September 14, 2016, and issued March 22, 2017*. Available at: http://www.google.it/patents/EP3143933A1

Aricò, P., Borghini, G., Di Flumeri, G., Bonelli, S., Golfetti, A., Graziani, I., et al. (2017b). Human factors and neurophysiological metrics in air traffic control: a critical review. *IEEE Rev. Biomed. Eng.* 10, 250–263. doi: 10.1109/RBME.2017.2694142

Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., Colosimo, A., and Babiloni, F. (2017c). Passive BCI in operational environments: insights, recent advances, and future trends. *IEEE Trans. Biomed. Eng.* 64, 1431–1436. doi: 10.1109/TBME.2017.2694856

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Bonelli, S., Golfetti, A., et al. (2016a). Adaptive automation triggered by eeg-based mental workload index: a passive brain-computer interface application in realistic air traffic control environment. *Front. Hum. Neurosci.* 10:539. doi: 10.3389/fnhum.2016.00539

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Pozzi, S., and Babiloni, F. (2016b). A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks. *Progr. Brain Res.* 228, 295–328. doi: 10.1016/bs.pbr.2016.04.021

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Graziani, I., Imbert, J. P., et al. (2015). "Reliability over time of EEG-based mental workload evaluation during Air Traffic Management (ATM) tasks," in *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (New York, NY: IEEE), 7242–7245. doi: 10.1109/EMBC.2015.7320063

Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., and Babiloni, F. (2018). Passive BCI beyond the lab: current trends and future directions. *Physiol. Meas.* 39:08TR02. doi: 10.1088/1361-6579/aad57e

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J. Mathemat. Psychol.* 12, 387–415. doi: 10.1016/0022-2496(75)90001-2

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Borghini, G., Aricò, P., Di Flumeri, G., and Babiloni, F. (2017a). *Industrial Neuroscience in Aviation: Evaluation of Mental States in Aviation Personnel.* Cham: Springer International Publishing Springer. doi: 10.1007/978-3-319-58598-7

Borghini, G., Aricò, P., Di Flumeri, G., Cartocci, G., Colosimo, A., Bonelli, S., et al. (2017b). EEG-based cognitive control behaviour assessment: an ecological study with professional air traffic controllers. *Sci. Rep.* 7:547. doi: 10.1038/s41598-017-00633-7

Borghini, G., Aricò, P., Di Flumeri, G., Sciaraffa, N., Colosimo, A., Herrero, M., et al. (2017c). A new perspective for the training assessment: machine learning-based neurometric for augmented user's evaluation. *Front. Neurosci.* 11:325. doi: 10.3389/fnins.2017.00325

Borghini, G., Aricò, P., Di Flumeri, G., Salinari, S., Colosimo, A., Bonelli, S., and Babiloni, F. (2015). "Avionic technology testing by using a cognitive neurometric index: a study with professional helicopter pilots," in *Proceedings of the Engineering in Medicine et al. (EMBC), 2015 37th Annual International Conference of the IEEE* (New York, NY: IEEE), 6182–6185. doi: 10.1109/EMBC.2015.7319804

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* 44, 58–75. doi: 10.1016/j.neubiorev.2012.10.003

Borghini, G., Vecchiato, G., Toppi, J., Astolfi, L., Maglione, A., Isabella, R., et al. (2012). "Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices," in *Proceedins of the Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA. 6442–6445. doi: 10.1109/EMBC.2012.6347469

Brookhuis, K. A., and de Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accid. Anal. Prevent.* 42, 898–903. doi: 10.1016/j.aap.2009.06.001

Bucchi, A., Sangiorgi, C., and Vignali, V. (2012). Traffic psychology and driver behavior. *Proc. Soc. Behav. Sci.* 53, 972–979. doi: 10.1016/j.sbspro.2012.09.946

Cestac, J., Paran, F., and Delhomme, P. (2014). Drive as i say, not as i drive: influence of injunctive and descriptive norms on speeding intentions among young drivers. *Trans. Res. Part F Traffic Psychol. Behav.* 23, 44–56. doi: 10.1016/j.trf.2013.12.006

Costa, M., Simone, A., Vignali, V., Lantieri, C., Bucchi, A., and Dondi, G. (2014). Looking behavior for vertical road signs. *Transport. Res. Part F Traffic Psychol. Behav.* 23, 147–155. doi: 10.1016/j.trf.2014.01.003

De Waard, D. (1996). *The Measurement of Drivers' Mental Workload*. Groningen: University of Groningen, Traffic Research Centre.

de Winter, J. C. F., Happee, R., Martens, M. H., and Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. *Transport. Res. Part F Traffic Psychol. Behav.* 27, 196–217. doi: 10.1016/j.trf.2014.06.016

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Di Flumeri, G., Aricò, P., Borghini, G., Colosimo, A., and Babiloni, F. (2016). "A new regression-based method for the eye blinks artifacts correction in the eeg signal, without using any EOG channel," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Orlando, FL.

Di Flumeri, G., Aricò, P., Borghini, G., Sciaraffa, N., Vignali, V., Lantieri, C., et al. (2018). "Mental workload assessment as taxonomic tool for neuroergonomics," in *Neuroergonomics. The Brain at Work and in Everyday Life*, 1st Edn, eds H. Ayaz and F. Dehais (Amsterdam: Elsevier).

Di Flumeri, G., Borghini, G., Aricò, P., Colosimo, A., Pozzi, S., Bonelli, S., et al. (2015). "On the use of cognitive neurometric indexes in aeronautic and air traffic management environments," in *Proceedings of the Symbiotic Interaction: 4th International Workshop, Symbiotic 2015*, Berlin, Germany. doi: 10.1007/978-3-319-24917-9_5

Dienes, Z. (2004). Assumptions of subjective measures of unconscious mental states: higher order thoughts and bias. *J. Consciousn. Stud.* 11, 25–45.

Elul, R. (1969). Gaussian behavior of the electroencephalogram: changes during performance of mental task. *Science* 164, 328–331. doi: 10.1126/science.164.3877.328

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recogn. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010

Gevins, A., and Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theor. Issues Ergon. Sci.* 4, 113–131. doi: 10.1080/14639220210159717

Göhring, D., Latotzky, D., Wang, M., and Rojas, R. (2013). "Semi-autonomous car control using brain computer interfaces," in *Intelligent Autonomous Systems 12*, vol. 194, eds S. Lee, H. Cho, K.-J. Yoon, and J. Lee (Berlin: Springer), 393–408. doi: 10.1007/978-3-642-33932-5_37

Gopher, D., and Braune, R. (1984). On the psychophysics of workload: why bother with subjective measures? *Hum. Fact. J. Hum. Fact. Ergon. Soc.* 26, 519–532. doi: 10.1177/001872088402600504

Hansen, F. (2007). Human error: a concept analysis. *J. Air Transport.* 11, 61–77.

Harms, L. (1991). Variation in drivers' cognitive load. Effects of driving through village areas and rural junctions. *Ergonomics* 34, 151–160. doi: 10.1080/00140139108967303

Hart, S. G., and Staveland, L. E. (1988). "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research," in *Advances in Psychology*, Vol. 52, eds P. A. Hancock and N. Meshkati (North-Holland: Human Mental Workload), 139–183. doi: 10.1016/S0166-4115(08)62386-9

Holm, A., Lukander, K., Korpela, J., Sallinen, M., and Müller, K. M. I. (2009). Estimating brain load from the EEG. *Sci. World J.* 9, 639–651. doi: 10.1100/tsw.2009.83

Horowitz, A. D., and Dingus, T. A. (1992). Warning signal design: a key human factors issue in an in-vehicle front-to-rear-end collision warning system. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* 36, 1011–1013. doi: 10.1177/154193129203601320

Kaplan, S., Guvensan, M. A., Yavuz, A. G., and Karalurt, Y. (2015). Driver behavior analysis for safe driving: a survey. *IEEE Trans. Intell. Transport. Syst.* 16, 3017–3032. doi: 10.1109/TITS.2015.2462084

Kirk, R. E. (2015). "Experimental design," in *The Blackwell Encyclopedia of Sociology*, ed. G. Ritzer (Oxford: John Wiley & Sons, Ltd.), doi: 10.1002/9781405165518.wbeose087.pub2

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3

Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K. L-R., Curio, G., et al. (2007). "Improving human performance in a real operating environment through real-time mental workload detection," in *Towards Brain-Computer Interfacing*, eds G. Dornhege, J. del R. Millá, T. Hinterberger, D. McFarland, and K.-Robert Mülle (Cambridge, MA: MIT Press).

Kong, W., Lin, W., Babiloni, F., Hu, S., and Borghini, G. (2015). Investigating driver fatigue versus alertness using the granger causality network. *Sensors* 15, 19181–19198. doi: 10.3390/s150819181

Kong, W., Zhou, Z., Jiang, B., Babiloni, F., and Borghini, G. (2017). Assessment of driving fatigue based on intra/inter-region phase synchronization. *Neurocomputing* 219, 474–482. doi: 10.1016/j.neucom.2016.09.057

Lantieri, C., Lamperti, R., Simone, A., Costa, M., Vignali, V., Sangiorgi, C., et al. (2015). Gateway design assessment in the transition from high to low speed areas. *Transport. Res. Part F Traffic Psychol. Behav.* 34, 41–53. doi: 10.1016/j.trf.2015.07.017

Lees, M. N., Cosman, J. D., Lee, J. D., Fricke, N., and Rizzo, M. (2010). Translating cognitive neuroscience to the Driver's operational environment: a neuroergonomic approach. *Am. J. Psychol.* 123, 391–411. doi: 10.5406/amerjpsyc.123.4.0391

Lin, C.-T., Ko, L.-W., and Shen, T.-K. (2009). Computational intelligent brain computer interaction and its applications on driving cognition. *IEEE Computat. Intell. Mag.* 4, 32–46. doi: 10.1109/MCI.2009.934559

Lin, C.-T., Wu, R.-C., Jung, T.-P., Liang, S.-F., and Huang, T.-Y. (2005). Estimating driving performance based on EEG spectrum analysis. *EURASIP J. Adv. Sig. Process.* 2005:521368. doi: 10.1155/ASP.2005.3165

Maglione, A., Borghini, G., Aricò, P., Borgia, F., Graziani, I., Colosimo, A., et al. (2014). "Evaluation of the workload and drowsiness during car driving by using high resolution EEG activity and neurophysiologic indices," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, 6238–6241. doi: 10.1109/EMBC.2014.6945054

Michail, E., Kokonozi, A., Chouvarda, I., and Maglaveras, N. (2008). *EEG and HRV Markers Of Sleepiness and Loss of Control During Car Driving*. New York, NY: IEEE. doi: 10.1109/IEMBS.2008.4649724

Moray, N. (ed). (2013). *Mental Workload: Its Theory and Measurement*. Berlin: Springer Science & Business Media, 8.

Parasuraman, R., and Rizzo, M. (2008). *Neuroergonomics: The Brain at Work*, 1 Edn. New York, NY: Oxford University Press.

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. *J. Cognit. Eng. Decis. Mak.* 2, 140–160. doi: 10.1518/155534308X284417

Paxion, J., Galy, E., and Berthelon, C. (2014). Mental workload and driving. *Front. Psychol.* 5:1344. doi: 10.3389/fpsyg.2014.01344

Petridou, E., and Moustaki, M. (2000). Human factors in the causation of road traffic crashes. *Eur. J. Epidemiol.* 16, 819–826. doi: 10.1023/A:1007649804201

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Åkerstedt, T., et al. (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep* 28, 1511–1516. doi: 10.1093/sleep/28.12.1511

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., and Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *Int. J. Aviat. Psychol.* 10, 393–410. doi: 10.1207/S15327108IJAP1004_6

Reason, J. (2000). Human error: models and management. *BMJ? Br. Med. J.* 320, 768–770. doi: 10.1136/bmj.320.7237.768

Robert, J. H. G. (1997). Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical framework. *Biol. Psychol.* 45, 73–93. doi: 10.1016/S0301-0511(96)05223-4

Shankar, V., Mannering, F., and Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Anal. Prevent.* 27, 371–389. doi: 10.1016/0001-4575(94)00078-Z

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences. Nonparametric Statistics for the Behavioral Sciences*. New York, NY: McGraw-Hill.

Subramanian, R. (2012). *Motor Vehicle Traffic Crashes as a Leading Cause Of Death in the United States, 2008 and 2009. HS-811 620. Traffic Safety Facts – Research Note. National Highway Traffic Safety Administration*. Available at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811620

Summala, H., and Mikkola, T. (1994). Fatal accidents among car and truck drivers: effects of fatigue, age, and alcohol consumption. *Hum. Fact.* 36, 315–326. doi: 10.1177/001872089403600211

Toppi, J., Borghini, G., Petti, M., He, E. J., De Giusti, V., He, B., et al. (2016). Investigating cooperative behavior in ecological settings: an EEG hyperscanning study. *PLoS One* 11:e0154236. doi: 10.1371/journal.pone.0154236

Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., and Hume, R. D. (1979). *Tri-level Study of the Caused of Traffic Accidents. Executive Summary*. Available at: https://trid.trb.org/view.aspx?id=144095

Vecchiato, G., Borghini, G., Aricò, P., Graziani, I., Maglione, A. G., Cherubino, P., et al. (2016). Investigation of the effect of EEG-BCI on the simultaneous execution of flight simulation and attentional tasks. *Med. Biol. Eng. Comput.* 54, 1503–1513. doi: 10.1007/s11517-015-1420-6

Verwey, W. B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics* 43, 187–209. doi: 10.1080/001401300184558

Vicente, K. J. (2013). *The Human Factor: Revolutionizing the Way People Live with Technology*. Hoboken: Taylor and Francis. Available at: https://www.taylorfrancis.com/books/9781135877262

Wall, T. D., Michie, J., Patterson, M., Wood, S. J., Sheehan, M., Clegg, C. W., et al. (2004). On the validity of subjective measures of company performance. *Pers. Psychol.* 57, 95–118. doi: 10.1111/j.1744-6570.2004.tb02485.x

Wang, H., Zhang, C., Shi, T., Wang, F., and Ma, S. (2015). Real-time EEG-based detection of fatigue driving danger for accident prediction. *Int. J. Neural Syst.* 25:1550002. doi: 10.1142/S0129065715500021

World Health, and Organization. (2015). *Global Status Report on Road Safety 2015: Supporting a Decade of Action.* Geneva: WHO.

World Health Organization (2013). *Global Status Report on Road Safety 2013: Supporting a Decade of Action.* Available at: http://site.ebrary.com/id/10931312

Xing, Y., Lv, C., Cao, D., Wang, H., and Zhao, Y. (2018). Driver workload estimation using a novel hybrid method of error reduction ratio causality and support vector machine. *Measurement* 114, 390–397. doi: 10.1016/j.measurement.2017.10.002

Zhang, H., Chavarriaga, R., Khaliliardali, Z., Gheorghe, L., Iturrate, I., and Millán, J. D. R. (2015). EEG-based decoding of error-related brain activity in a real-world driving task. *J. Neural Eng.* 12:066028. doi: 10.1088/1741-2560/12/6/066028

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 30, 2019, accepted December 19, 2019, date of publication January 13, 2020, date of current version January 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2965940

Analysis of Road-User Interaction by Extraction of Driver Behavior Features Using Deep Learning

ARIANNA BICHICCHI [1], RACHID BELAROUSSI [2], ANDREA SIMONE [1], VALERIA VIGNALI [1], CLAUDIO LANTIERI [1], AND XUANPENG LI [3]

[1]Department of Civil, Chemical, Environmental and Materials Engineering, School of Engineering and Architecture, University of Bologna, 40136 Bologna, Italy
[2]COSYS-GRETTIA, French Institute of Science and Technology for Transport, Development and Networks, Gustave Eiffel University, 77447 Marne-la-Vallée, France
[3]School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

Correspondin g author: Arianna Bichicchi (arianna.bichicchi2@unibo.it)

ABSTRACT In this study, an improved deep learning model is proposed to explore the complex interactions between the road environment and driver's behaviour throughout the generation of a graphical representation. The proposed model consists of an unsupervised Denoising Stacked Autoencoder (SDAE) able to provide output layers in RGB colors. The dataset comes from an experimental driving test where kinematic measures were tracked with an in-vehicle GPS device. The graphical outcomes reveal the method ability to efficiently detect patterns of simple driving behaviors, as well as the road environment complexity and some events encountered along the path.

INDEX TERMS Deep learning, driver behavior, event detection, road safety, workload.

I. INTRODUCTION

Road safety is today one of the most actual and challenging field of research, as road fatalities continue to increase year after year with dramatic social and economic impacts [1]. Because of the factors associated to fatal road accidents, most studies are addressed to human factors who aspire to analyze the driver behaviour. The entire process of observation, modelling, visualization and prediction of driving behaviour unavoidably presuppose the development of experimental tests who produce large amount of data. Both simulated and semi- or naturalistic tests are provided with different types of sensors aiming to record all possible information from the driver and from the vehicle, together with their interactions.

As deep learning methods may help in the processing of high-dimensional data, their application in transportation field has been increasing recently. Applications are mainly related to traffic flow forecasting, crashes prediction and driver behaviour analyses.

Studies focusing on driver behavior analysis presuppose that the design and selection of the features is based on researchers experience and finding an appropriate method for their representation is often difficult, especially for driving behaviors that are obtained from a driver–vehicle–environment system [2].

Most of the studies aiming to extract latent features from multi-dimensional time-series data were performed by using the Principal Component Analysis. PCA is used to decompose a multivariate dataset in a set of successive orthogonal components that explain a maximum amount of the variance [3]. Despite the several conducted studies, it is difficult to adopt PCA for extracting time series of latent features from driving behaviour data because vehicle dynamics and the driver behaviour have non-linear properties, whereas PCA is based on linear transformations. This problem can be solved by using Kernel Principal Component Analysis (KPCA), because it uses a non-linear kernel function that involves a non-linear transformation for mapping the data to a high-dimensional space. Then, KPCA employs PCA to extract latent features in the high-dimensional space. Indeed, Zhao successfully extracted latent features for driver mental fatigue classification using KPCA, proving that this method is more accurate than PCA [4]. Nevertheless, when a large amount of driving behaviour data is used for analysis, the computational cost of KPCA is high because the kernel method has to compute a matrix in RN×N, where N is the total number of frames of data.

Dong made the first attempt by adopting a deep neural architecture, based on Convolutional Neural Network (CNN)

The associate editor coordinating the review of this manuscript and approving it for publication was Haiyong Zheng.

and Recurrent Neural Network (RNN), for the first time in order to extract features directly from GPS. This because those neural networks can learn high level driving style features from the low level feature matrices requiring less human work than the previous methods, that rely on handcrafted driving behaviour features [5].

A particular case of driver behaviour analysis is the one by Dwivedi who proposed a vision based on CNN to detect driver drowsiness. According to this, the caption of various latent facial features and complex non-linear feature interactions were possible [6].

Other studies have proved that an algorithm based on the Back Propagation (BP) provides good results because it can approximate any non-linear continuous function with arbitrary precision. Meseguer used this kind of neural network to dynamically and automatically analyze users data in order to identify the driving style as well as the category of road segment profile [7]. This type of neural network results are effective not only for vehicle dynamics and geographic data, but also for training and testing features in order to improve the road type recognition rate based on images [8].

The development of several applications including tasks based on unsupervised methods for driver behaviour detection have started only a few years ago [9]. In case of large dataset generated by non-linear transformations, unsupervised learning methods are able to extract latent features of driving behavior without using label information. In particular, recent studies involved this method to enable reliable driving behaviour visualization output. The review of driving behaviour can be a key practice for the improvement of driving behaviour and safe driving promotion [10].

Considered the importance of using a denoising criterion as a tractable unsupervised objective to guide the learning of useful higher-level representations [11], in this study has been chosen to exploit a Denoising Stacked-Autoencoder (SDAE) to extract the latent features for a deep driving behavior analysis.

It is worth noting the importance of such research, which also aims at being able to provide real time information and consequently safety advices for the prevention of road crashes. Indeed, it is supposed that deep learning will be able to accurately predict driver behaviour patterns, attracting relevant attention for the potential role in autonomous driving applications.

II. METHODS

The objective of this study is to use SDAE for extracting the driving behaviour features from a dataset from a real driving test and recorded by an in-vehicle GPS sensor. Regarding data source, current research suggests that in-vehicle data (CAN-BUS) can be used as an effective representation of driving behaviour for recognizing different drivers [12]. Similarly, other studies involve GPS receiver [13] for simple data extraction and potential usability in large scale research.

The proposed method implements a deep sparse autoencoder (SDAE) to extract the lowdimensional high-level representation from high-dimensional raw driving behavioral dataset. According to the resulting low-dimensional representation, two visualization methods are suggested. The first is a cubic representation displaying extracted three-dimensional features. The second is a colored trajectory showing on the path driven the color expression of the extracted features. The color results from an RGB color space combination corresponding to the extracted three-dimensional features.

FIGURE 1. VBOX GPS/camera data logger.

A. EXPERIMENTAL SETUP

The data collection used the Racelogic Video V-Box Pro device (FIGURE 1), an in-car video system installed on the test vehicle. The reliability of this device for data analysis have been tested several road safety studies [14], [15]. The device combines a 10Hz GPS data logger with a two cameras video system, with an accuracy of 0.5 meters and 0.2 km/h. The output consisted in a .csv file with a data recording period of 0,1s. Every recording included information on positioning coordinates (latitude and longitude), time and several kinematic data. With reference to this study, the six typologies of kinematic data considered are:

- Longitudinal speed (km/h);
- Vertical speed (km/h);
- Longitudinal Acceleration (m/s^2);
- Transversal Acceleration (m/s^2);
- ComboG (combination of g forces);
- Heading of the vehicle (deg);

For the scope of this research, a real driving test has been conducted and driving behavior data of 10 participants have been collected.

The driving tests were run within the industrial zone of Casalecchio di Reno (Bologna – Italy) on a circuit route of 2500 meters. The first 1000 meters, red in FIGURE 2 consisted of a complex road stretch with many entrances and traffic while the last 1500 meters, blue in FIGURE 2, until the starting point consisted of a road stretch with a simpler driving complexity [16], [17].

All participants were asked to drive for two testing sessions (approximatively one in the morning and one in the afternoon). As summarized in TABLE 1, each driving session included three laps of the circuit route: the first dedicated to the user' adaptation to the experimental conditions, the second to test the driving performance without any external event occurring (''baseline lap'') and the third to test the driving

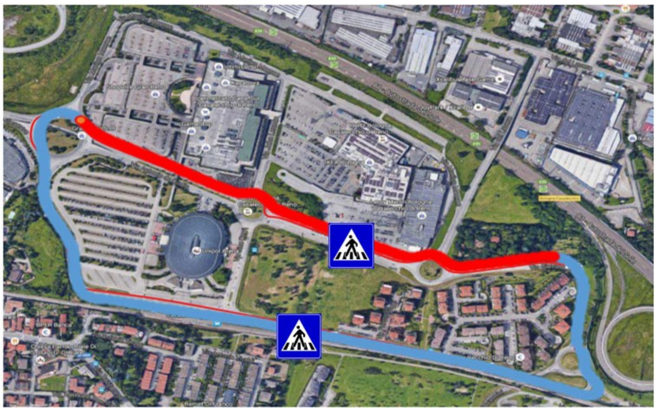

FIGURE 2. Driving route.

TABLE 1. Driving test features and variables.

LAP SCHEDULE	
Lap 0	Adaptation to the experimental settings (dataset not used)
Lap 1	Baseline
Lap 2	Test with external events (2 pedestrians crossing the road on crosswalks)
ROAD COMPLEXITY	
Complex (red color)	Road section with three lanes for direction, many entrances and traffic jams, commercial area
Easy (blue color)	Road section with one lane for direction, residential area

performance in reaction to two workload inducing simulated events (''test lap'') consisting of a pedestrian crossing the road on a crosswalk (one on the ''complex'' segment and the other on the ''easy'' segment''). At the end of each driving session, participants were asked to subjectively evaluate their mental load (workload) while accomplishing the driving task throughout the standardized NASA-TLX survey.

One dataset sheet was elaborated for each test participant, accordingly to the scheduled data processing activities.

B. MODEL AND HYPERPARAMETERS TUNING

The developed DSAE model is able to extract time series of latent features through an encoding process, one for each hidden layer, minimizing the error computed with the cost function between the input time-series data and the decoded time-series data. According to the visualization method here proposed, the extracted features are drawn on a roadmap representing a colored trajectory.

In the following subparagraphs the developed model is described together with hyperparameters tuning activity. The goal of hyperparameters tuning is to select hyperparameters that will give good generalization performance. Typically, this works by estimating the generalization performance for different choices of hyperparameters (e.g. using a validation set), and then choosing the best.

Théano library was employed for model development of this study.

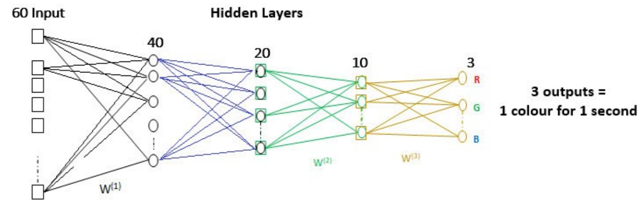

FIGURE 3. Architecture of the SDAE.

a: ACTIVATION FUNCTION

The chosen activation function consists in a hyperbolic tangent function $f(\cdot) = \tanh(\cdot)$ as has been evaluated that outperforms the traditional sigmoid function.

b: ARCHITECTURE

The chosen architecture consisted in the scheme in FIGURE 3:

- Input: 6 data inputs x 10 data measurements over 1s = 60 inputs for a sliding window.
- Encoding hidden layers (40, 20, 10);
- Output layer: 3 RGB colors. The RGB color space is ideal to represent driving behaviors, being a three-dimensional space. As the range of the RGB color space is [0, 1], the three-dimensional hidden features have been normalized into [0, 1]. In summary, the three-dimensional hidden features could be mapped to the RGB space by:

$$rgb_{t,d} = \frac{h_{t,d}^{(final)} - h_{min_d}^{(final)}}{h_{max_d}^{(final)} - h_{min_d}^{(final)}} \tag{1}$$

where:

- $rgb_{t,d}$ is a d-th element of a three-dimensional vector in the RGB space that represents the driving behaviour at the t-th time step.
- $\boldsymbol{h}_{t,d}^{(final)}$ is the d-th element of the extracted three-dimensional hidden feature's vector at the t-th time step;
- $\boldsymbol{h}_{\mathrm{max,d}}^{(\mathrm{final})}$ and $\boldsymbol{h}_{\mathrm{min,d}}^{(\mathrm{final})}$ are the minimum and maximum values of the d-th dimension in $\boldsymbol{h}$*(final),* respectively.

c: NORMALIZATION

The measured driving behavior input data are defined as Y ∈ RDY×NY, where DY is the dimensionality and NY is the quantity of data (frames) in Y. Each dimension of Y represents one type of feature time-series data. The t-th frame of Y is defined as:

$$\boldsymbol{y_t} = (y_{t,1}, y_{t,2}, \ldots, y_{t,DY})^T \in RDY \tag{2}$$

Considering the use of a hyperbolic tangent as activation function, the output range of the normalization process results is [-1,1].

To reconstruct the input data using the tanh function, each dimension of Y is independently normalized into [-1,1] by using the maximum and minimum values. Thus, the t-th frame of the normalized data is expressed as:

$$\boldsymbol{x_t} = (x_{t,1}, x_{t,2}, \ldots, x_{t,DY})^T \in \mathrm{R}DY \tag{3}$$

where it is normalized by:

$$x_{t,d} = 2\left(\frac{y_{t,d} - y_{d_{min}}}{y_{d_{max}} - y_{d_{min}}}\right) - 1 \tag{4}$$

$$y_{d_{max}} = \max\left(y_{1,d}, \ldots, y_{NY,d}\right) \tag{5}$$

$$y_{d_{min}} = \min\left(y_{1,d}, \ldots, y_{NY,d}\right) \tag{6}$$

where y_{dmax} and y_{dmin} are the maximum and minimum values of the d-th dimension of Y, respectively.

d: WINDOWING

In the windowing process, the normalized data are aggregated with a slide window that converts the data of w frames into a vector.

Thus, the windowing time-series data in the t-th frame are expressed as:

$$\boldsymbol{h}_t^{(1)} = \left(\boldsymbol{x}_{t-w+1}^{\boldsymbol{T}}, \boldsymbol{x}_{t-w+2}^{\boldsymbol{T}}, \ldots, \boldsymbol{x}_{DV}^{\boldsymbol{T}}\right) \boldsymbol{T} \in \boldsymbol{R}^{D\mathbf{H}} \tag{7}$$

and $\boldsymbol{DH} = w \times \boldsymbol{DY}$, $t \geq w$.

Finally, the obtained windowing time-series data are:

$$\boldsymbol{H}^{(1)} = \left\{\boldsymbol{h}_1^{(i)}, \boldsymbol{h}_2^{(i)}, \ldots, \boldsymbol{h}_{NH}^{(i)}\right\} \in \mathrm{R}^{D\mathrm{H}xN\mathrm{H}} \tag{8}$$

when the slide window moves along the time axis frame by frame.

Hence, $N\mathrm{H} = N\mathrm{Y}\text{-}w\text{+}1$ frames of windowing time-series data are obtained.

e: REGULARIZATION AND GENERATION OF A DRIVING COLOR MAP

Since loss functions are a key part of any machine learning model, we define an objective against which the performance of the model is measured. The set of weight parameters learned by the model is determined by minimizing a chosen loss function. In this research, the chosen cost function to train layer (l) is the following:

$$O^{(l)}(\Sigma) = \frac{1}{2N_V}\sum_{t=1}^{N_V}\left\|\mathbf{W}_t^{l(T)}\mathbf{h}_t^{(l)} - \mathbf{h}_t^{(l)}\right\|^2 + \frac{\alpha}{2}\sum_{l=1}^{L-1}\left\|\mathbf{W}^{(l)}\right\|_2^2 + \beta\sum_{i=1}^{D_H^{(l)}}\mathrm{KL}(\omega||\bar{h}_i^{(m)}) \tag{9}$$

where:

- L is the number of layers;
- $\mathbf{h}_t^{(l)}$ is the activity of hidden layer (l);
- $\sum = \{\mathbf{W}\ (1), \cdots, \mathbf{W}\ (L\text{-}1), \mathbf{b}\ (1), \cdots, \mathbf{b}\ (L\text{-}1)\ \}$ and the three terms represent respectively the Reconstruction Error Term and two Regularization Terms;
- the second term is used as a penalty to prevent overfitting, limiting the elements of all the weights W(l) by the L2 norm. In addition, the parameter α controls the strength of the penalty term.

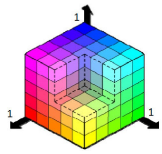

FIGURE 4. RGB Color Space.

- the third term is a sparse term ensuring data sparsity in the m-th layer and allows more obvious features to be obtained. The sparse term includes Kullback–Leibler divergence between two Bernoulli random variables with ω and $\mathrm{h}_\mathrm{i}^{(\mathrm{m})}$, where:

$$KL\left(\omega \parallel \bar{h}_i^{(m)}\right) = \omega\log\frac{\omega}{\bar{h}_i^{(m)}} + (1-\omega)\log\frac{1-\omega}{1-\bar{h}_i^{(m)}} \tag{10}$$

With ω as the sparsity target of the median layer and $\mathbf{h}\ (t,i\ m)$ as the i-th element of $\mathbf{h}(tm)$. Further, β controls the strength of the sparse term and when the sparse term is minimized, $\mathbf{h}_i^{(m)}$ is close to ω.

$$\bar{\boldsymbol{h}}_i^{(m)} = \frac{1}{2}\left(1 + \frac{1}{N_V}\sum_{t=1}^{N_V} h_{t,i}^{(m)}\right) \tag{11}$$

To generate a driving color map with different colors, an average value of the hidden features has been supposed as located in the center of the RGB color space. Therefore, a value of 0.5 has been set for ω because the center of each axis of the RGB color space is 0.5 (FIGURE 4). Thus, in the visualization method–driving color map, the generated colors do not tend to appear biased (e.g. reddish, bluish, etc.).

In order to monitor the Kullback–Leibler divergence, the plots reported in FIGURE 5 show the latent features moving towards the center of the RGB space, preserving information while ω increases). Since the range of the tanh function is [-1,1], the latent features result closer to 0, namely the center of the space.

Similarly, the same latent features can be plotted as time-series. In this case, the three extracted latent features are represented as time-series of three variables, namely the Red, Green and Blue color (FIGURE 6).

Therefore, in training the proposed model a backpropagation (BP) is implemented to raise reconstruction accuracy and, in the meantime, reduce the overfitting problem. The BP method performs partial differentiations of the weight matrices and biases for the objective function through chain rule. Therefore, the weight matrix $\mathbf{W}(l)$ and the bias vector $\mathbf{b}(l)$ between the l-th and $(l\text{+}1)$-th layers are updated by:

$$\mathbf{W}^{+(l)} \leftarrow \mathbf{W}^{(l)} - \eta^{(l)}\frac{\partial O(\sum)}{\partial \mathbf{W}^{(l)}}, \tag{12}$$

$$\mathbf{b}^{+(l)} \leftarrow \mathbf{b}^{(l)} - \eta^{(l)}\frac{\partial O(\sum)}{\partial \mathbf{b}^{(l)}}, \tag{13}$$

where $\mathbf{h}(l)$ represent the Learning Rate, equal for each hidden layer. To prevent the weight and bias from converging to an inaccurate local minimum, a greedy layer-wise training method is used.

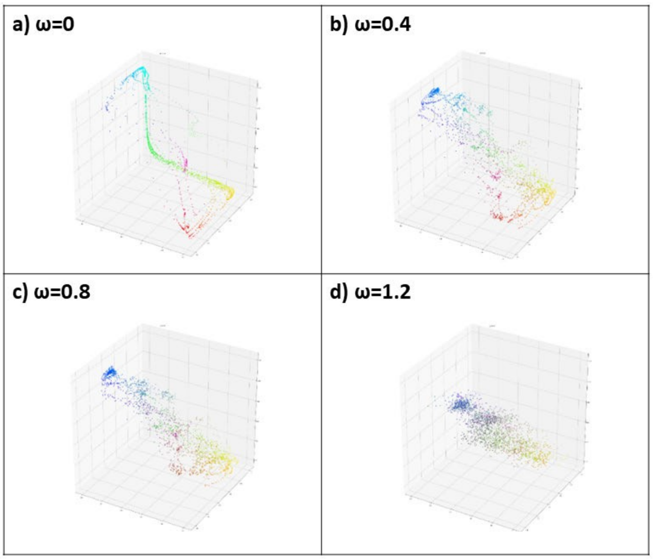

FIGURE 5. Latent features visualization with different ω values.

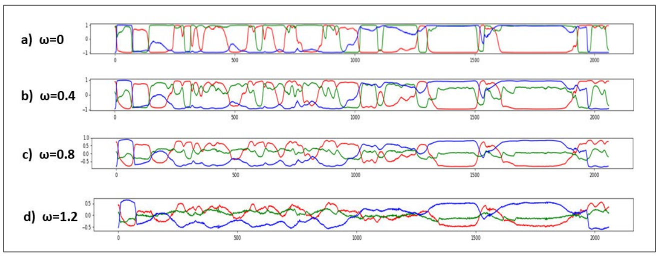

FIGURE 6. Latent features time-series with different ω values.

f: TRAINING AND VALIDATION

After the setting of all the necessary hyper-parameters, training and validation have been done on representative datasets. To evaluate the effectiveness of low-dimensional representation, the SDAE has been compared to other conventional methods from the viewpoint of linear separability of elemental driving behaviour. As a result, our methods outperformed other conventional methods in processing large amount of data.

The driving color maps generated by PCA, kernel PCA and DSAE for the 2nd and 3rd lap are shown in the FIGURE 7. Also, the observation of the extracted colors allowed the association with different selected "basic" driving behaviors, summarized in Table 2.

It is noticeable that other methods do not permit a differentiation of the same basic driving behaviors, as the generated colors are similar. For instance, more than one driving behavior (high speed forward and change in acceleration) correspond to a similar color (■) using the PCA. Similarly, the kernel PCA method characterized many driving behaviors with the same color (■) with slightly different shades.

Looking at the maps created with the PCA and kernel PCA it is possible to notice same segments with different colors. Obviously, this may be due to the driver behavior itself or just to a different combination of the input, but theoretically should not lead to different colors.

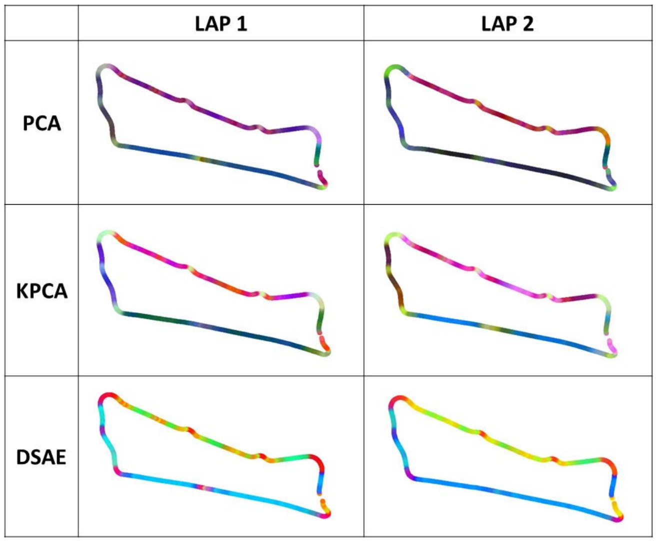

FIGURE 7. Model validation with driving color maps.

TABLE 2. Basic driving behaviors with representative colours.

	PCA	Kernel PCA	SDAE
High speed forward			
Low speed forward			
Accelerate / Decelerate			
Turning left / right			
Stopping vehicle			

The latent features of the different methods (Table 2), as a last confimation, allow to discourage the use of PCA and kernel PCA for non-linear dataset as driving behaviour. Vice versa, the latent features extracted with DSAE look like roughly placed with the same criterion for each lap and, moreover, may allow to distinguish the different driver behaviors by their relative position in the latent space. That kind of result is mainly due to the regularization techniques applied to the neural network, that allow the DSAE to connect and arrange the features in the same way despite the different origin of the input feature.

III. DISCUSSION

The consequent testing of the model is discussed, as the main scope of this research is the identification of the previously mentioned experimental variables (TABLE 1) on the generated driving color maps. Testing has been done exploiting all the datasets from driving tests.

Trials were performed on datasheets containing one single lap of a driver. At a first glance, the chosen model resulted able to work any series of input data that concerns with the same features.

a: IDENTIFICATION OF ROAD COMPLEXITY

In order to verify the method capability to identify the type of road scenario (as previously anticipated in TABLE 1 and schematically reported in FIGURE 9, the route is composed

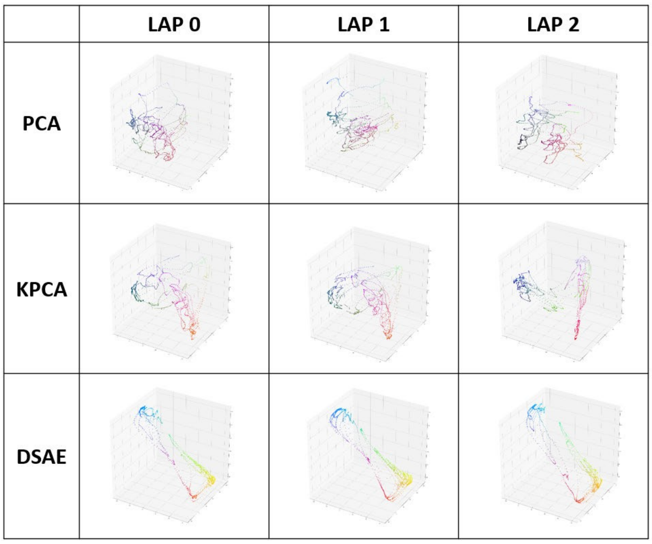

FIGURE 8. Model validation with features extraction in the latent space.

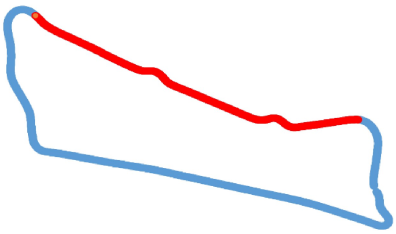

FIGURE 9. Complex (red) and easy (blue) driving situations.

by two segments defined as hard and simple context), have been tested the dataset of all drivers on the lap 3.

The complex contest refers the street Antonio de Curtis and have a complex road geometry (several lanes for both directions, intersections, roundabouts.). The simple contest refers instead to the streets Fausto Coppi and Giovannini that regards a residential area, with one lane for each direction of travel, low traffic volume.

The analysis of the obtained maps for all laps of all test participants (examples in FIGURE 10), fully confirms the identification of both contexts, as for all the maps the two road stretches have significantly different colors:

- The complex scenario - despite roundabouts and pedestrian crossings - results mainly in red or yellow/green color;
- The easy scenario results in blue colors.

b: DRIVER'S WORKLOAD

Subjective assessments have been proposed to measure driver effort during the driving task. The most common techniques are scales for the subjective mental workload. Examples are the NASA task load index, subjective workload assessment technique (SWAT) and the rating scale mental effort (RSME).

Among them, the NASA Task Load Index (NASA-TLX) is the most commonly used tool to rates the workload and most studies choose the standard NASA-TLX scales to conduct the subjective evaluations. More in detail, the tool includes

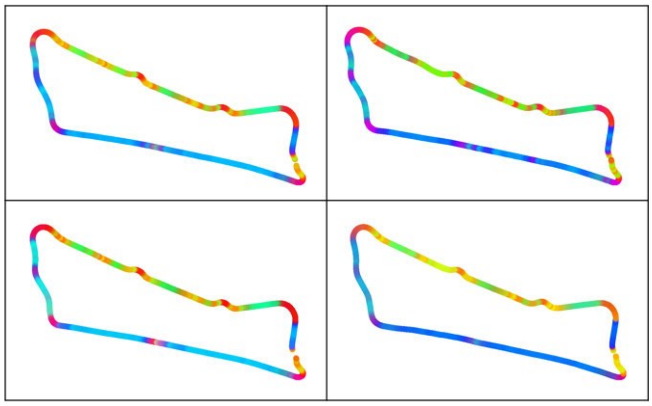

FIGURE 10. Example of Driving color maps for 4 drivers on Lap 3.

TABLE 3. NASA-TLX resulting scores.

Driver	TEST1	TEST2	Difference
1	39.0	24.3	-14.7
2	41.0	39.0	-2.0
3	28.0	23.3	-4.7
4	57.0	51.3	-5.7
5	31.7	24.7	-7.0
6	70.0	40.3	-29.7
7	34.7	20.3	-14.3
8	56.3	26.7	-29.7
9	53.0	38.3	-14.7
10	44.7	22.0	-22.7

a rating on six different subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. They are rated for each task within a 100-points range with 5-point steps. The ratings are combined to the task load index by create an individual weighting of these subscales by letting the subjects compare them pairwise based on their perceived importance. This requires the user to choose which measurement is more relevant to workload. The number of times each is chosen is the weighted score. This is multiplied by the scale score for each dimension and then divided by 15 to get a workload score from 0 to 100, the overall task load index [18].

In this paper, the workload obtained with the NASA-TLX Test was compared to driving color map of each driver in order to evaluate if colors are predictive of the drivers' cognitive load.

The numerical results of NASA-TLX show a significant difference in workload between the two driving sessions for all the participants. In particular, the first test result is more demanding than the second, coherently with an increasing confidence with experimental conditions, and determining an average difference in workload score of -14.52 between all drivers (Table 3).

A comparison between the first and the second driving session for each driver was carried out. Indeed, any color difference resulted was attributable to a variation in workload.

The results obtained show that in the first test session the road path followed by the users is not particularly stressful from the point of view of mental load, considering that the evaluation range goes from 0 to 100 and the average workload of users is around 50 (with some exceptions such as user 6).

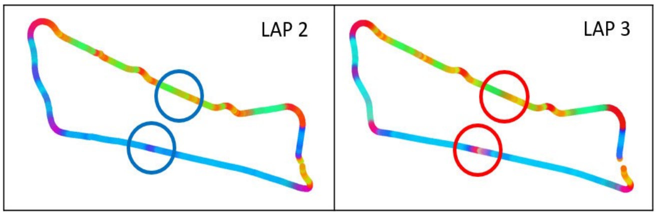

FIGURE 11. Pedestrians recognition between lap 2 and 3.

This indicates that drivers have been subject to an average level of strain. Comparing test 1 with test 2, we can see that, despite the increase in traffic during test 2, the value of the NASA-TLX score decreases for all users. This indicates that drivers tend to relax with respect to the initial test due to the previous knowledge of the track, the habitual effect of the track and the already known testing modes.

This trend is certainly caused by the habitual effect, i.e. the fact that in general a person who normally drives a road is subjected to a lower workload than a non-regular driver.

c: IDENTIFICATION OF EXTERNAL EVENTS

To verify the capability to graphically identify the presence of a pedestrian on the crosswalks, the 2nd and 3rd lap maps for each driver were compared with the expectation of a difference in the color pattern between the two maps (FIGURE 11). This difference results only for some drivers: color maps without pedestrian (lap 2) maps show orange color for the hard road context and dark blue color for the easy, while maps with pedestrian (lap 3) show dark orange shading into grey for the hard road context and reddish shades for the easy.

The overall precision of the method in recognizing pedestrians have been evaluated considering also false positive and negatives cases (True Positive = 13; False Positive = 4; False Negative = 3) and it resulted in a True Positive rate equal to 0,76 (TP/[TP+FP]).

IV. CONCLUSION

This study proposed an approach for extracting low-dimensional time series of latent features from multi-dimensional driving behaviour data using DSAE where Hyperbolic Tangent is set as activation function, the cost function integrated a L2 penalization term and a Kullback-Leibler divergence term.

From a theoretical point of view, the low-dimensional time series of latent features extracted using DSAE proved useful for driving behaviour visualization. Feature extraction were robust against defects and outliers. This is a direct consequence of the training method used on the DSAE, namely the back-propagation method that minimize the square error between the input data and the reconstructed data. The research demonstrated also that dataset with high correlated inputs features obtained best results in term of defects reparability and latent features extraction.

The obtained driving color maps represent an immediate visualization tool considering the potential impacts on road safety of driver behaviour recognition from large datasets. It is possible to evaluate this first attempt as a successfully one, as resulting in marked capability of the method to recognize road complexity and a satisfying capacity to visualize external events (i.e. pedestrians walking on crosswalk).

For future studies is envisaged the necessity to involve different categories of experimental variables in order to go beyond the limit of using only one typology of data (kinematic data). In particular, it is expected that physiological drivers' measurements (i.e. oculometry, direct measure of workload) and road conditions, if implemented, would add significance to the graphical output.

ACKNOWLEDGMENT

The authors thank the students Simone Frisco and Giorgia Bulgarelli for their valuable effort.

REFERENCES

[1] *Global Status Report on Road*, World Health Organization, Geneva, Switzerland, 2018, p. 20.
[2] U. Kramer, "On the application of fuzzy sets to the analysis of the system driver-vehicle-environment," *Automatica*, vol. 21, no. 1, pp. 101–107, Jan. 1985.
[3] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 37–52, 1987.
[4] C. Zhao, C. Zheng, and M. Zhao, "Classifying driving mental fatigue based on multivariate autoregressive models and kernel learning algorithms," in *Proc. 3rd Int. Conf. Biomed. Eng. Informat.*, Oct. 2010, pp. 2330–2334.
[5] W. Dong, J. Li, R. Yao, C. Li, T. Yuan, and L. Wang, "Characterizing driving styles with deep learning," Cornell Univ., New York, NY, USA, Tech. Rep. 1607.03611, 2016.
[6] K. Dwivedi, K. Biswaranjan, and A. Sethi, "Drowsy driver detection using representation learning," in *Proc. IEEE Int. Advance Comput. Conf. (IACC)*, Feb. 2014, pp. 995–999.
[7] J. Meseguer, C. T. Calafate, J. C. Cano, and P. Manzoni, "Characterizing the driving style behavior using artificial intelligence techniques," in *Proc. 38th Annu. IEEE Conf. Local Comput. Netw. (LCN)*, 2013, pp. 4–6.
[8] H. Wu, H. Gao, S. Wang, and D. Jiang, "Recognition of road type for unmanned vehicle based on texture and color features," in *Proc. 2nd Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2017, pp. 189–193.
[9] Z. Camlica, A. Hilal, and D. Kulic, "Feature abstraction for driver behaviour detection with stacked sparse auto-encoders," in *Proc. IEEE Int. Conf. Syst., Man, (SMC)*, Oct. 2016, pp. 3299–3304.
[10] H. Liu, T. Taniguchi, Y. Tanaka, K. Takenaka, and T. Bando, "Visualization of driving behavior based on hidden feature extraction by using deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2477–2489, Sep. 2017.
[11] P. Thomas, B. Price, C. Paine, and M. Richards, "Stacked denoising autoencoder," *J. Educ. Technol.*, vol. 33, no. 5, pp. 537–549, 2002.
[12] J. Zhang, Z. Wu, F. Li, C. Xie, T. Ren, J. Chen, and L. Liu, "A deep learning framework for driving behavior identification on in-vehicle can-bus sensor data," *Sensors*, vol. 19, no. 6, p. 1356, Mar. 2019.
[13] J. Guo, Y. Liu, L. Zhang, and Y. Wang, "Driving behaviour style study with a hybrid deep learning framework based on GPS data," *Sustainability*, vol. 10, no. 7, p. 2351, Jul. 2018.
[14] V. Vignali, F. Cuppi, E. Acerra, A. Bichicchi, C. Lantieri, A. Simone, and M. Costa, "Effects of median refuge island and flashing vertical sign on conspicuity and safety of unsignalized crosswalks," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 60, pp. 427–439, Jan. 2019.
[15] M. Costa, A. Bichicchi, M. Nese, C. Lantieri, V. Vignali, and A. Simone, "T-junction priority scheme and road user's yielding behavior," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 60, pp. 770–782, Jan. 2019.
[16] G. Di Flumeri, "EEG-based mental workload assessment during real driving: A taxonomic tool for neuroergonomics in highly automated environments.," *Neuroergonomics*, vol. 12, pp. 121–126, Jan. 2019.
[17] G. Di Flumeri, "EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings," *Front. Hum. Neurosci.*, vol. 12, pp. 1–18, Dec. 2018.
[18] S. G. Hart and L. E. Staveland, "Development of NASA-TLX," *Hum. Ment. Workload. Adv. Psychol.*, no. 52, pp. 139–183, Jan. 1988.

ARIANNA BICHICCHI received the master's degree in civil engineering from the University of Bologna, in 2011, with a specialization in infrastructures and transportation. She is currently pursuing the Ph.D. degree with the XXXI cycle at the DICAM Department, School of Engineering, University of Bologna, with a research project entitled innovative technologies for road—users interaction analysis, who proposes to analyze the contribution of the human factors in road safety. She has coauthored several scientific publications focusing on the use of Eye Tracker and electroencephalographic headset (EEG) for the evaluation of drivers' attention, inattention, and workload.

RACHID BELAROUSSI received the master's degree in physics from ESPCI, the master's degree in electronics from UPMC, and the Ph.D. degree in computer vision on appearance-based model (auto-encoder) applied to face detection, in 2006. He joined the French Institute of Science and Technology for Transport (IFSTTAR), in 2011. He has worked since on vehicle perception and driver characterization.

ANDREA SIMONE has been an Associate Professor with the School of Engineering and Architecture, University of Bologna, since 2004. He has served on European and international (AIPCR-PIARC) technical committees. He has authored about 100 publications. His main research areas are mobility analysis and design, highway design, road safety analysis, road safety design, and human factor interaction.

VALERIA VIGNALI received the Ph.D. degree in transportation engineering (civil engineering). She is currently an Assistant Professor with the Department of Civil, Chemical, Environmental, and Materials Engineering, School of Engineering and Architecture, University of Bologna. Her research interests concern road, railway and airport infrastructures, with particular reference to the analysis of the road-drivers interaction by innovative technologies. She authored numerous scientific articles published in international journals.

CLAUDIO LANTIERI received the Ph.D. degree in transportation engineering. He has been a Junior Assistant Professor, since 2016 and a Senior Assistant Professor, since 2019 with the School of Engineering and Architecture, University of Bologna. He has been teaching and Teaching Assistant in all courses of the DICAM—Roads area, with activities assistant supervisor in several dissertations. He has been a member of the SIIV (Italian Society of Road Infrastructures), since 2006, he has participated in several conferences and seminars and has authored publications in journals and national and international conferences on topics related to the subject area ICAR04.

XUANPENG LI received the Ph.D. degree from University Paris Sud, in 2014, on the study of vehicle safety based on multisource information fusion for ADAS. In 2015, he joined Southeast University as Assistant Professor. His is interested in evidence theory for driver inattention estimation and machine learning for 3D semantic mapping and visual signal language recognition.

• • •