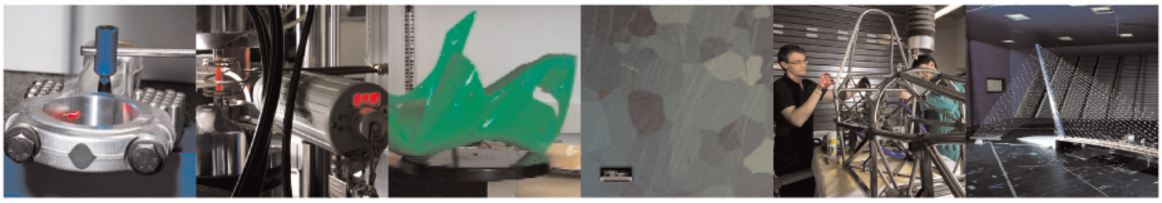
POLITECNICO MILANO 1863

DIPARTIMENTO DI MECCANICA



Perception of driving simulations: Can the Level of Detail of virtual scenarios affect the driver's behavior and emotions?

Y. Shi, M. Boffi, B. Piga, L. Mussone, G. Caruso

This is a post-peer-review, pre-copyedit version of an article published in **IEEE Transactions on Vehicular Technology**. The final authenticated version is available online at: http://dx.doi.org/10.1109/TVT.2022.3152980

Perception of driving simulations: Can the Level of Detail of virtual scenarios affect the driver's behavior and emotions?

Y. Shi, *Member, IEEE*, M. Boffi, B. Piga, L. Mussone, G. Caruso

***Abstract*—Human factors studies are becoming more and more crucial in the automotive sector due to the need to evaluate the driver's reactions to the increasingly sophisticated driving-assistant technologies. Driving simulators allow performing this kind of study in a controlled and safe environment. However, the driving simulation's Level of Detail (LOD) can affect the users' perception of driving scenarios and make an experimental campaign's outcomes unreliable. This paper proposes a study investigating possible correlations between driver's behaviors and emotions, and simulated driving scenarios. Four scenarios replicating the same real area were built with four LODs from LOD0 (only the road is drawn) to LOD3 (all buildings with real textures for facades and roofs are inserted together with items visible from the road). 32 participants drove in all the four scenarios on a fixed-base driving simulator; their performance relating to the vehicle control (i.e. speed, trajectory, brake and gas pedal use, and steering wheel), their physiological data (electrodermal activity, and eye movements), their subjective perceptions, opinions and emotional state (questionnaires concerning the research and Self-Assessment Manikin Scale) were measured. The results showed that driver's behavior changes in a very complex way. Geometrical features of the route and environmental elements constrain much more driving behavior than LOD does, as observed for vehicle trajectory and speed; skin conductance and gas pedal appear to be more sensitive to LODs. Gaze position changes according to LODs, but differently for rectilinear and curvilinear segments. Pupil diameter increases with higher LODs. Emotions are not affected by LODs. Generally, different signals showed different correlations with the LOD level, suggesting that future studies should consider their measures while modeling the virtual scenario. It is hypothesized that scenario realism is more relevant during leisurely environmental interaction, whilst simulator fidelity is crucial in task-driven interactions.**

Index Terms— Driving simulator; simulation reliability; level of detail (LOD); driver behavior; eye tracker; environmental psychology

I. INTRODUCTION

IMPROVING roads' safety and the overall driving experience are important objectives that the automotive sector has been addressing over the decades, even through the analysis of driver's behavior. In driving behavioral assessment studies, car simulators are widely used, rather than conventional field test methods, due to their advantages in terms of safety and reliability [1]–[6]. The benefits of reliable driving simulators, with their safe and highly replicable scenarios, brought to increased use of such technologies for behavioral studies, especially in the transition from manual to autonomous driving [7]. Indeed, the number of car simulators worldwide is continuously increasing, as well as the literature on driving performance and behavior through simulator studies [8]. Even if simulators cannot fully replicate the real-world experience, they offer the researcher an advantage that real-world studies cannot match: the ability to control and repeat experimental conditions and create prescribed scenarios with selected variables [9]. This enables manipulating several variables for revealing different driver responses to dissimilar conditions in a short timeframe, which would be substantially impossible otherwise [8], [10]. The multiple variables influencing drivers' behaviors [11] might be subjective or environmental, that in turn can be temporary and non-temporary [12]; for instance, in the first case, we can assess the driving performance according to age [13] or the effects of alcohol, drugs, and/or fatigue on driving behavior [14], whereas in the second case it is possible to appraise the driving performance in different environmental conditions, such as different types of adverse weather conditions or day/night hours [15]. Driving simulators can be effectively used to evaluate these variables [16] both at the end of the simulation experience or/and in real-time along with the testing session, allowing scholars to make comparisons with an analogous experience in the real environment [17]–[19].

Usually, driving simulator validation studies tended to assess speed, speed adaptation [20], and lane-keeping [21] to measure the human abilities in vehicle control. However, the driver's behavior also involves other elements, associated with the car controlling and decision-making situations (e.g., speed control, driving maneuver, overtaking, stopping), that requires the definition of the driver's psychophysical states [22]. The subjective state is described through psychological and physiological measures, which can account for the effects of individual variability and/or type of driving situation. For instance, higher-level physiological measures, like the frequency-domain measure of Heart Rate Variability (HRV) [23] and Skin Conductance (SC), are valuable indexes to measure cognitive effort and workload [24]. Each signal could be correlated with a specific psychological parameter that informs about a driver's emotional or cognitive state. For example, by measuring the SC, the psychological or physiological arousal level related to a specific event during the experiment can be captured numerically, e.g. the increase of SC indicates that the event enhances the stress level. Through the analysis of these data the arousal level of the user can be derived with statistical accuracy.

This article addresses the issue of scenario's reliability of car simulators in the mobility field in relation to the notion of Level of Detail (LOD) of the virtual scenario, i.e. the visual simulation. The hypothesis is that LODs affect drivers' behavior and emotions and, consequently, the reliability of the driving appraisal. Indeed, in driving simulation, the 3D environment is crucial to provide the user with an immersive experience during the driving tasks; however, modeling is a demanding and expensive part of the overall simulation procedure, and this can affect its proper applicability. This cost/effectiveness relationship is important. Reducing the effort in producing virtual scenarios also means decreasing time and costs, which are indubitably highly relevant for professionals. At the same time, producing accurate driving scenarios from scratch is crucial for getting reliable outcomes. Thus, defining the minimum requirement (i.e. minimum effort/cost) is profitable both for professionals and decision-making of public authorities in the urban planning field. In this perspective, this article aims at contributing to the investigation of the ecological validity of visual simulation and in detail how and to which extent the LOD affects the driving behavior and subjective experience. This study presents a comparative analysis to evaluate the impact on drivers' experience of four different LODs of a road visual simulation navigated with a fixed-base driving simulator. Thus, four driving scenarios representing the same area were modeled with varied graphical details and physical elements. During the in-vitro driving experiment, several data were collected: vehicular information, driver's physiological signals, eye-tracking signals, and psychological variables. Statistical analysis and comparison of the outcomes enabled the evaluation of the driving experience.

The article is organized as follows. Section II depicts the state-of-the-art review, Section III presents the methodology adopted to carry out the experiments, Section IV explains the data processing methodology, Section V shows the results achieved, Section VI presents the discussion, and Section VII draws the conclusions.

II. State of the Art

To design a sound methodology, three main elements, namely participants, instruments, and procedure, should be carefully considered to gain reliable results in in-vitro experiments aiming at investigating people's experience through simulation. Although the core of this article is the scenario, some reflections on the simulator and the instruments for data collection are useful when reasoning about the ecological validity of simulations. The instruments are indeed crucial in determining the overall reliability of the experiments through simulation in the field of mobility and urban design. These may concern: (i) the simulator, i.e. the device or interface used for running the experiment, e.g. car, bike, or other types of structured simulators, but also simple visualizations or audio-visual media that act as surrogates of reality administered through immersive or non-immersive devices; (ii) the scenario, i.e. the medium representing the environment, e.g. the urban or extra-urban environment to drive through, with its specific temporary and non-temporary characteristics; and (iii) the instruments for data collection, i.e. the specific tools used for gathering the data, e.g. physiological sensors, psychological scales, and/or behavioral observations.

According to the instruments' combination and their level of accuracy, the reliability of the overall simulation process can vary. Indeed, the simulator and the simulation (i.e. scenario) together produce a combined level of perceived realism (see also Section II-C phycological fidelity). To assess the overall simulation reliability it is possible to refer to the 'objective method' and the 'performance matrix' [25]. The first method "calculates the number of identical elements shared between the real world and the simulation; the greater the number of shared identical elements, the higher the simulation fidelity" [25] (p.62), whereas the second compares the users' reactions to the simulation and the reality it represents, the higher the similarity, the greater the fidelity [26], [27].

A. Simulator Fidelity

There is not yet a full and definitive agreement on the variables to consider for producing or assessing the reliability of simulators [28]–[30], even if the term fidelity is used as an umbrella of all the issues contributing to replicating the actual environment [25] in its specific contextual conditions. Anyhow, it is worth mentioning some well-established references addressing some of the subcategories contributing to the overall simulator fidelity. Indeed, the dichotomy of 'physical fidelity' and 'functional fidelity' by Fink & Shriver [31], recalled by Hays [32] and Allen et al. [33], is considered a pivotal distinction [25]. Allen et al. [33] clearly define physical fidelity as "the degree to which a training simulator 'looked like' actual equipment", whereas functional fidelity as "the extent to which it 'acted like' real equipment" (p. 498). These two elements contribute to generating the perceived realism of the device or interface, even though the functional fidelity is more important than the physical one for cognitive tasks [34]. 'Motion fidelity', i.e., "the degree to which a simulator can reproduce the sense of motion felt by humans in the operational environment" [25] is part of physical fidelity. The sense of speed is a key factor in driving simulation studies, that is why the common taxonomy in this field refers to the motion fidelity levels of simulators, typically defined as low-level, mid-level, and high-level. The low-level driving simulator is usually equipped with a fixed-base, namely with 0 Degrees of Freedom (DOF), the mid-level driving simulator starts to get more DOF in one or two directions, whereas the high-level driving simulator consists of a motion platform of at least 6 DOF [6], [16], [35]. The influence of the motion fidelity levels on driving behaviors has been investigated over the years in the mobility research field. Besides the dichotomy mentioned above, Kinkade and Wheaton [36] refer to the 'equipment fidelity', i.e. "the degree to which the simulator duplicates the appearance and "feel" of the operational equipment", the 'environmental fidelity', i.e. "the degree to which the simulator duplicates the sensory stimulation (excluding control feel) which is received from the task situation", and the 'psychological fidelity', which is "the degree to which the simulator is perceived by the trainee as being a duplicate of the operational equipment and the task situation" (p. 679) (see Section II-C).

B. Scenario Reliability

A reliable scenario in the experiential simulation field should activate reactions comparable to those that would occur in the same actual environment. In the field of environmental psychology, McKechnie [37] defined these types of scenarios as 'perceptual simulation' (also named by scholars as 'experiential simulation' or 'realistic simulation'), i.e. the "attempts to provide tangible, concrete replicas or isomorphs of environments - often future environments - that can be displayed to observers for their evaluation or other response" (p.169). The ecological validity, i.e. "the extent to which findings utilizing the simulation laboratory are generalizable to the real environment represented in the model" [37] (p.183), is thus a crucial aspect of reliable simulations, even if, as Appleyard noticed [26], "a simulation attempts to represent reality. It does not, indeed cannot, reproduce that reality in toto. Rather it selects critical aspects of that reality for the particular purposes at hand" (p. 43). Following this approach, the identification of the needed scenarios' characteristics for a reliable replica of reality, that enables to gain sound outcomes, are addressed over the years by several authors in the field of architecture and urban planning [12], [26], [38], [39]. As highlighted by Piga and Morello [12], it is possible to identify in the literature some crucial elements for unbiased - or almost unbiased - simulated scenarios, namely (i) the geometrical coherence between the simulated and actual environment, (ii) the proper level of detail and the graphical quality of the representation, (iii) the attention in depicting the atmosphere of places, (iv) the inclusion of dynamic aspects within the simulated environment, (v) the simulation size and distance from the observer in relation to the human optics. However, as Sheppard noticed [38], "a certain amount of bias in people's responses is to be expected even with a good simulation" (p. 62). At the same time, the identification of the proper realism of the scenario should also be evaluated with regard to the specific context of the application [40]. This is particularly relevant when we move from the field of environmental psychology and architecture to mobility studies, where the goal of the investigation through simulation might be, for instance, the evaluation of the driving performance instead of the driving experience in the environment itself; indeed, in this case, the level of realism might be lower than in other experiential assessments. As for the simulator, it is hard to establish an ever-valid framework of the scenario necessary realism [41]. Undoubtedly, as remarked by Lange [42] "even simulations with a lower degree of realism can still contain the most important information needed for a specific purpose" (p. 165). In addition, a high degree of realism is not necessarily required for gathering useful information on how an individual will act in each situation. For example, high vs. low realism in visual scenes (given comparable screen size and viewing angle) does not generally seem to have a major impact on driving performance variables [43]. Moreover, beyond the scenario characteristics per se, it is relevant to notice that to properly set or assess the quality of an experiential simulation it is necessary to consider how it is navigated by the final user, i.e., simulation size, distance, and point of view of the viewer [39]. In this respect, it is worth noting that from the mono screen to the 360-degree projection view, the visual fidelity in the driving simulator improves greatly [44]. Some elements of the scenario are in any case relevant, for instance textures and entities (for example the trees and plants), in the virtual environment since these impact driving behaviors [44]. These considerations highlight the importance of choosing the correct Level of Detail (LOD) of the simulated scenarios, i.e. of the virtual 3D model, for the specific purpose. Identifying the proper LOD in a logic of a cost-effective procedure is crucial also because the creation of the virtual environment is often time demanding and typically expensive [44]. The term LOD was originally used in describing the hierarchical structure of 3D model polygons [45]. The impact of different LODs on the human visual perception has been conducted in various virtual environments [46]–[49]. Specifically, graphics quality of driving simulation generated by three rendering methods were evaluated in a subjective evaluation of their impacts on the perceived visual fidelity and the over-all experience [50]. This term was imported from 3D computer graphics to city modeling [51]; in this field, the Open Geospatial Consortium (OGC) defined the CityGML 2.0 (City Geography Markup Language) standard, i.e. "an open data model and XML[Extensible Markup Language]-based format" (p. 9) for digitally representing and sharing structured information. In detail, CityGML is an application schema of the GML3 geometry model based on the ISO 19107 "Spatial Schema", and it is specifically designed for representing and sharing 3D virtual geometries of city and landscape models [52], [22]. LODs are defined according to five consecutive architectural scales with their related details: LOD0 regional, landscape; LOD1 city, region; LOD2 city districts, projects; LOD3 architectural models (outside), landmarks; LOD4 architectural models (interior). In parallel to the identification of geometrical entities related to the above-mentioned architectural scales, a taxonomy of semantic features, based on the ISO 19109, is identified to thematically characterize the entities in spatial or non-spatial. The experimental approach proposed in this article refers to the well-established OGC original distinction of geometrical LODs, from LOD0 to LOD3.

C. Psychological fidelity

In parallel to the studies mentioned above on the simulator and the scenario, scholars examined fidelity issues from the subjective perspective. According to Hays [53], the psychological fidelity is not strictly related to the actual realism of the simulator since "the level of psychological fidelity would be high if the trainee perceives the simulator as being highly realistic, even though it might deviate substantially from the actual equipment it is supposed to represent" (p. 3). This approach emphasizes the importance of integrating people's perception among the criteria for evaluating the overall simulation method: such a psychological-cognitive fidelity describes how much the psychological factors and cognitive effort are duplicated during the simulation testing phase [54]. In such a perspective, the scenario's objective features are not a reference for the direct assessment of fidelity, they are rather the independent variables for comparing the people's reaction to the actual and the simulated physical environment [25]. In this vein, Lombard and Ditton [55] developed the fruitful concept of 'presence', defined as a perceptual illusion of non-mediation which allows the users of any digital support to interact in a realistic way with a simulated environment. In psychological terms, it must not be conceived simply as a reaction to sensory data, but instead as a complex experience including various cognitive processes [56]. Studies on the psychological experience in simulated environments showed how the same VR scenario elicited a greater sense of presence with a more realistic representation (e.g., more foliage on the ground for the natural environment, more abundant and sophisticated shop interiors for the urban environment) [57].

Hence, research goals should define the type of simulator and scenario fidelity needed and also the type of psychological fidelity. Depending on such goals, scholars should identify the most appropriate task to assign to participants, the relevant aspects of the subjective experience to monitor, and thus select the assessment tools accordingly. For instance, when one is interested in studying the driving experience, emotions certainly play a key role, and their assessment can provide relevant information. Indeed, the driving performance is significantly affected by emotions, which can lead to a diminished capability of controlling the steering wheel [58] and more intense use of gas and brake pedals [59]. Such a relationship between emotions and driving performance can have important implications, as it can inform the improvement of training for driver's licenses [60] or the design of safer environments detecting near-miss accidents [61].

III. Method

A. Participants

The experiment initially involved 32 volunteer students from the Politecnico di Milano, yet, due to errors in the data acquisition of some participants, only 29 were counted as valid users, 66% males and 34% females, due to the collected data quality. Their age ranges from 21 to 26 years (M=23.31, SD=1.17). All participants hold a valid driving license, and 45% have more than 5 years of driving experience, whereas 55% have 2-5 years of driving experience. 35% of them drive every day, 17% 3 times a week, 24% once a week, 17% 2 times a month, whereas 7% never drove during the last three months before the experiment. Only 14% of them previously experienced a crash, and 28% were fined in the past for breaking highway rules, but not the same as those involved in car crashes.

B. Tools and techniques

1) Simulator

The proposed method is based on the use of a fixed-based driving simulator with a realistic driver seat, including a force-feedback steering wheel, a gear shifter, pedals, a surrounding audio system, and a three-monitor visualization system, which provides a 175-degree horizontal field of view. Monitors' diagonal size is 32 inches, and their resolution is 1920x1080 pixels. Drivers are seated at about 80cm from these screens by adjusting their posture according to their habits. All the lighting facilities, screen brightness, environmental temperature, and noise had been maintained the same during the experiments. The software used for the simulation is the IPG CarMaker [62], which is a professional driving simulator software that enables to build a realistic driving experience. The simulator's hardware includes the commercial physiological sensor ProComp Infiniti System [63] and the eye tracker Pupil Labs Core [64].

2) Driving scenarios

The driving scenarios relate to a real urban area, which is about 350m x 350m wide, and the path is 1.107 km long. The 3D models were built starting from the area's 2D map, and textures were elaborated from photos of the visible facades. Four scenarios with different LODs were elaborated: from LOD 0 to LOD 3, where LOD 0 is the testing scenario with minimal details, LOD 3 has the highest level of details (Table I, Fig. 1).

3) Vehicular, physiological, and eye-tracking measures

Vehicular and physiological measures are commonly used to assess the driver's behavior [4], [18]. While vehicular data provide information about the execution of the driving task [65], physiological measurements like the frequency-domain measure of Heart Rate Variability (HRV) [23], Skin Conductance (SC), are valuable indexes to measure cognitive effort and workload [24]. Also, physiological eye movements are a valuable index to evaluate the cognitive effort of drivers. Pupil size is sensitive to the mental workload changes in controlled environments, such as driving simulators [66], [67]. There is a correlation between pupil diameter increase and the increase of the cognitive load level [68]. The number of fixations and the fixation duration are both parameters related to drivers' mental workload [66]. Long fixation duration is typically associated with a high processing load [69].

During the experiment the following data were collected by monitoring the driver and the virtual vehicle: 1) Speed (speed of vehicle) [m/s]; 2) DMGas (use of gas pedal), varying in the range [0, 1]; 3) RoadDist (distance from the road median line) [m]; 4) SC (Skin Conductance) [μSiemens]; 5) Pupil size [mm]; 6) Gaze position, varying in the range [0, 1]; 7) Number of fixations; 8) Fixation duration [ms].

Physiological and vehicle data have been collected using external software running on a dedicated computer. This software can synchronize both signals at 10 Hz and manages the synchronization of the Pupil Labs data by monitoring specific trigger events on IPG CarMaker software through a network connection between the two computers. This synchronization allows simplifying the subsequent processing of the acquired data.

TABLE I
Four Levels Of Detail (LOD) and their meaning.

LOD	Architectural details
0	Route (including road markings, vertical signs, zebra crossings)
1	Route + Blocks of the building without roofs
2	Route + Blocks of the building with roof
3	Route + Building with roof and textures + Supplementary details

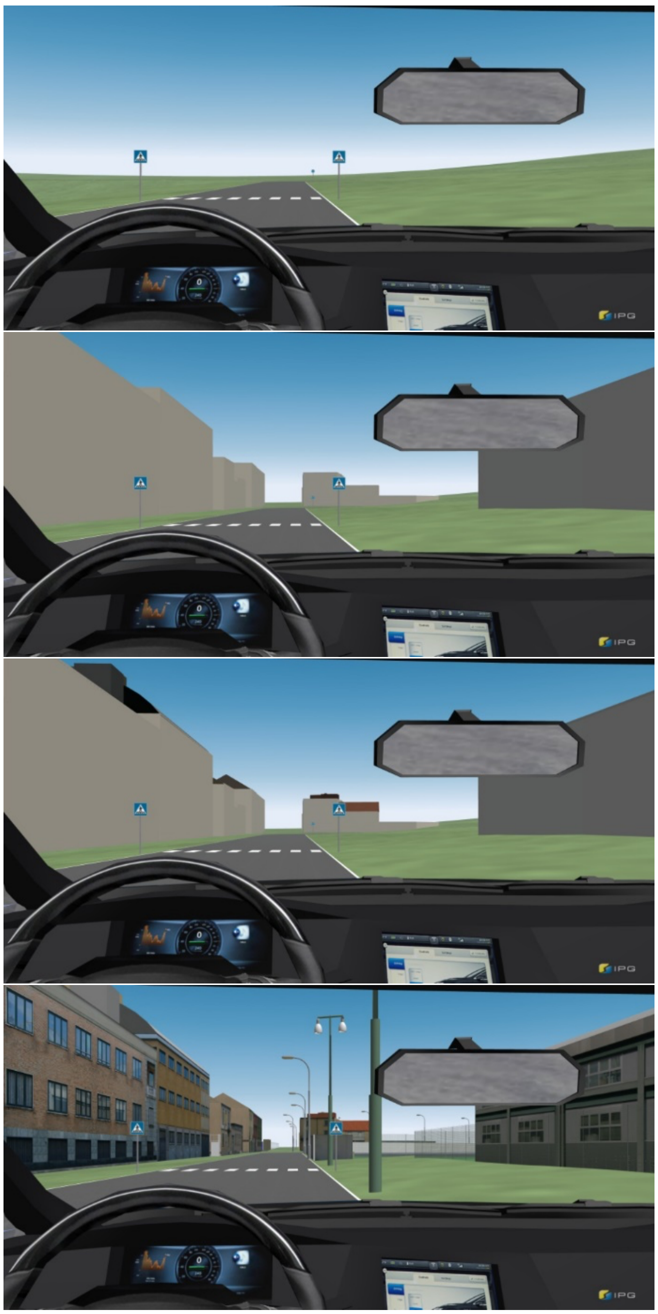


Fig. 1. The four LODs considered in the experiment: from the top LOD0, LOD1, LOD2, LOD 3 of the driving scenario.

4) Subjective psychological measures

The subjective reaction of participants to each LOD was described with a questionnaire comprising two parts. The first part is focused on emotions experienced during the driving phase, assessed with the Self-Assessment Manikin (SAM) [70], which is among the most used self-report tools in a broad range of theoretical and applied fields [71], [72]. It is a pictorial tool designed to measure the emotional state drawing on a threefold conception of emotions by Mehrabian & Russell [73], including pleasure (from happy to unhappy), arousal (from excitement to relaxation), and dominance (from autonomy to submissiveness). A stylized human character represents the subjective reaction to affective stimuli: fifteen pictures offer a non-verbal description to the respondent with five different degrees of intensity along the continuum for each of the three dimensions. The tool allows scholars to gather a quantitative assessment of these three dimensions of emotions, whose combination provides a multifaceted description of the emotional state. The second part aims at evaluating the driving experience and the scenarios in a broader sense with the following questions rated on a 5-points Likert scale (1=very bad, 5=very good): 1) How do you evaluate your driving experience? 2) How do you evaluate the realistic level of the driving circumstances? 3) How do you evaluate the adequacy of the virtual scenario to perform the driving task?

C. Procedure

The participants were invited to sign the informed consent, where the procedures and tasks of the experiment are described. Afterward, the participants were invited to fill in a questionnaire including socio-demographic data and the first round of SAM questions to get the participants' status at the beginning of the experiment. After these preliminary activities, the participants were invited to sit on the simulator and for adjusting the height, inclination, and distance from the steering wheel of the seat to reach a comfortable posture.

All participants started the experiment by driving for 3 minutes within an adaptation scenario. During this preliminary phase researchers elicited the participants' familiarity with the simulator by inviting them to check visibility, gas pedal, brake pedal, and steering wheel reactions. Then, the participants carried out the actual experimental task, which was driving 2 laps within the testing scenario developed with the different LODs at a maximum speed of 50Km/h. They were invited to stop driving and answer the questions at the end of the second lap. After the questionnaire, SC was recorded for 30 seconds in a relaxed state to reconstruct the baseline and then remove the bias in the collected physiological signal.

To assess the LODs' impact avoiding order bias, four different experimental sequences were adopted according to the following experimental design with codes A, B, C, and D. Sequence A identifies the series of LODs (0, 1, 2, and 3), sequence B LODs (1, 2, 3, 0), sequence C LODs (2, 3, 0, 1), and sequence D LODs (3, 0, 1, 2). This design balances the participants to start with different LODs and enables us to get data of different LODs' sequences. The impact of sequence on the acquired signals has been evaluated and never exceeds the impact of LODs. A debriefing session with participants about the driving experience was conducted at the end of the experiment. Fig. 2 shows the protocol relative to sequence with code A.

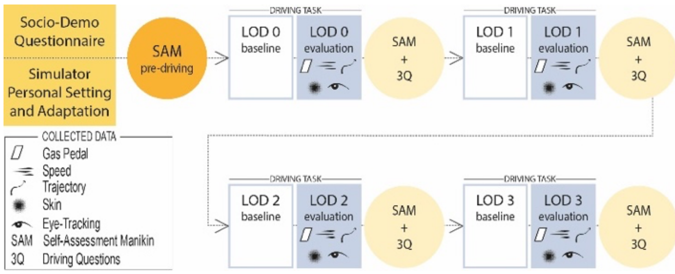


Fig. 2. The sequence of the experimental activities (code A).

IV. Data Processing

A. Segmentation of the path

The circuit was subdivided into 111 elementary segments 10 meters long. Then, an aggregation is made in order to consider rectilinear (1/r=0) and curvilinear stretches (r < 100m), resulting in 8 rectilinear segments (total 730m) and 8 curvilinear segments (total 390m) (Fig. 3).

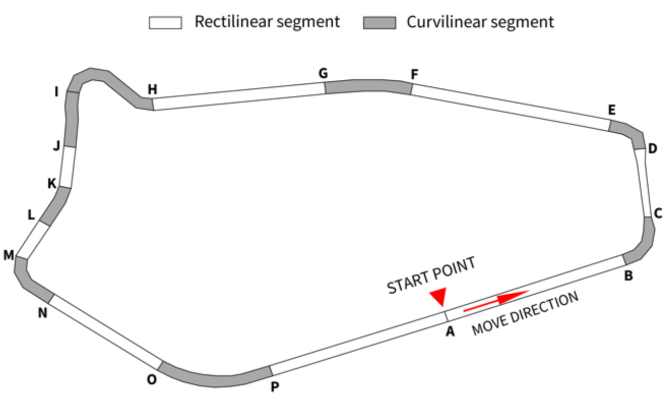


Fig. 3. Segmented path aggregated by type (rectilinear and curvilinear stretches) for the test circuit.

B. Time to space transformation

A pre-processing phase is required for enabling data comparison since data are collected at regular time intervals (100ms) during the driving simulation sessions. Since the speed changes according to driver and road characteristics, both physiological and vehicular measures may result in a different number of records. The slower the vehicle's speed, the higher the number of measures for the same portion of the path. For the sake of data comparison, we distributed the data at regular space intervals. By doing this, it is possible to use each road segment to compare the overall trend of the measures; indeed, all data series have the same length and, in this way, events (e.g. curving) are more clearly comparable and attributable to road geometry. Since the relationship time-space strictly increases monotonic (vehicles can go only in one direction and do not stop) and then bijective, the reciprocal function could be easily built using a classical fitting function. Then, we could calculate the new corresponding values of time by using an array of equally spaced distance intervals (e.g. we use a 0.1m step to increase resolution).

C. Find-Spots analysis

The effect of LODs on drivers has been preliminarily analyzed by a specific program named FindSpot. This routine compares behavioral and physiological measures of each driver, except for the eye movements, to find the spots along the path where the local max or min values (peaks) for the four LOD curves overlap. Two main parameters control the FindSpot routine: 1) the percentage difference of measures, 2) the breadth of the spot. Whether the difference between two peaks is less than 20% in the around 20 meters of the road (two elementary segments), the driver's behavior can be considered similar. The outcomes of this analysis give a first interpretation and a wide overview for understanding whether and where the behavior of every single driver is similar according to the four LOD testing scenarios.

D. Correlation analysis

Considering that Find-Spot analysis allows evaluating only local similarities of the measures, a correlation analysis was performed to have a more general overview of the measures' similarities through different LODs. The relationship between variables has been evaluated by Pearson's linear correlation coefficient. Correlation coefficients are in the range [-1, +1]; the 0 value means there is no relationship between variables, while -1 or +1 means there is a perfect negative or positive relationship between two series of measures. The correlation analysis was performed between all series for each pair of LODs. The average of the correlation values, which is needed when comparing groups of different measures to improve the normality of the data, was calculated by applying the Fisher's Z transformation [2] and the Oklin-Pratt formula [74]. Finally, an ANOVA analysis was performed to determine whether variation in the series of measures arises between the different groups.

E. Eye movements

The eye movements data required real-time and offline video elaboration to extract gaze position, pupil diameter, and fixations. The gaze positions were measured based on the x-coordinate and y-coordinate of the driver's gaze focus on the central display screen of the simulator. Markers were physically applied to the screens of the driving simulator to help the eye tracker track the gaze according to the monitor's position. The origin of the axis is located at the lower-left corner, and the value of the x-axis and y-axis coordinates ranges from 0 to 1.

The value of pupil diameters (left and right eye) of the driver was collected in real-time during the experiment. The data with a confidence level higher than 80% were filtered and are reliable for the analysis. Subsequently, the average diameter of the pupil was calculated based on an interval of 100 ms to make them consistent with those collected from the simulator (i.e. speed, gas, road distance).

The fixations detector was implemented with the dispersion-based method proposed by [75]. Three parameters for the detection threshold were set: maximum dispersion, minimum duration, and maximum duration. According to the definition in ISO15007, the fixation is the alignment of the eyes so that the image of the fixated area of interest falls on the fovea for a given time, for example, from 100 ms to 2000 ms [76]. It must be noticed that the eyes' alignment is not fixed but moves in a tiny dispersion. The max of this dispersion is 1.0 deg in the current study, taking reference from the study of [77]. Although the general threshold of the sampling time is wide, in this study the threshold of the sampling time is narrowed down to 140ms-300ms, considering the driving velocity and the route setup [75], [78]. The fixations ranging from 140 milliseconds to 300 milliseconds were taken into consideration. The data with a confidence level higher than 80% were taken as reliable data, and those data were filtered for the analysis.

F. Subjective measures

To appraise the emotional experiences across the four LODs, the average values of pleasure, arousal, and dominance are separately compared. Descriptive statistics represent the broader evaluation given of the driving experience after each LOD. Inferential statistics are applied to identify significant differences, with a one-way ANOVA run with LODs as independent variables, and the three emotional dimensions and the single items about the driving experience and the scenarios as dependent variables.

V. Results

A. Driving performance: behavioral and physiological reactions

The results of the Find Spot analysis (Table II) reveal that Speed and SC are the most varying measures between LODs, and spots are unlikely to occur. Generally, Speed spots occur in rectilinear segments and just after the curvilinear ones. SC spots, instead, are observed before the demanding parts of the road: curves with a small curvature ray. DMGas and RoadDist are less sensitive to LODs, and spots occur frequently. For DMGas, they occur mainly close to more challenging stretches of the path, where rectilinear and curvilinear segments alternate closely. A similar pattern also occurs for RoadDist, but max values occur just before curves and min values just after curves. A first conclusion is that driver's behavior is less or not at all conditioned by LOD in curvilinear segments, and the degree of conditioning depends on the type of measures too. Fig. 4 graphically shows the above observations relative to the second loop of one of the participants. Red marks refer to max values, whereas the blue ones refer to the min values.

The correlation analysis shows the relationship between Speed, DMGas, RoadDist, and SC behavior of two different LODs (0,1,2,3). The combination between the different LODs generates 6 clusters of correlation pairs (LOD 0 and LOD 1, LOD 0 and LOD 2, LOD0 and LOD 3, LOD 1and LOD 2, LOD 1 and LOD 3, LOD 2 and LOD 3). The synthesis of outcomes for the whole analysis is reported in Table II whereas Fig. 5 shows the ANOVA box-plot graphs of correlations for each measure.

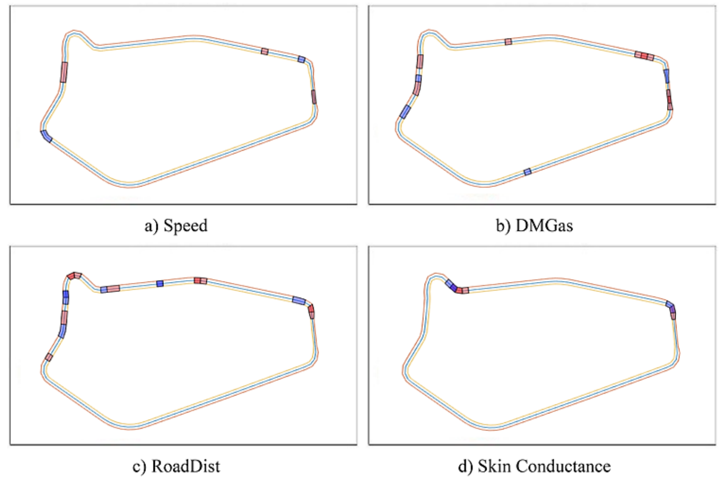


Fig. 4. Find spots relative to the second loop of one of the participants. Red marks refer to max values, whereas the blue ones refer to the min values. a) Speed, b) DMGas, c) RoadDist, and d) Skin Conductance analysis.

TABLE II

Average correlation values of each LOD pair, overall average value of correlations between LODs calculated by using the Fisher's Z formula, and number of min and max spots for each measure.

	Speed	***SC***	***DMGAS***	***RoadDist***
LOD 0 - LOD1	*0.680*	*0.351*	*0.418*	*0.779*
LOD 0 - LOD2	*0.692*	*0.356*	*0.406*	*0.779*
LOD 0 - LOD3	*0.670*	*0.341*	*0.394*	*0.753*
LOD 1 - LOD2	*0.705*	*0.430*	*0.499*	*0.822*
LOD 1 - LOD3	*0.706*	*0.413*	*0.427*	*0.755*
LOD 2 - LOD3	*0.716*	*0.342*	*0.451*	*0.766*
Average value	*0.715*	*0.384*	*0.454*	*0.752*
min spots	*1.34*	*1.19*	*4.98*	*8.70*
max spots	*2.02*	*1.75*	*4.70*	*6.09*

It can be observed that RoadDist and Speed are the most correlated between LODs, SC and DMGas the least, and consequently, it can be considered that the last two measures are more affected by LOD. The ANOVA confirms this last hypothesis (*p=0*). Fig. 5 shows how RoadDist and Speed have the highest correlation values and RoadDist also the smallest range (apart from a few outliers); on the opposite, DMGas and SC have the lowest correlation values. This can be explained by considering RoadDist and Speed more constrained by geometrical road features.

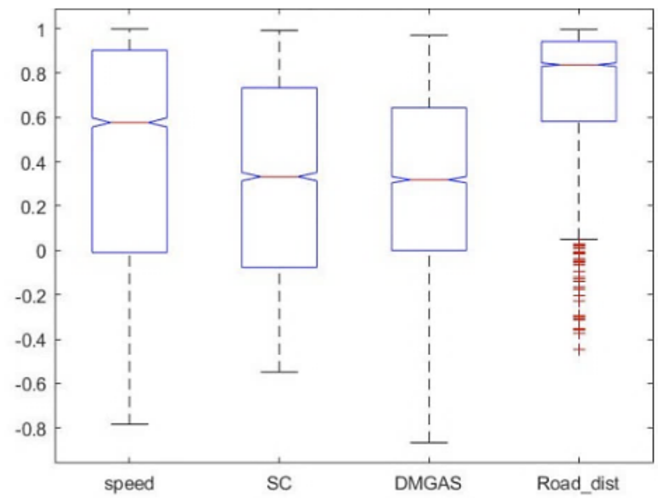


Fig. 5. ANOVA box-plot graphs of correlations for each measure.

B. Visual exploration behavior

1) Gaze position

A specific program was implemented to represent all the data relative to the gaze positions in a synthetic way. This program calculates the rectangular area of maximum frequency gaze position in an aggregated and segment-wise format for a driver. 65% of the most frequent data was taken as the essential data of gaze position. The coordinate of maximum frequency gives the location of the maximum area of visual interest for a certain road segment. It gives an idea about how driver visual interest mainly focuses during driving. It also calculates the area of maximum data frequency at an aggregated situation with the area of maximum frequent data at each road segment for all participants. The overlapping area between the maximum frequency of the aggregated segment and the segment-wise data was also calculated. The peak value of aggregated data and the peak value of segment-wise data were in the outcomes, and the difference of position between them was also measured. This program also checks whether the position of the peak value of the data lies inside or outside the respective rectangular area of the most frequent data (Fig. 6a, b).

A recurrent difference between data distributions is observed for rectilinear and curvilinear segments. The red rectangle is generally inside the blue one for the rectilinear segments (Fig. 6c) and outside for the curvilinear ones. In curvilinear segments, the position of the red square is always on the left side (Fig. 6d) since all curves of the path are on the left, except for the one in segment H-I. The exploration of these data does not highlight different behaviors among the 4 LODs. However, this type of analysis does not consider the breadth of gaze distribution and the presence of other local maxima which could occur along the road.

However, considering the data related to specific segments of the path, the distribution of the gaze has a different behavior, according to the LODs. By elaborating attentional maps through the eye-tracker software in the rectilinear segment E-F the gaze generally focused on the center of the road in LOD 0. In LOD 1 and LOD 2 there is a little focus dispersion due to the presence of the building that becomes more evident in LOD 3. In the curvilinear segment H-I, instead, the gaze was mostly allocated on the edges of the road, especially in LOD 1, LOD 2, and LOD 3, which coincide with the findings of [79]. In LOD 0, instead, participants were forced to change visual strategy to seek reference. The comparison of these results confirms the difference in driver behavior between rectilinear and curvilinear segments and how LOD differently conditions it. However, this different gaze behavior seems not to affect the vehicle trajectory. Fig. 7 shows the gaze attentional maps for the 2 segments for each LOD extracted from one of the participants.

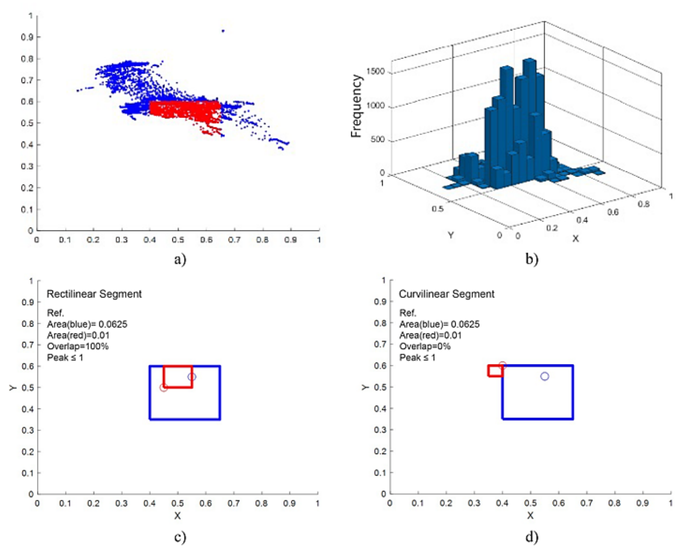


Fig. 6. Gaze data collected for a single driver along one loop; a) data distribution in blue and red; the red area contains 65% of data; b) the same as in a) but showing data frequency; c) localization of 65% data collected in a rectilinear segment (red rectangle) with respect to the 65% of overall data (blue rectangle); d) the same as in c) but for a curvilinear segment.

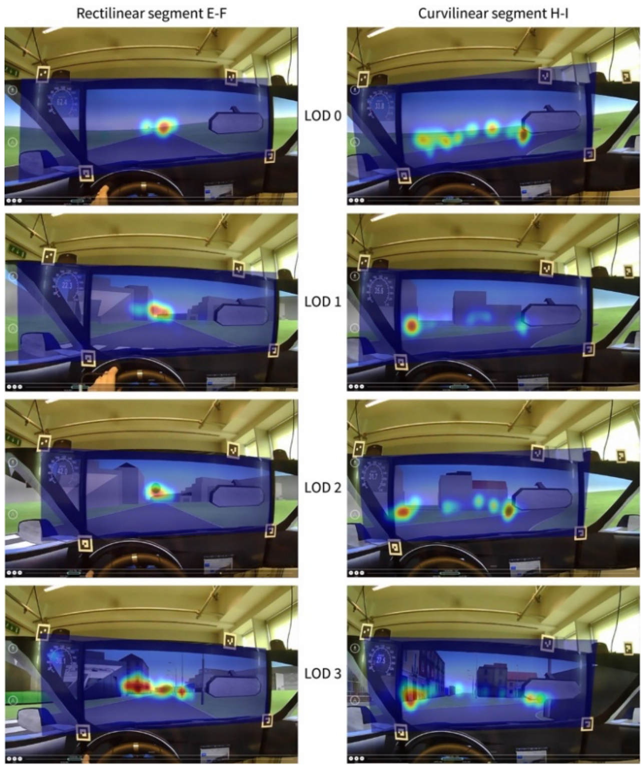
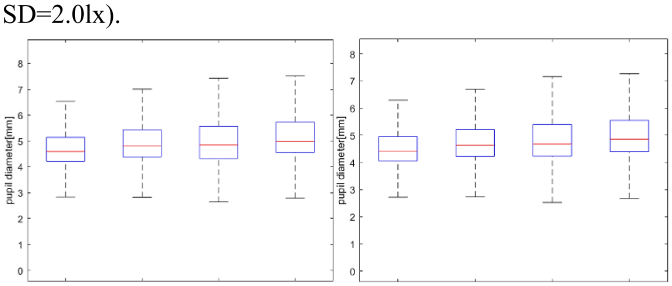


Fig. 7. Attention maps of a participant in different LODs.

2) Pupil diameter

Pupil diameter is usually considered a measure of the driver's visual attention. Considering the accuracy level of the eye-tracking device, a threshold of 80% on the confidence level was applied to filter the effective data. The calculation is done separately due to a subtle difference between human's two eyes diameters. In Fig. 8, the results of all participants' average pupil diameter are shown; a slight increase of both eyes from LOD 0 to LOD 3 can be observed ($p<<0.05$). This result suggests that the participants' visual attention increased with the increasing complexity of the scenarios. It is worth noting that this parameter could also be influenced by the ambient brightness, including the image brightness of the different LODs; without buildings, the portion of the sky within the scene is greater. As introduced before, in this study the ambient brightness is highly controlled; a quick check on the 4 LODs images has been conducted by applying the Minolta CHROMA METER CL-200A. The different brightness between the 4 LODs is very small (M=228.6lx, SD=2.0lx).

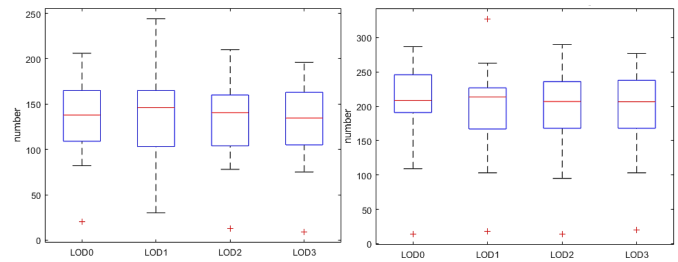


a) b)

Fig. 8. Pupil diameters of all participants in a) rectilinear segments and b) curvilinear segments by LOD.

3) Eye Fixations

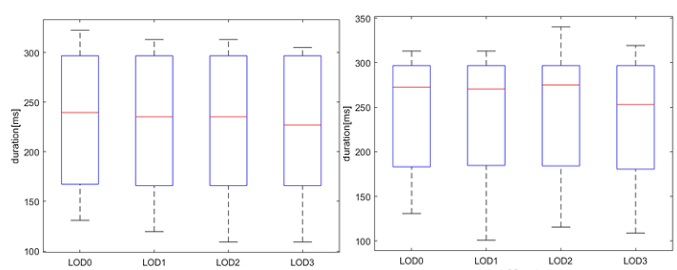
An eye fixation could be defined with different thresholds; as mentioned before, in this study, our constraints are set up based on various previous research: maximum dispersion of 1.0 deg, and duration from 140ms to 300ms. To calculate the number of fixations (NoF), the sampling period is set as the entire testing session. In Fig. 9, the result shows that, in general, the number of fixations in curvilinear segments is greater than that in rectilinear segments, still, the difference between the LODs is not statistically significant ($p >>0.05$).



a) b)

Fig. 9. NoF of all participants in rectilinear segments a) and curvilinear segments b) by LOD.

As introduced, the minimum eye fixation duration in this study is set as 140ms, and the maximum is set as 300ms, but there is a fluctuation in this interval, and the fixation duration most often reflects the fact that the brain is processing certain information. In this case, longer fixation duration values refer to an increase in the driver's attention at that screen spot. The average fixation duration in each LOD is shown in Fig. 10 in units of milliseconds. In general, the fixation duration was longer in the curvilinear segments than in the rectilinear segments, similar to the number of fixations. This result suggests that curvilinear segments require a higher mental workload. The decrease in the fixation duration in LOD 3 ($p<0.05$) indicates that the participants spent less mental effort on the driving task, i.e. with all elements depicted in the road environment (higher level of realism).



a) b)

Fig. 10. Fixation duration [ms] of all participants in rectilinear segments a) and curvilinear segments b) by LOD.

C. Subjective measures

The results of the subjective measures are presented in Table III. The first three rows of the table present the values of the three components of emotions. The average values of pleasure between the four LODs range between 3.58 and 3.72, corresponding to a moderately positive emotional state. The arousal values vary from 1.41 to 1.52, which is a strongly deactivating state. The value of dominance goes from 4.03 to 4.14, showing a high sense of autonomy experienced by the participants. The ANOVA shows no significant differences between the LODs for any of the emotional dimensions related to the driving experience.

The second three rows present the values of the items about the driving experience and the scenarios. In general, the driving experience is perceived as positive, as the average values range from 3.44 to 3.93. The driving experience is evaluated more positively than realism (average values from 3.24 to 3.55) and scenario adequacy (average values from 2.97 to 3.17). The ANOVA shows no significant effects of LODs are observed for the two items about realism and adequacy. The effect for the explicit evaluation of the driving experience is instead significant ($p<0.05$), and it is considered more positive as the LODs increase.

TABLE III
MEAN AND STANDARD DEVIATION BY LOD OF SUBJECTIVE MEASURES AND F STATISTICS ABOUT DIFFERENCE OF MEANS.

	LOD 0	**LOD 1**	**LOD 2**	**LOD 3**	**F (df), p**
Pleasure	M = 3.58 SD = 0.57	M = 3.79 SD = 0.62	M = 3.72 SD = 0.59	M = 3.59 SD = 0.63	0.86 (115), 0.47
Arousal	M = 1.52 SD = 0.63	M = 1.48 SD = 0.69	M = 1.41 SD = 0.68	M = 1.45 SD = 0.63	0.13 (115), 0.94
Dominance	M = 4.03 SD = 0.68	M = 4.14 SD = 0.74	M = 4.14 SD = 0.63	M = 4.10 SD = 0.72	0.14 (115), 0.93
How do you evaluate your driving experience?	M = 3.44 SD = 0.63	M = 3.79 SD = 0.62	M = 3.90 SD = 0.56	M = 3.93 SD = 0.59	3.90 (115), 0.0108
How do you evaluate the realistic level of the driving circumstances?	M = 3.24 SD = 0.69	M = 3.55 SD = 0.87	M = 3.41 SD = 0.98	M = 3.34 SD = 0.90	0.65 (115), 0.56
How do you evaluate the adequacy of the virtual scenario to perform the driving task?	M = 2.97 SD = 0.68	M = 3.00 SD = 0.70	M = 3.17 SD = 0.97	M = 3.17 SD = 1.00	0.49 (115), 0.69

VI. DISCUSSION

Despite the crucial effort by several pioneering authors, the scientific debate on the scenario's needed realism (in general and concerning 3D models' LOD) from an ontological perspective is still open; indeed, it is far from being a universal concept shared by the different fields dealing with simulations. Undoubtedly, the variables to consider are many, and moreover, they are also related to the technical tools and devices that continuously evolve. Nevertheless, what is generally agreed is that the simulation requirements highly depend on the specific goal of the investigation, which is in turn commonly related to and influenced by the disciplinary perspective. For instance, as emerged from the state of the art, in the architectural field a common goal of such investigation is to study people's experience in an actual or designed environment; to do so, the use of visual simulations with a high level of realism is generally adopted. The same high degree of scenario realism is generally not necessary in the mobility field focusing on driving performance. Rather, the simulator fidelity seems to play a major role in this case. This difference of outcomes relies on assessing driving as a 'task' and of driving as an 'environmental experience' in motion. This change of perspective is enough to shift the need from a low to a high LOD, a consideration confirmed by our research. As usual, correctly posing and framing the research question is crucial to choosing the proper tool and achieving reliable results. Undoubtedly, to build up an interdisciplinary ontological framework as an efficient reference for speeding up a proper simulation process, a shared effort of researchers is needed. Starting from the goal of the investigation might be a smart way to organize the framework.

Our research contributes to the analysis of drivers' reactions for in-vitro testing via simulations. The goal is to understand how to make the simulation process outcomes reliable, adopting solutions that are the most time and cost-effective as possible. The number of data collected for this study enabled us to assess people's reactions from different perspectives, i.e., by analyzing vehicular and physio/psychological data. By varying the LOD of the scenario during this experimental phase, the variety of collected data and their combined analysis enabled us to investigate the driver's emotional reactions, for each of the four LODs, during and after the assigned task. To note that the psychological fidelity of the simulation in relation to the actual context has been the objective of a previous study by the authors [18].

The data analysis highlights how some of them differ according to the LOD of the virtual scenario. Vehicular and Skin Conductance data, which represent human behavior in performing the driving task, highlight a slight difference according to the LODs. Vehicle trajectory (RoadDist) and speed data keep a high correlation value, i.e. they are barely affected by LODs, whereas Skin Conductance and the gas pedal activity do not, i.e. they are more affected by LODs. In addition, the FindSpot analysis showed how the differences in all measures are quite spread out along the road, and common trends are identified only in specific segments, e.g. curvilinear vs rectilinear segments. It is also worth noting that collected data intrinsically differ in their trends from measure to measure. Vehicle trajectory and speed smoothly change according to the road geometry, whereas the other data present different trends: Skin Conductance quickly increases and then decreases as a reaction to external events, such as a curve, whereas the gas pedal is continuously moved to keep the desired speed. Consequently, trajectory and speed data seem to be more affected by the road geometry than the LODs of the virtual scenario. Further investigations should clarify if these measures are affected by interacting effects of circulation, like pedestrians or other vehicles.

The data collected with the eye-tracker system show a similar mixed pattern. The gaze position data along the whole path are more sensitive to the road geometry than to the

LODs; the driver looks at the center of the road in the rectilinear segments, whereas the gaze anticipates the direction of the curves in the curvilinear segments. Despite this, the gaze position in curvilinear segments shows different trends across the LODs. In particular, in LOD0 the only reference to accomplish the driving task is the road; hence, the drivers look for more reference points within the environment to set the vehicle trajectory. By increasing the LOD the presence of the buildings' facades facilitates the trajectory setting, and consequently, the number of focus points is drastically reduced. Conversely, buildings' facades become a visual attractor along with the rectilinear segments, where the drivers tend to look also at them instead of solely focusing on the center of the road, as in LOD0. The analysis of the pupil diameters showed that higher LODs better engaged participants. Hence, visual attention increases with the increasing complexity of the scenario since pupil diameter is related to visual attention [69].

Likewise, the subjective assessment reveals a complex picture. The three components of emotions experienced while driving, namely pleasure, arousal, and dominance, do not vary across the LODs. In the same vein, explicit evaluation of the realism of the driving sensation and the adequacy of the scenario to perform the assigned task remain the same, regardless of the LODs. However, when asked to assess the overall driving experience, participants prefer higher LODs. These data suggest that the influence of LODs on subjective measures is not significant when investigating aspects more closely related to the driving task. A possible explanation for the significant effect of LODs on the general evaluation of the driving experience lies in its broader meaning. Unlike previous measures, it is less focused on a specific feature of the driving task and implies a more general reflection on the environmental interaction. Such a consideration relies on a comparison among different literature traditions investigating the person-environment relationship. On the one side, studies on restorative environments showed that increased realism in the same VR scenario elicited more positive emotions [57]. Other studies with a similar psychological approach did not specifically investigate the issue of realism, still, they demonstrated that 3D VR simulations can more positively impact affect than 360° videos of the same environment [80], [81]. A common trait of such studies, which is widely spread in psychology methods, concerns the instructions given to the participants, who are invited to observe the surroundings sitting or freely wandering in the area (when the experimental setting allows them in doing so). As the focus of the investigation is the general experience in the environment, no specific assignments are traditionally foreseen, and the participants loiter for a certain amount of time in the simulated environment. A similar methodological approach can be observed in urban planning when VR technologies are often used to present design solutions to citizens. In these circumstances, participants are generally invited to explore an area to get an idea of how it will look like. They generally have different degrees of autonomy (e.g. free exploration simulating a walk in the area, free exploration with digital shortcuts such as teleportation, pre-determined exploration paths with free movement of head or torso for visual examination), yet, the procedure asks to physically and visually explore the environment without providing specific predefined goals (e.g. [82], [83]). On the other side, most studies about driving simulation provide specific tasks requiring to drive from one point to the other, focusing on the driving performance rather than on the global environmental experience [84]. As the main goal in this field lies in investigating the actual driving behavior, the environment is conceived as a means for reaching a purpose: participants have an instrumental relationship with the simulated environment, which is in addition mediated by the simulated car [85]–[87]. We argue that such a difference in the experimental procedure, namely a leisurely versus a task-driven environmental interaction, and the object of investigation, that is, the environmental experience versus the task experience and the performance, might affect the importance of the scenario realism. According to the literature, scenario realism seems relevant when scholars investigate the environmental experience as a whole during carefree interactions, whilst its importance decreases when studying the performance or the training effectiveness related to a certain task carried out in the environment. In the latter case, a more relevant role of simulator fidelity over scenario realism is to be considered. This stance is consistent with studies on avatars, which showed how a higher degree of control on the movements of the avatar when performing physical tasks in a simulated environment is associated with an improved sense of embodiment [88], [89], that includes spatial, motor and affective factors [90]. Although the relationship that a person establishes with an avatar and with a simulated vehicle entails many differences, such studies pave the way for integrating the knowledge on this issue and deepening the relationship between scenario realism and simulator fidelity in affecting the subjective experience in simulated environments. This would offer new insights in understanding the role of simulations with different means of transportation and the relevance of including avatars in some circumstances.

VII. Conclusion

The current article investigates people's behavioral and emotional reactions during a driving task in a simulated urban environment, comparing four different LODs. The results showed that vehicle trajectory and speed appear to be more influenced by road geometry, whereas Skin Conductance and the gas pedal activity are more sensitive to LODs. Gaze position is differently affected by LODs in curvilinear segments, where higher LODs offer more reference points requiring less visual exploration, and rectilinear segments, where the richer environment calls for more exploration. Pupil diameter increases with higher LODs, suggesting increased attention. The number of fixations is not affected by LODs, whereas their duration only decreases for LOD3. Subjective evaluation of the emotions experienced while driving does not vary across the LODs. The experimental investigation suggests that the LODs' influence is lower than hypothesized for many variables, and in some cases, it arises according to the geometry of the road.

The comparison of these results with previous studies from the mobility field and other disciplinary sectors emphasizes that the importance of LOD cannot be considered *per se* but is related to the experimental design and the objective of the investigation. Depending on the specific simulation goals, the relevance of the scenario, simulator, or even avatar might play different roles. Such conclusions are a first attempt to connect different research traditions, which will require better deepening the interaction among the observed variables, for instance, including as experimental variable the simulator fidelity together with the scenario LODs. In addition, some limitations are worth mentioning, as they can affect the results. The study investigated a specific type of urban and social context, with a limited number of participants who are quite homogeneous in socio-demographical terms. In these regards, future works applying the same methodology should deal with different variables for the sake of direct comparison: different types of driving simulators (low and high quality); urban typo-morphological patterns (e.g. compact, historical) and landscape environments (e.g. rural, industrial) and relative roads' infrastructures; environmental conditions (e.g. sunny, foggy as well as the presence of people, cars, and so on); means of transportation (e.g. bike, walk); sample of drivers (e.g. people with a wider range of ages from young to older people). These types of variables, i.e. related to the context and its fruition, would contribute to a more exhaustive framework for informing the analysis via simulation of people's emotional and behavioral reactions in motion. This would be a useful reference for choosing the proper instruments out of the simulation toolkit according to the specific goal in order to achieve reliable scientific outcomes. Indeed, a reference for setting up or evaluating simulations is essential, even if specific decisions should be made according to the specific research goals, the contextual case with its peculiarities, and the available resources.

Authors' statement

Yuan Shi developed the first draft of the article. Since the research team and the contribution is highly interdisciplinary, all authors contributed equally to the whole research and article writing from their own specific disciplinary perspective, i.e. M. Boffi from the psychological perspective, G. Caruso and Yuan Shi from the mechanical engineering perspective, L. Mussone from the mobility engineering perspective, B. Piga from the urban and representation perspective.

References

[1] A. H. Jamson, N. Merat, O. M. J. Carsten, and F. C. H. Lai, "Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions," *Transp. Res. Part C Emerg. Technol.*, vol. 30, pp. 116–125, May 2013, doi: 10.1016/j.trc.2013.02.008.

[2] D. L. Fisher, *Handbook of driving simulation for engineering, medicine, and psychology*. Boca Raton: Taylor & Francis Group, 2011. Accessed: Oct. 07, 2021. [Online]. Available: http://0-marc.crcnetbase.com.fama.us.es/isbn/9781420061017

[3] M. Mulder, D. A. Abbink, and E. R. Boer, "The effect of haptic guidance on curve negotiation behavior of young, experienced drivers," in *2008 IEEE International Conference on Systems, Man and Cybernetics*, Singapore, Singapore, Oct. 2008, pp. 804–809. doi: 10.1109/ICSMC.2008.4811377.

[4] D. Ariansyah, G. Caruso, D. Ruscio, and M. Bordegoni, "Analysis of Autonomic Indexes on Drivers' Workload to Assess the Effect of Visual ADAS on User Experience and Driving Performance in Different Driving Conditions," *J. Comput. Inf. Sci. Eng.*, vol. 18, no. 3, p. 031007, Sep. 2018, doi: 10.1115/1.4039313.

[5] Y. Shi, W. Huang, F. Cheli, M. Bordegoni, and G. Caruso, "Do Autonomous Vehicle Driving Styles Affect User State?: A Preliminary Investigation," in *Volume 1: 39th Computers and Information in Engineering Conference*, Anaheim, California, USA, Aug. 2019, p. V001T02A084. doi: 10.1115/DETC2019-98186.

[6] Y. Shi, M. Bordegoni, and G. Caruso, "User Studies by Driving Simulators in the Era of Automated Vehicle," *Comput.-Aided Des. Appl.*, vol. 18, no. 1, pp. 211–226, May 2020, doi: 10.14733/cadaps.2021.211-226.

[7] Y. Shi, J. Maskani, G. Caruso, and M. Bordegoni, "Explore User Behaviour in Semi-autonomous Driving," *Proc. Des. Soc. Int. Conf. Eng. Des.*, vol. 1, no. 1, pp. 3871–3880, Jul. 2019, doi: 10.1017/dsi.2019.394.

[8] O. Shechtman, S. Classen, K. Awadzi, and W. Mann, "Comparison of Driving Errors Between On-the-Road and Simulated Driving Assessment: A Validation Study," *Traffic Inj. Prev.*, vol. 10, no. 4, pp. 379–385, Jul. 2009, doi: 10.1080/15389580902894989.

[9] B. E. Porter, *Handbook of traffic psychology*, 1st ed. London Waltham, MA: Academic Press, 2011.

[10] G. B. Sætren, J. P. Wigum, R. Robertsen, P. Bogfjellmo, and E. Suzen, "The future of driver training and driver instructor education in Norway with increasing ADAS technology in cars," in *Safety and Reliability – Safe Societies in a Changing World*, CRC Press, 2018. [Online]. Available: https://www.taylorfrancis.com/books/e/9781351174657/chapters/10.1201%2F9781351174664-181.

[11] A. B. Ellison, S. P. Greaves, and M. C. J. Bliemer, "Driver behaviour profiles for road safety analysis," *Accid. Anal. Prev.*, vol. 76, pp. 118–132, Mar. 2015, doi: 10.1016/j.aap.2015.01.009.

[12] B. Piga and E. Morello, "Environmental Design Studies on Perception and Simulation: an Urban Design Approach," *Int. J. Sens. Environ. Archit. Urban Space*, Sep. 2015, Accessed: Oct. 19, 2015. [Online]. Available: http://ambiances.revues.org/647

[13] M. H. Smolensky, L. Di Milia, M. M. Ohayon, and P. Philip, "Sleep disorders, medical conditions, and road accident risk," *Accid. Anal. Prev.*, vol. 43, no. 2, pp. 533–548, Mar. 2011, doi: 10.1016/j.aap.2009.12.004.

[14] H. L. Lew, J. H. Poole, E. H. Lee, D. L. Jaffe, H.-C. Huang, and E. Brodd, "Predictive validity of driving-simulator assessments following traumatic brain injury: a preliminary study," *Brain Inj.*, vol. 19, no. 3, pp. 177–188, Mar. 2005, doi: 10.1080/02699050400017171.

[15] P. Konstantopoulos, P. Chapman, and D. Crundall, "Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving," *Accid. Anal. Prev.*, vol. 42, no. 3, pp. 827–834, May 2010, doi: 10.1016/j.aap.2009.09.022.

[16] N. A. Kaptein, J. Theeuwes, and R. van der Horst, "Driving Simulator Validity: Some Considerations," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1550, no. 1, pp. 30–36, Jan. 1996, doi: 10.1177/0361198196155000105.

[17] C. J. Robbins, H. A. Allen, and P. Chapman, "Comparing drivers' visual attention at Junctions in Real and Simulated Environments," *Appl. Ergon.*, vol. 80, pp. 89–101, Oct. 2019, doi: 10.1016/j.apergo.2019.05.005.

[18] D. Ruscio *et al.*, "Collection and comparison of driver/passenger physiologic and behavioural data in simulation and on-road driving," in *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Naples, Italy, Jun. 2017, pp. 403–408. doi: 10.1109/MTITS.2017.8005705.

[19] Y. Wang, B. Mehler, B. Reimer, V. Lammers, L. A. D'Ambrosio, and J. F. Coughlin, "The validity of driving simulation for assessing

differences between in-vehicle informational interfaces: A comparison with field testing," *Ergonomics*, vol. 53, no. 3, pp. 404–420, Mar. 2010, doi: 10.1080/00140130903464358.

[20] S. T. Godley, T. J. Triggs, and B. N. Fildes, "Driving simulator validation for speed research," *Accid. Anal. Prev.*, vol. 34, no. 5, pp. 589–600, Sep. 2002, doi: 10.1016/S0001-4575(01)00056-2.

[21] J. D. Lee, "Simulator fidelity: How low can you go," 2004.

[22] G. Gröger, T. H. Kolbe, C. Nagel, and K.-H. Häfele, "OGC City Geography Markup Language (CityGML) Encoding Standard," 2012.

[23] M. J. Johnson, T. Chahal, A. Stinchcombe, N. Mullen, B. Weaver, and M. Bédard, "Physiological responses to simulated and on-road driving," *Int. J. Psychophysiol.*, vol. 81, no. 3, pp. 203–208, Sep. 2011, doi: 10.1016/j.ijpsycho.2011.06.012.

[24] J. K. Lenneman and R. W. Backs, "Cardiac Autonomic Control During Simulated Driving With a Concurrent Verbal Working Memory Task," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 51, no. 3, pp. 404–418, Jun. 2009, doi: 10.1177/0018720809337716.

[25] D. Liu, N. Macchiarella, and D. Vincenzi, "Simulation Fidelity," in *Human factors in simulation and training*, 2008, pp. 61–73. doi: 10.1201/9781420072846.ch4.

[26] D. Appleyard, "Understanding Professional Media," in *Human Behavior and Environment*, I. Altman and J. F. Wohlwill, Eds. New York: Plenum Press, 1977, pp. 43–88. Accessed: Feb. 03, 2014. [Online]. Available: http://link.springer.com/chapter/10.1007/978-1-4684-0808-9_2

[27] K. Mania, T. Troscianko, R. Hawkes, and A. Chalmers, "Fidelity Metrics for Virtual Environment Simulations Based on Spatial Memory Awareness States," *Presence*, vol. 12, no. 3, pp. 296–310, Jun. 2003, doi: 10.1162/105474603765879549.

[28] B. C. Schricker, R. W. Franceschini, and T. C. Johnson, "Fidelity evaluation framework," in *Proceedings. 34th Annual Simulation Symposium*, Seattle, WA, USA, 2001, pp. 109–116. doi: 10.1109/SIMSYM.2001.922122.

[29] Z. C. Roza, *Simulation fidelity theory and practice: a unified approach to defining, specifying and measuring the realism of simulations*. Delft: Delft University Press, 2004.

[30] A. J. Rehmann, R. D. Mitman, and M. C. Reynolds, "A Handbook of Flight Simulation Fidelity Requirements for Human Factors Research," Technical note ADA303799, 1995.

[31] C. D. Fink and E. L. Shriver, "Simulators for Maintenance Training: Some Issues, Problems and Areas for Future Research," KINTON INC ALEXANDRIA VA, Jul. 1978. Accessed: Sep. 27, 2021. [Online]. Available: https://apps.dtic.mil/sti/citations/ADA060088

[32] R. T. Hays, "Simulator Fidelity: A Concept Paper," U.S. Army Research Institute for the Behavioral and Social Sciences, 1980. [Online]. Available: http://archive.org/details/DTIC_ADA101475

[33] J. A. Allen, R. T. Hays, and L. C. Buffardi, "Maintenance Training Simulator Fidelity and Individual Differences in Transfer of Training," *Hum. Factors*, vol. 28, no. 5, pp. 497–509, Oct. 1986, doi: 10.1177/001872088602800501.

[34] J. A. Allen, L. Buffardi, and R. Hays, "The Relationship of Simulator Fidelity to Task and Performance Variables," U.S. Army Research Institute for the Behavioral and Social Sciences, 1991.

[35] J. J. Slob, "State-of-the-Art Driving Simulators, a Literature Survey," DCT report 107, 2008.

[36] R. G. Kinkade and G. R. Wheaton, "Human Engineering Guide to Equipment Design (Revised Edition).," H. P. Van Cott and R. G. Kinkade, Eds. American Institutes for Research Washington D C, 1972. Accessed: Sep. 27, 2021. [Online]. Available: https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/AD758339.xhtml

[37] G. E. McKechnie, "Simulation Techniques in Environmental Psychology," in *Perspectives on Environment and Behavior*, 1°., D. Stokols, Ed. New York, London: Plenum Press, 1977, pp. 169–189. Accessed: Jan. 21, 2014. [Online]. Available: http://link.springer.com/chapter/10.1007/978-1-4684-2277-1_7

[38] S. R. J. Sheppard, *Visual Simulation: A User Guide for Architects, Engineers, and Planners*. New York: Van Nostrand Reinhold, 1989.

[39] P. Bosselmann, *Representation of Places: Reality and Realism in City Design*. Berkeley, Los Angeles, London: University of California Press, 1998.

[40] R. Hall, "Algorithms for realistic image synthesis," in *Computer Graphics Techniques: Theory and Practice*, D. F. Rogers and R. A. Earnshaw, Eds. Springer Science & Business Media, 1990.

[41] P. C. Lai, K.-H. Kwong, and A. S. H. Mak, "Assessing the applicability and effectiveness of 3D visualisation in environmental impact assessment," *Environ. Plan. B Plan. Des.*, vol. 37, no. 2, pp. 221–233, 2010, doi: 10.1068/b34141.

[42] E. Lange, "The limits of realism: perceptions of virtual landscapes," *Landsc. Urban Plan.*, vol. 54, no. 1–4, pp. 163–182, May 2001, doi: 10.1016/S0169-2046(01)00134-7.

[43] M. P. Reed and P. A. Green, "Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task," *Ergonomics*, vol. 42, no. 8, pp. 1015–1037, Aug. 1999, doi: 10.1080/001401399185117.

[44] M. van Gisbergen, M. Kovacs, F. Campos, M. van der Heeft, and V. Vugts, "What We Don't Know. The Effect of Realism in Virtual Reality on Experience and Behaviour," in *Augmented Reality and Virtual Reality: The Power of AR and VR for Business*, M. C. tom Dieck and T. Jung, Eds. Cham: Springer International Publishing, 2019, pp. 45–57. doi: 10.1007/978-3-030-06246-0_4.

[45] T. A. Funkhouser and C. H. Séquin, "Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments," in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques - SIGGRAPH '93*, Not Known, 1993, pp. 247–254. doi: 10.1145/166117.166149.

[46] M. Usoh *et al.*, "Walking > walking-in-place > flying, in virtual environments," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99*, Not Known, 1999, pp. 359–364. doi: 10.1145/311535.311589.

[47] Insu Yu, J. Mortensen, P. Khanna, B. Spanlang, and M. Slater, "Visual Realism Enhances Realistic Response in an Immersive Virtual Environment - Part 2," *IEEE Comput. Graph. Appl.*, vol. 32, no. 6, pp. 36–45, Nov. 2012, doi: 10.1109/MCG.2012.121.

[48] P. Zimmons and A. Panter, "The influence of rendering quality on presence and task performance in a virtual environment," in *IEEE Virtual Reality, 2003. Proceedings.*, Los Angeles, CA, USA, 2003, pp. 293–294. doi: 10.1109/VR.2003.1191170.

[49] M. Slater, B. Spanlang, and D. Corominas, "Simulating virtual environments within virtual environments as the basis for a psychophysics of presence," *ACM Trans. Graph.*, vol. 29, no. 4, pp. 1–9, Jul. 2010, doi: 10.1145/1778765.1778829.

[50] K. Debattista, T. Bashford-Rogers, C. Harvey, B. Waterfield, and A. Chalmers, "Subjective Evaluation of High-Fidelity Virtual Environments for Driving Simulations," *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 1, pp. 30–40, Feb. 2018, doi: 10.1109/THMS.2017.2762632.

[51] F. Biljecki, H. Ledoux, J. Stoter, and J. Zhao, "Formalisation of the level of detail in 3D city modelling," *Comput. Environ. Urban Syst.*, vol. 48, pp. 1–15, Nov. 2014, doi: 10.1016/j.compenvurbsys.2014.05.004.

[52] A. Stadler and T. H. Kolbe, "Spatio-semantic coherence in the integration of 3D city models," 2007.

[53] R. T. Hays, "Simulator Fidelity: A Concept Paper," U.S. Army Research Institute for the Behavioral and Social Sciences, 1980. [Online]. Available: http://archive.org/details/DTIC_ADA101475

[54] M. K. Kaiser and J. A. Schroeder, "Flights of fancy: The art and science of flight simulation.," in *Principles and Practice of Aviation Psychology*, 2002, p. 435.

[55] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *J. Comput.-Mediat. Commun.*, vol. 3, no. 2, p. JCMC321, 1997.

[56] G. Riva, F. Davide, and W. A. IJsselsteijn, "Being there: The experience of presence in mediated environments," *There Concepts Eff. Meas. User Presence Synth. Environ.*, vol. 5, 2003.

[57] M. Newman, B. Gatersleben, K. J. Wyles, and E. Ratcliffe, "The use of virtual reality in environment experiences and the importance of realism," *J. Environ. Psychol.*, vol. 79, p. 101733, Feb. 2022, doi: 10.1016/j.jenvp.2021.101733.

[58] A. N. Stephens and J. A. Groeger, "Situational specificity of trait influences on drivers' evaluations and driving behaviour," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 12, no. 1, pp. 29–39, Jan. 2009, doi:

10.1016/j.trf.2008.06.005.

[59] A. N. Stephens and J. A. Groeger, “Do emotional appraisals of traffic situations influence driver behaviour?,” in *Behavioural research in road safety 2006: sixteenth seminar*, Great Britain and Department for Transport, Eds. London: Dept. for Transport, 2006, pp. 49–62.

[60] G. M. Hancock, P. A. Hancock, and C. M. Janelle, “The impact of emotions and predominant emotion regulation technique on driving performance,” *Work*, vol. 41, no. Supplement 1, pp. 3608–3611, Jan. 2012, doi: 10.3233/WOR-2012-0666-3608.

[61] S. De Nadai *et al.*, “Enhancing safety of transport by road by on-line monitoring of driver emotions,” in *2016 11th System of Systems Engineering Conference (SoSE)*, Jun. 2016, pp. 1–4. doi: 10.1109/SYSOSE.2016.7542941.

[62] “IPG CarMaker.” https://ipg-automotive.com/ (accessed Feb. 09, 2022).

[63] “ProComp Infiniti System.” https://thoughttechnology.com/ (accessed Feb. 09, 2022).

[64] “Pupil Labs Core.” https://pupil-labs.com/products/core/ (accessed Feb. 09, 2022).

[65] M. A. Perez, Z. R. Doerzaph, C. K. Gaylord, and J. M. Hankey, “Extracting information from continuous naturalistic driving data: sample applications,” in *2010 IEEE Intelligent Vehicles Symposium*, Jun. 2010, pp. 1201–1208. doi: 10.1109/IVS.2010.5547964.

[66] G. Marquart, C. Cabrall, and J. de Winter, “Review of Eye-related Measures of Drivers’ Mental Workload,” *Procedia Manuf.*, vol. 3, pp. 2854–2861, Jan. 2015, doi: 10.1016/j.promfg.2015.07.783.

[67] S. Chandra, G. Sharma, S. Malhotra, D. Jha, and A. P. Mittal, “Eye tracking based human computer interaction: Applications and their uses,” in *2015 International Conference on Man and Machine Interfacing (MAMI)*, Dec. 2015, pp. 1–5. doi: 10.1109/MAMI.2015.7456615.

[68] T. Čegovnik, K. Stojmenova, G. Jakus, and J. Sodnik, “An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers,” *Appl. Ergon.*, vol. 68, pp. 1–11, 2018.

[69] B. Reimer, B. Mehler, Y. Wang, and J. F. Coughlin, “The Impact of Systematic Variation of Cognitive Demand on Drivers’ Visual Attention across Multiple Age Groups,” *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 54, no. 24, pp. 2052–2055, 2010, doi: 10.1177/154193121005402407.

[70] M. M. Bradley and P. J. Lang, “Measuring emotion: The self-assessment manikin and the semantic differential,” *J. Behav. Ther. Exp. Psychiatry*, vol. 25, no. 1, pp. 49–59, Mar. 1994, doi: 10.1016/0005-7916(94)90063-9.

[71] R. B. Rubin, A. M. Rubin, E. E. Graham, E. M. Perse, and D. Seibold, *Communication Research Measures II: A Sourcebook*. Routledge, 2010.

[72] T.-M. Bynion and M. T. Feldner, “Self-Assessment Manikin,” in *Encyclopedia of Personality and Individual Differences*, V. Zeigler-Hill and T. K. Shackelford, Eds. Cham: Springer International Publishing, 2020, pp. 4654–4656. doi: 10.1007/978-3-319-24612-3_77.

[73] A. Mehrabian and J. A. Russell, *An approach to environmental psychology.* the MIT Press, 1974.

[74] I. Olkin and J. W. Pratt, “Unbiased Estimation of Certain Correlation Coefficients,” *Ann. Math. Stat.*, vol. 29, no. 1, pp. 201–211, Mar. 1958, doi: 10.1214/aoms/1177706717.

[75] D. D. Salvucci and J. H. Goldberg, “Identifying fixations and saccades in eye-tracking protocols,” in *Proceedings of the 2000 symposium on Eye tracking research & applications*, New York, NY, USA, Nov. 2000, pp. 71–78. doi: 10.1145/355017.355028.

[76] “ISO15007 Road vehicles — Measurement of driver visual behaviour with respect to transport information and control systems.” 2014.

[77] A. Borowsky, D. Shinar, and T. Oron-Gilad, “Age, skill, and hazard perception in driving,” *Accid. Anal. Prev.*, vol. 42, no. 4, pp. 1240–1249, Jul. 2010, doi: 10.1016/j.aap.2010.02.001.

[78] B. M. Velichkovsky, A. Rothert, M. Kopf, S. M. Dornhöfer, and M. Joos, “Towards an express-diagnostics for level of processing and hazard perception,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 5, no. 2, pp. 145–156, 2002, doi: 10.1016/S1369-8478(02)00013-X.

[79] Y. Ren, X. Li, X. Zheng, Z. Li, and Q. Zhao, “Analysis of Drivers’ Eye-Movement Characteristics When Driving around Curves,” *Discrete Dyn. Nat. Soc.*, vol. 2015, p. e462792, Mar. 2015, doi: 10.1155/2015/462792.

[80] T. Nukarinen *et al.*, “Physiological and Psychological Restoration in Matched Real and Virtual Natural Environments,” in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, Apr. 2020, pp. 1–8. doi: 10.1145/3334480.3382956.

[81] N. L. Yeo *et al.*, “What is the best way of delivering virtual nature for improving mood? An experimental comparison of high definition TV, 360° video, and computer generated virtual reality,” *J. Environ. Psychol.*, vol. 72, p. 101500, Dec. 2020, doi: 10.1016/j.jenvp.2020.101500.

[82] D. Edler, J. Keil, T. Wiedenlübbert, M. Sossna, O. Kühne, and F. Dickmann, “Immersive VR Experience of Redeveloped Post-industrial Sites: The Example of ‘Zeche Holland’ in Bochum-Wattenscheid,” *KN - J. Cartogr. Geogr. Inf.*, vol. 69, no. 4, pp. 267–284, Dec. 2019, doi: 10.1007/s42489-019-00030-2.

[83] M. Loyola, B. Rossi, C. Montiel, and M. Daiber, “Use of Virtual Reality in Participatory Design,” in *Blucher Design Proceedings*, Porto, Portugal, Dec. 2019, pp. 449–454. doi: 10.5151/proceedings-ecaadesigradi2019_156.

[84] J. de Winter, P. Wieringa, J. Dankelman, M. Mulder, M. M. Paassen, and S. D. Groot, “Driving simulator fidelity and training effectiveness,” presented at the Proceedings of the 26th European annual conference on human decision making and manual control, Lyngby, Denmark, 2007. Accessed: Dec. 23, 2021. [Online]. Available: https://www.semanticscholar.org/paper/Driving-simulator-fidelity-and-training-Winter-Wieringa/b2f756d5af8cd7ef5462da623f5239d606470f2c

[85] R. W. Allen, G. D. Park, M. L. Cook, and D. Fiorentino, “The Effect of Driving Simulator Fidelity on Training Effectiveness,” *N. Am.*, p. 15, 2007.

[86] H. Cai, Y. Lin, and R. Mourant, “Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators,” *DSC N. Am. – Iowa City – Sept. N. Am. – Iowa City – Sept.*, Jan. 2007.

[87] H. Tu, Z. Li, Y. Sun, and X. Dai, “Driving Behavior in Case of Haze: A High Fidelity Driving Simulator Study,” pp. 2408–2421, Jul. 2014, doi: 10.1061/9780784413623.230.

[88] R. Fribourg, F. Argelaguet, A. Lécuyer, and L. Hoyet, “Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View,” *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 5, pp. 2062–2072, May 2020, doi: 10.1109/TVCG.2020.2973077.

[89] R. Fribourg *et al.*, “Virtual Co-Embodiment: Evaluation of the Sense of Agency while Sharing the Control of a Virtual Body among Two Individuals,” *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 10, pp. 4023–4038, Oct. 2021, doi: 10.1109/TVCG.2020.2999197.

[90] F. de Vignemont, “Embodiment, ownership and disownership,” *Conscious. Cogn.*, vol. 20, no. 1, pp. 82–93, Mar. 2011, doi: 10.1016/j.concog.2010.09.004.