Università degli Studi di Padova

Department of General Psychology

Ph.D. Course in Psycological Sciences

XXXII Series

A tool for training hazard perception and for assessing driving behaviors in adolescents and inexperienced drivers: A simulation study

Coordinator: Prof. Giovanni Galfano

Supervisor: Prof. Michela Sarlo

Co-Supervisor: Dr. Mariaelena Tagliabue

Ph.D student : Evelyn Gianfranchi

A Berny, che mi ha guidata.

E a Federico, che mi sostiene.

Contents

Abstract

All over the world, young and inexperienced road users are overrepresented among the victims of crashes. They are reported to have a worse hazard perception, being at the same time more prone to risk-taking behaviors. Moreover, various patterns of personality traits seem to affect their on-road behaviors. During the years, lots of efforts have been devoted to the identification of strategies to cope with such a phenomenon, usually employing psychosocial interventions and frontal lessons. Meanwhile, the progress in technology led to the development of driving simulators, allowing to test driving behaviors in a safe environment. However, to date no specific training or assessment driving simulator protocols have yet been developed.

The present Ph.D. project aimed at developing and testing a hazard perception training and a driving profile assessment protocol for inexperienced (young and adolescents) road users by mean of a specific driving simulator, psychophysiological indices, and psychological measures. The development of the training and of the assessment protocol followed several steps, sometimes running parallel to each other. The features and issues of the inexperienced road users are discussed in Chapter 1, where a distinction between experience and exposure is made and their links with on-road behaviors are explored. In the second part of Chapter 1, the pros and cons of driving simulators are described; Moreover, it is introduced the Honda Riding Trainer - HRT simulator (employed in this Ph.D. project), followed by an overview of the previous studies that have previously used this simulator. In Chapter 2 the most important factors influencing inexperienced road users' behaviors (hazard perception, attention, and personality traits) are introduced and discussed.

The research work can be divided into two main branches: The first (Chapter 3) includes three studies. Specifically, the first two constituted the two parts of a single research project aimed at assessing changes in the skin conductance (SC) considered as a somatic marker of hazard perception in novice drivers during a two and a three-session training on the HRT, respectively. In the case of a two-session training, the main results show that the reduction of the crashes while

driving the simulator (when participants faced again the same courses during the second session) is paralleled by a significant anticipation in the onset of skin conductance responses (SCRs) in the hazardous scenes included in the courses, testifying that participants learned to anticipate the hazards that they had already experienced. In the subsequent study, three sessions were employed (the first two with the same courses and the last with different ones) and two groups compared: The training group drove the simulator whereas the control group just had to watch the video-clips of the same courses and detect hazards. Here, the reduction in risky driving behaviors and the increase in the safe ones in the training group were paralleled by significant changes in the SC in terms of percentages of SCRs, but no anticipation of the SCR onset was present in the third session. The last study included in Chapter 3 aimed at assessing the event-related potentials (ERPs) of attentional monitoring and shifting in a one-session training on the HRT, during which participants were administered with a multi-feature oddball paradigm. The deviant stimuli in the oddball were divided into three categories: Deviant pure tones, human sounds and traffic-related sounds. The main results showed that, at the end of the training, the participants avoided a higher percentage of accidents and had enough cognitive resources available to monitor the environment and to shift attention toward salient stimuli (traffic-related sounds), as attested by the increase in the amplitude of the P3a.

The second branch of the research work (Chapters 4 and 5) includes four studies: In the first two (Chapter 4) an assessment method to identify driving profiles by means of 18 driving indices extracted from the HRT is proposed. In the first study reported in Chapter 4, two profiles emerged and were compared in terms of self-reported aberrant driving behaviors, whereas, in the second study, three profiles were identified and compared in terms of sensation seeking and decision-making ability. In the last two studies (Chapter 5), the method previously proposed was employed to identify three "embryonic" driving profiles among adolescents without any license, also finding personality patterns that can predict the inclusion in the driving profiles; Overall, the results of this second branch indicated that the HRT can be employed as a tool to assess different driving profiles among low-experienced and totally inexperienced drivers and that these profiles show consistent links with the personality traits that are usually reported as predictors of risky driving behaviors. Finally, a first evaluation of the effectiveness of the training on the simulator in terms of on-road accidents and near misses (follow up) is reported, by means of an exploratory comparison of the self-reported real road performance of a trained and an untrained group of adolescents. The follow up study showed that, in a 12-months post-training period, participants reported less frequent bike crashes independently of the driving profile in which they were included. Moreover, an exploratory analysis showed that a subgroup of the trained participants, compared to an untrained control group, matched for age, gender and on-road exposure, reported less frequent bike crashes.

Overall, the results of the present Ph.D. project suggest that: (1) a training of at least two sessions on the HRT simulator leads to a reduction in the risky and to an increase in the safe driving behaviors, in novice and inexperienced road users; (2) these changes are supposed to be due to an improvement in participants' hazard perception, as attested by the psychophysiological changes in the participants who drove the simulator; (3) a one-session training on the HRT allows to have enough available resources to monitor the environment and to shift attention toward salient stimuli; (4) the behavioral improvements in the simulator seem to generalize to some degree to the real world, as attested by the reduction in the bike crashes for the trained participants and by the comparison of their crash rate with that of a matched control group; (5) the HRT can be employed as a tool to identify different driving profiles in people with different degrees of road exposure.

The training and the assessment protocol presented in this research work can be considered useful tools to cope with the inexperienced road users crash rate issue, from a preventive perspective. The possibility, provided by the simulator, to test and train drivers in a safe and simulated environment, may lead to the development of *ad hoc* protocols, even customized on the basis of the participant's driving profile and personality traits. As a result, this opportunity may provide a crucial contribution in preventing road crashes.

Chapter 1

The inexperienced road users: Features, issues and possible solutions

1.1 The inexperienced road users' problem

Although recent data (European Road Safety Observatory – ERSO, 2018) show a percentage decrease between 2006 and 2017 in the number of crashes for each vehicle category in the European Union (EU), the year-to-year decrease in the fatalities rate for the so-called vulnerable road users (*i.e.*, pedestrians, bicycles, moped- and motorcycle-drivers) in the same period is the least prominent (ERSO, 2018). Moreover, 15.4% of the on-road deaths in EU in 2016 is reported among people aged between 15 and 24 years old: Moped or motorcycle drivers and passengers represent about a third of these deaths (ERSO, 2018). Concerning Italy (Istituto Nazionale di Statistica –ISTAT, 2018), although the data regarding the vulnerable road users resemble the EU situation, in 2017 an increase of the number of deaths is present only among motorcyclists (+ 11.9%) and pedestrians (+ 5.3%), with unchanging rates for the other vulnerable categories. Moreover, although the number of on-road deaths of young and adolescents did not change compared to 2016, the highest amount of injures is reached by people between 20 and 29 years old, with an increase of 2% in the 15-19 age range (ISTAT, 2018). Overall, these data suggest that young and inexperienced road users are frequently involved in road crashes and that the most at-risk age range is between 15 and 29 years old. The data also highlight the issues of the vulnerable road users, who are physically more exposed to risks.

Therefore, the next question is "Why are young people more exposed to risky on-road situations?". Many efforts have been devoted by the researchers to understand the causes of this phenomenon. For example, McKenna (2012) stated that both developmental factors and experience must be considered when assessing the issue of young drivers' on-road fatality. From a developmental point of view, the high crash rate of young people may derive from a combination of increased reward sensitivity, low impulse control, higher levels of antisocial tendencies, and changes in the sleep rhythm (McKenna, 2012).

However, these factors may be crucial when we consider the "problem young driver" (*i.e.*, the most problematic and most at-risk sub-samples of young drivers; Scott-Parker, Watson, King, & Hyde, 2013; p. 144) and less important for the "young driver problem" (*i.e.*, the young drivers meant as a cohort that represents *per se* an issue for road safety; Scott-Parker *et al*., 2013; p. 144). To illustrate, when we talk about the young drivers' issues in general, we need to identify common characteristics that may explain, at least in part, the high crash rates of this category. For sure, on-road (in)experience is a common factor that need to be considered when assessing the young driver problem.

1.2 On-road experience or exposure?

Although crucial for the road-safety research, there is no common agreement on the definition of driving experience. Usually, driving experience is measured in terms of years since the obtainment of a driving/riding license. Carrol (1971) proposed a distinction between the concepts of driving experience and exposure. Specifically, driving exposure, differently from experience, is defined as the frequency with which the drivers face events that can produce a risk of accident. Thus, the best measure for driving experience is the driving mileage, considered as proportional to the frequency of risky traffic events (Carrol, 1971).

However, as pointed out also by Wolfe (1982), the expression "driving" exposure totally ignores pedestrians' and bicyclists' exposure to accidents. Thus, when measuring driving experience and exposure it is important to consider not only the data regarding motor vehicles, but also those regarding at least the use of the bicycle. Therefore, the terms "on-road" experience or exposure are to be preferred over "driving" experience or exposure. Moreover, since experience and exposure are two different concepts, it is necessary to consider both these variables to obtain unambiguous data regarding their influence on driving behaviors.

On-road exposure plays a crucial role especially in the case of inexperienced road users. As pointed out by McKenna (2012), many evidences show that the crash rate of young drivers significantly decreases over the first few months of driving. For example, Mayhew, Simpson and Pak (2003) demonstrated that crash rate dramatically decreases during the first 6 months of driving, and that for certain types of crashes (*e.g.*, night crashes or run-off-the-road crashes) the decrease is more rapid. These data would indicate that the exposure to traffic events play a more substantial role in influencing driving behaviors than experience. Unfortunately, this also indicates that young and inexperienced drivers are facing a paradox: They need to increase their exposure to risky traffic configurations to reduce crash risk, yet in the process they are exposed to a high risk of crash (McKenna, 2012).

1.3 The relationship between on-road experience/exposure and driving behaviors

How can on-road experience/exposure concretely influence driving behaviors? As we will see, there are two main reasons why a low level of on-road experience and exposure influence driving behaviors. First, many evidence (for a brief review see Deery, 1999) showed that inexperienced drivers have low hazard perception abilities, that is they are less able in detecting and dealing with traffic hazards. Second, on-road experience and exposure seem to affect the amount of distraction that can be tolerated (Crundall, Underwood, & Chapman, 2002; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003).

Concerning the relation between driving experience and hazard perception, Crundall (2016) demonstrated that novice and experienced drivers can be discriminated on the basis of hazard prediction (*i.e.*, the ability to predict an incoming hazard). In a series of experiments, participants were administered with video-clips showing on-road scenes characterized by different degrees of risk; Participants had to spot for hazards that could vary in terms of either source, timing, or type of their clues, *i.e.*, hazard "precursors" (Crundall, 2016; p. 49). Across all the studies, experienced drivers always performed better than novice drivers, showing higher accuracy in spotting hazards (in terms of source of the hazard, localization, and prediction of the consequences) and proving that hazard perception is modulated by different degrees of on-road experience.

As regards attention and distraction, Crundall *et al.* (2002) demonstrated that experienced drivers show a peculiar attentional pattern of gaze behavior, that seems to indicate an advantageous strategy of attention allocation to hazard position in peripheral parts of the visual field, whereas inexperienced drivers hardly allocate attention in extra-foveal regions. In line with this evidence, Underwood *et al.* (2003) found that novice drivers seem less able than experienced drivers in detecting peripheral stimuli under demanding conditions. Moreover, Tagliabue, Da Pos, Spoto, and Vidotto (2013) showed similar results during a simulated moped driving task, proving that, after a learning phase, participants improved their ability to detect peripheral stimuli.

Finally, relations between driving experience/exposure and aberrant driving behaviors (*i.e.*, violations, errors, and lapses) have been reported by different studies (Parker, Reason, Manstead, & Stradling, 1995; Reason, Manstead, Stradling, Baxter, & Campbell, 1999). Interestingly, it seems that a high frequency of on-road violations is predicted by a higher annual mileage (Parker *et al.*, 1995; Reason *et al.*, 1990). This effect is probably due to the fact that an increase in the mileage leads to an increase in the chances to commit violations and, at the same time, it may lead to an overconfidence that, in turn, may make people more prone to violate road rules. Moreover, the highest number of violations was reported by novice drivers, specifically by people aged between 17 and 20 years old (Parker *et al.*, 1995), confirming the importance of experience and exposure in influencing driving behaviors.

Overall, we can state that on-road experience and exposure must be considered in the road-safety research, particularly when the focus of the research is the so-called young driver problem. As seen, one of the causes of the high crash rate among young drivers is the inexperience, meant not only as the time past since the driving license obtainment, but also as the on-road exposure to risky traffic configurations. Low levels of driving experience and exposure are linked to less efficient hazard perception ability and attention allocation that, in turn, may lead to an increased crash rate. On the other hand, when it comes to the assessment of aberrant on-road behaviors, it seems that exposure and experience influence in an opposite way the frequency of violations.

As pointed out by McKenna (2012), a way to reduce crash probability is experiencing risky situations. However, risky situations, in turns, expose to crash risk. Therefore, researchers and organizations tried to solve this paradox developing methods to gain experience at low risk. Among them, the extension of the learning period and the introduction of the so-called graduate-licensing programs seem to have reached good results in reducing novice drivers' crash rate (McKenna, 2012). However, these countermeasures do not consider the inter-individual variability and a wide range of factors that can influence inexperienced road users' behavior. These factors are usually considered when assessing the abovementioned "problem young driver" (Scott-Parker *et al.*, 2013) and they will be discussed in the next chapter.

1.4 Virtual reality for driving training: Definitions

Another important tool that can be employed to develop driving training programs suitable for inexperienced road users is virtual reality (VR). As seen, experiencing hazardous situations is an important way to reduce crash likelihood among people with low or no on-road exposure. The reconstruction of virtual hazardous scenes should allow the users to experience risky situations without the need to be physically exposed to them.

However, since the definition of VR is not univocal, one could wonder whether a specific VR tool might be mostly suitable for driving training. Indeed, a lot of research has developed around virtual reality tools since 1965, the year in which a first glimpse of what VR could represent was given by Ivan Sutherland. At that time, VR was generally defined as a virtual world that looks, sounds, and feels real, and responds realistically to the viewer's actions (Sutherland, 1965). However, to date, no general agreement over the specific definition of "virtual reality" or "virtual environment" (usually considered as synonyms) has been reached.

All the definitions of VR that have spread since the nineties are based on some common elements (Mazuryk & Gervautz, 1996): Interaction (possibly immersive), and a simulated (autonomous) world that the user can experience through different devices. Thus, it seems that two key concepts characterize VR: The first is the interaction, and the second is the number of devices and sensory modalities through which the virtual world is presented. Both these factors concur in determining the "illusion of immersion".

Clearly, different levels of immersion and involvement can be recognized, ranging from the so-called Desktop VR (*i.e.*, the use of a conventional monitor to display the image of the world, that can vary depending on the user's behavior), up to the most complex and immersive systems, where the user is totally involved in a computer-generated world with a stereoscopic view of the scene, that varies as a function of the user's position and orientation. These immersive systems may be enhanced also by auditory, haptic, and sensory interfaces (Mazuryk & Gervautz, 1996). Depending on the purpose of the specific VR tool (*e.g.*, training, entertainment, professional purposes), sometimes users can also interact with the virtual environment by using standard input devices, such as a keyboard, a mouse, or a controller (Gandhi & Patel, 2018).

1.5 Driving simulators: Pros and cons

Undoubtedly, driving simulators are one of the most employed VR tools for research purposes. Driving simulators are usually distinguished in low-level, mid-level, and high-level categories, depending on the degree of immersion that they provide (Kaptein, Theeuwes, & Van Der Horst, 1996). Low-level simulators typically consist of a PC or a graphic work station, a monitor, and simple controls. Mid-level simulators include advanced imaging techniques, realistic controls, and possibly a simple motion system. High-level simulators are those usually employed to train specific professional figures (*e.g.*, pilots), providing the most realistic experience of driving.

As stated by Kaptein *et al.*, 1996, given the wide range of characteristics that driving simulators can have, any use should be preceded by questioning whether the simulator is sufficiently valid for the task or ability investigated. The validity is not directly linked to the degree of involvement and immersion that a simulator can provide. Because people rarely need all the range of available information to correctly perform a task, it is generally not necessary that the environment reconstructed by a simulator is identical to what would be in a real vehicle (Kaptein *et al.*, 1996). Nevertheless, different types of validity of driving simulators are normally considered and defined (Kaptein *et al.*, 1996), such as absolute validity (*i.e.*, if the absolute size of the effect is comparable to that found in reality), relative validity (*i.e.*, if the direction or relative size of the effect of the measure is the same in the simulator and in reality) and, more frequently, internal validity (*i.e.*, the identification of a possible relation between a manipulation and an effect on the driving behavior in the simulator) and external validity (*i.e.*, the extent to which the results obtained in a specific environment during a specific period of time can be generalized to other persons, environments, and times).

Beyond the aspects linked to their validity, simulators have a series of advantages (de Winter, van Leeuwen, & Happee, 2012) that make them very useful for the road safety research. For instance, they allow controllability, reproducibility, and standardization of the administered scenarios and stimuli (in terms of weather conditions, road layout, and traffic configurations). Moreover, driving simulators can produce accurate measurement data of driving performance, also in risky driving conditions that would be impossible to test in the real world. They can be also easily employed to train and to assess inexperienced drivers, allowing to face risky traffic configurations in a safe environment. Finally, they offer the opportunity for feedback (*e.g*., replay) over a wide range of situations.

Nevertheless, some disadvantages must be considered when driving simulators are employed for research and training (de Winter *et al*., 2012). For instance, low-level simulators have usually a limited physical and perceptual fidelity and therefore it is important to test generalizability of collected data to the real world. On the other hand, high-immersion simulators may cause motion sickness, leading to a reduction in the reliability of the recorded data. Moreover, only few studies have specifically investigated the validity and generalizability to the real world of the effects found in simulated environments. For instance, Shechtman, Classen, Awadzi, and Mann (2009) reported no differences between the rate and type of errors made by experienced participants when turning right or left at a crossroad, as measured by various parameters such as lane maintenance and speed regulation. On the other hand, Underwood, Crundall, and Chapman (2011) compared hazard detection in three experimental conditions, *i.e*., while driving on real roads, watching video-clips recorded from a real vehicle, and driving in a simulator. Their results indicate that, in all three conditions, professional drivers show a peculiar pattern of visual scanning and early eye fixations on hazardous road objects. This comparability between the results in the three conditions led the authors to consider the simulators as valid tools for driving abilities training and testing.

Taken together, these considerations suggest that, first, simulators can be useful tools to train and to assess driving behaviors because of their characteristics, such as standardization and replicability of conditions, accuracy in the measurement, opportunity for feedback and for facing risky situations in a safe context. Second, since high level immersion is not crucial to obtain good validity and generalizability, as attested by Kaptein *et al*. (1996), who obtained good results in terms of both absolute and relative validity in a mid-level simulator validation study, low- and mid-level simulators can be employed for research and training. Third, at least some aspects of validity and generalizability of the effects need to be considered when a driving simulator is employed, independently of its degree of immersion.

1.6 The Honda Riding Trainer simulator: Features and previous studies

The Honda Riding Trainer (HRT) simulator employed in this Ph.D. project is a non-immersive, mid-level driving simulator, developed with the aim to train driving abilities in novice drivers. It includes a Pentium 4 PC with a Windows XP operating system and an LCD monitor (1924 × 768) placed on a base connected to a chassis that is equipped with moped-like controls (Figure 1).

The controls allow the participants to ride along a range of virtual courses. Two speakers are placed next to the monitor: Through them, the acoustic effects of the engine and the traffic are reproduced, and instructions are given about the path to follow in the simulated environment. The users can choose between three different engine sizes (moped, average, and large engine size), three conditions (daylight, night, and fog) and three types of courses (six main roads, five secondary roads, and five touristic roads), that can be preceded by an exercitation course (without traffic) to familiarize with the environment and the controls. Moreover, the users can choose between automatic or manual transmission.

Each virtual course of the HRT includes 7 or 8 hazardous scenes (Figure 1) *i.e.*, reconstruction of the most common potentially hazardous on-road situations, as catalogued by the Motorcycle Accidents in Depth Study (MAIDS, 2004). The users must drive along the courses trying to avoid crashes and to comply with traffic rules. When a crash occurs, it is possible to visualize a replay of the crash circumstances and, at the end of each course, it is possible to watch a replay of the whole course.

The simulator automatically records a wide range of driving parameters (*e.g.*, speed, horizontal position, crashes, use of the brakes) with a 30-Hz sampling frequency. These parameters are employed by the simulator itself to calculate a letter score for the performance in each hazardous scene, depending on how well the user has prevented a crash. Specifically, the scores can be A (safe performance), B (almost safe), C (near miss), and D (crash).

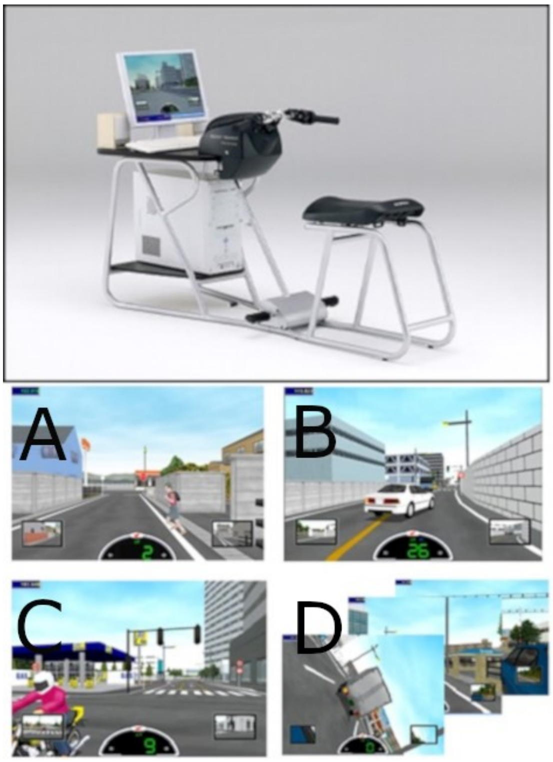

Figure 1: The HRT simulator (top panel) and some examples of HRT hazardous scenes classified by the simulator itself on the basis of participants' performance, reflecting its risk degree, namely A (totally safe), B (almost safe), C (near miss), and D (crash). Adapted Gianfranchi, Tagliabue, & Vidotto (2018).

A wide number of studies have employed the HRT in the last decade. The main research lines that included the HRT among their tools were focused on the study of driving behaviors and hazard perception of both novice and experienced drivers. For instance, the HRT was employed to study cognitive correlates of driving, such as attention, decision-making, and workload (Di Stasi, Álvarez-Valbuena, Cañas, Maldonado, Catena, Antolí, & Candido, 2009; Megías, Di Stasi, Maldonado, Catena, & Cándido, 2014; Megías, Maldonado, Catena, Di Stasi, Serrano, & Cándido, 2011; Tagliabue *et al*., 2013). Other studies started to focus on hazard perception and hazard perception training on the HRT (Tagliabue & Sarlo, 2015; Vidotto *et al*., 2011), also employing psychophysiological measures (Tagliabue & Sarlo, 2015) and eye movements (Di Stasi, Contreras, Cándido, Cañas, & Catena, 2011). Finally, a first validation of the HRT and a first attempt to employ it as a tool to measure the behavioral counterparts of different personality profiles was made by Miceli, Settanni, and Vidotto (2008) and by Marengo *et al.* (2012), respectively.

However, to date no studies have focused on the generalizability of the found effects to the real world, even though, as seen, this aspect should be explored when a simulator is employed for research. Moreover, the classical approach of the identification of personality profiles differentially correlated to self-reported measures of driving (aberrant) behaviors has frequently produced inconsistent results (see Chapter 2). Finally, more evidences on the development of the process of hazard perception are needed: The HRT simulator, with the wide range of hazardous scenes that can be administered to the participants, represents an opportunity to deeply study and understand this process.

Chapter 2

Factors influencing inexperienced road users' behaviors

As stated in the previous chapter, the causes of the young and inexperienced road users' high crash rate rely on a combination of factors such as on-road inexperience, developmental and personality factors. The aim of this chapter is to introduce and discuss the role of three main psychological variables (namely, hazard perception, attention, and personality traits) in influencing inexperienced road users' behavior.

2.1 Hazard perception

A central question for road safety researchers is whether there is any quantifiable element of driving skills that can explain the difference in the likelihood to crash among drivers. If so, this component should be the core of drivers' assessment and training. Only one element of driving skills has been consistently correlated with drivers' accidents records across many studies: Hazard perception (Horswill & McKenna, 2004).

Hazard perception can be defined as the ability to detect in advance dangerous situations in the traffic (Horswill & McKenna, 2004), so as to be ready to prevent the crash. A wide amount of studies has focused on the differences in the hazard perception ability between inexperienced and experienced road users. Hazard perception has been investigated with a wide range of methodologies, but the most common paradigm requires detecting hazards by pressing a button while watching video-clips of first-person driving scenes (Crundall, Chapman, Phelps, & Underwood, 2003). The Hazard Perception Test developed by McKenna and Crick (1991) represents the beginning and the best expression of this methodology. In their study, McKenna and Crick (1991) found that novice drivers with up to three years of experience reacted significantly slower to hazards showed in a series of first-person driving video-clips than ten-year experienced drivers. Moreover, experienced drivers detected a higher number of hazards than novices. Similar results were reported by Crundall *et al.* (2003), who compared hazard ratings, eye movements, and psychophysiological responses (SC, which reflects an implicit, automated way of detecting hazards that parallels cognitive hazard recognition) of police drivers with those of novice drivers and of an age-matched control group. The results showed that, although the three groups did not differ in terms of hazard ratings, both the police drivers and the matched control had shorter durations in eye fixation than novice drivers across different types of driving clip, such as police vehicles during pursuit and during response to an emergency call, and control clips showing normal driving situations. This was interpreted as a proof of the reduction in the time needed by the two experienced groups to process hazards. Moreover, the police drivers showed a significant higher number of SCRs, meaning that their peripheral system responded to a higher number of hazardous scenes than the other groups.

Taken together, the previous results suggest that hazard perception is modulated by the degree of on-road experience and exposure. This effect, as stated by McKenna and Crick (1991), can be explained as follows. Experienced drivers have access to a wide range of hazardous traffic situations. Whenever they face one of these situations, they can quickly recognize it, with almost no effort. In contrast, novice drivers should be less familiar with hazardous situations and, therefore, they need a greater effort (and time) to recognize the hazard. This explanation implies that hazard perception increases as a function of experience and exposure and that it is a trainable ability. Before talking about how to train hazard perception, however, it is necessary to understand the cognitive and psychophysiological mechanisms underpinning hazard perception.

2.1.1 Risk as feelings and the Somatic Marker Hypothesis

One of the most famous theories regarding hazard perception and decision-making in risky contexts postulates that humans perceive risk and, consequently, decide how to act on risk, in two ways (Loewenstein, Weber, Hsee, & Welch, 2001; Slovic, Finucane, Peters, & MacGregor, 2004; Slovic, Peters, Finucane, & MacGregor, 2005; Slovic & Peters, 2006). The first and so-called "risk as feelings" way is referred to the instinctive and intuitive reaction to danger, whereas the second ("risk as analysis") involves logic and reason to manage the risk (Slovic & Peters, 2006). The first path relies on the so-called "experiential system" which is fast, mostly automatic and unconscious, whereas the second is based on the "analytic system", characterized by a conscious control that requires more effort and more time (Slovic *et al*., 2005).

Usually, everyday risk management follows the first path (Slovic *et al*., 2005; Slovic & Peter, 2006), that is based on the so-called affect heuristic. An affect indicates the quality of "goodness" or "badness" experienced as a feeling state (both consciously and unconsciously), that defines a positive or negative quality of the stimulus (Slovic *et al*., 2004; Slovic *et al*., 2005; Slovic & Peter, 2006). These affects act as cues for the decision-making process, especially when the required decision is complex or needs to be taken rapidly (Slovic *et al*., 2004), which is exactly the case of risky contexts such as road traffic.

One of the most complete theoretical and evidence-based accounts of the role of affect and emotions in (risky) decision-making was developed by Antonio Damasio and his research group (Bechara, Damasio, & Damasio, 2000a; Bechara, Damasio, Damasio, & Anderson, 1994; Bechara, Damasio, Tranel, & Damasio, 2005; Bechara, Tranel, & Damasio, 2000b; Damasio, 1994). Starting from the evidence that, after a damage to the ventromedial prefrontal cortex (VMPFC), humans develop an impairment in real-life decision-making, becoming unable to foresee the negative consequences of their actions, Bechara *et al*. (1994) developed a task, the Iowa Gambling Task (IGT), that attempts to simulate everyday decision-making processes in ambiguous contexts.

At the beginning of the IGT, the participants are given a $2,000 loan of play money and they are told that their aim is to maximize their profit by making a series of card selections from the four decks placed in front of them. After turning each card, the subjects receive some money, whose amount is known only after the turning and varies depending on the deck. However, in each of the four decks, participants encounter unpredictable money loss. All the cards from decks A and B always give a reward of $100, whereas the cards from decks C and D yield $50. However, since the penalties are higher in decks A and B and lower in decks C and D, the first two are considered disadvantageous decks in the long term, whereas C and D are considered advantageous decks.

Comparing the performance at the IGT of a group of patients with lesions to the VMPFC with that of a healthy control group, Bechara *et al*. (1994; 2000a; 2000b) found that control participants chose more frequently from the advantageous decks (C and D), avoiding the disadvantageous ones, whereas the patients showed the opposite pattern. Moreover, considering the trial-by-trial performance, the participants in the control group seemed to start the task by choosing cards from each deck, eventually switching to more selections from the good decks. Conversely, the patients did not increase their selections from the advantageous decks but remained “stuck” to the short-term reward of the disadvantageous ones (Bechara *et al*., 1994; 2000a; 2000b). This behavior was named “myopia for the future” (Bechara *et al*., 1994; 2000a; 2000b) and was attributed to a problem in the somatic marker mechanism, as hypothesized by Damasio (1994).

Bechara and Damasio (2005) were able to distinguish four periods of performance as control subjects went from the first to the last trial in the task, on the basis of the participants’ awareness of the reward/penalty mechanisms of the IGT, thus proving that the task resembles everyday risky/ambiguous situations. The first period is a “pre-punishment” period (*i.e*., participants sample the decks, before encountering any punishment). The second is a “pre-hunch” period, when participants began to encounter punishment, but still had no clue about what is going on in the game. The third is a “hunch” period (*i.e*., the participants begin to express a hunch about which decks are riskier), followed by the fourth "conceptual" period, when subjects know very well that there are good and bad decks, and which decks are good and bad. Although 30% of controls did not reach the conceptual period, they still performed advantageously, whereas although 50% of VM patients did reach the conceptual period, they still performed disadvantageously (Bechara & Damasio, 2005).

The Somatic Marker Hypothesis (SMH) provides a neuroanatomical and cognitive framework to account for decision-making in risky and ambiguous contexts and the influence that emotions have on it. As explained by Bechara *et al*. (2000a; 2000b), the VMPFC would be responsible to store the links between factual knowledge that an individual has experienced in the past and bioregulatory states, including emotional states. Basically, the VMPFC would hold the connections between the features of a given situation experienced in the past and the emotions, in terms of physiological response patterns, paired with it. These links are implicit and can reactivate an emotion by acting on the appropriate cortical and subcortical structures. Thus, when one faces a situation for which some features have been previously categorized and stored, the links to the bioregulatory states are activated. This leads to the reactivation of the pertinent emotional dispositions, reconstructing a previously learned factual-emotional set (Bechara *et al*., 2000a). This reactivation can be carried out *via* actual changes in the body ("body loop") or *via* the so-called "as-if body loop" in which the body is bypassed, and the reactivation is conveyed to some neural structures (*e.g.*, somatosensory and insular cortices, amygdala, brainstem nuclei, basal ganglia) that, at a later stage, adopt the appropriate somatic pattern.

Either when this process is overt or when it is covert, the somatic state acts as a marker (*i.e.*, a signal of alarm or incentive), alerting us of the goodness or badness of a certain option-outcome pair, so as that some options can be rapidly excluded from the decision-making process (Bechara *et al*., 2000a). This mechanism is activated particularly in risky and/or ambiguous situations, when the outcomes of our decisions are unclear and thus it is impossible to employ the "risk as analysis" path (Bechara & Damasio, 2005).

The role of the affects for decision-making in risky and ambiguous contexts was clarified by a series of studies (for a brief review, see Bechara *et al*., 2000a; Bechara & Damasio, 2005) in which the skin conductance (SC) of patients with lesions to the VMPFC was measured during the IGT. The overall results showed that three types of skin conductance responses (*i.e*., a phasic modification of the electrodermal activity, in terms of increase, occurring in proximity of an identifiable stimulus; SCR) can be identified during the task: (1) the reward SCR, that is elicited up to 5 s after turning a card with reward only; (2) the punishment SCR, occurring *after* turning a card with both reward and punishment; (3) the anticipatory SCR, that occurs *before* turning a card from a deck and that is supposed to represent the somatic marker associated with the outcomes of the previous choices (Bechara *et al*., 2000a). Both healthy controls and VMPFC patients show the first two types of SCRs. However, only the controls begin to generate the anticipatory SCRs as they become experienced with the task, showing larger SCRs before selecting cards from the disadvantageous decks (Bechara *et al*., 2000a). Importantly, the increase in anticipatory responses begins during the so-called pre-hunch period, that is *before* any conscious knowledge regarding the IGT has developed (Bechara & Damasio, 2005). On the basis of this (and many other) evidence, SC and, particularly, SCRs have been considered indices of changes in the somatic state controlled by the autonomic system. In risky or ambiguous situations, the failure of this mechanism of somatic marking can lead to disadvantageous decisions, such as continuing to pick cards from the bad decks.

Thus, it is possible to hypothesize that, since traffic is one of the riskiest everyday situations, on-road hazard perception (and anticipation) develops as a function of a mechanism in which on-road experience allows to have a set of features and traffic configurations, linked to somatic states, that, after being experienced, can be reactivated when necessary (hazardous situations). The outcomes of the previous choices in similar situations guide our decision-making by reactivating the correspondent somatic markers and somatic states, without the need of conscious knowledge and reasoning. Therefore, the SC and the anticipatory SCRs of experienced road users should show a peculiar pattern. Moreover, being exposed to previously experienced on-road hazardous situations should lead to the development of anticipatory SCRs, that, in turn, should guide the behavior so as to prevent the development of actual risks and crashes.

2.1.2 Skin conductance as a somatic marker of on-road hazard perception

As seen, Crundall *et al.* (2003) reported that police drivers showed a significant higher frequency of SCRs when watching video-clips of hazardous traffic scenes. However, this was not the first work that studied the relationship between SC and driving. In 1964, Taylor measured the SC of a group of participants during a real-world driving task, finding that traffic events increased the frequency of the SCRs and that, in contrast with the results by Crundall *et al.*, 2003, inexperienced drivers showed more SCRs than experienced drivers. However, both these studies focused on the general psychophysiological activation of the drivers during the task, without considering hazard anticipation.

Kinnear, Kelly, Stradling, and Thomson (2013) attempted to understand how drivers learn to anticipate on-road hazards. The authors compared three groups: Learner (mean annual mileage lower than 200 km/year), inexperienced (mean annual mileage lower than 5,000 km/year), and experienced (mean annual mileage lower than 15,000 km/year) drivers. Participants had to watch a series of 60s video-clips of traffic hazardous scenes and to rate the level of hazard of each clip from “Safe” to “Hazardous”. During the task, the SC of the participants was measured, looking for SCRs in a time-window preceding the beginning of each maneuver performed to avoid the hazard by the drivers in the video-clip. The results supported the existence of a system based on affective appraisal, able to anticipate hazards when driving. Experienced drivers showed significant higher percentages of anticipatory SCRs, that is almost the double than those shown by the inexperienced drivers and three times larger than those elicited by the learners. Moreover, the percentages of anticipatory SCRs linearly increased with the increase of annual mileage.

Thus, it seems that on-road experience and exposure allow to develop the somatic marker underpinning hazard perception: This mechanism allows to anticipate the presence of a potentially hazardous scene and, possibly, to act to prevent the collision. Is it possible, however, to train this ability in a safe environment, minimizing the risks of the on-road exposure? This is the question addressed in a study by Tagliabue and Sarlo (2015), in which a group of inexperienced participants was involved in a training on a driving simulator, whereas a gender-matched control group just had to watch the video-clips of the training group's performance, spotting for hazards. The SC was measured in all participants. The training on the driving simulator was divided into five courses involving 7 or 8 hazardous scenes each, in which the aim of the participants was to prevent collisions. The results showed that the trained group had a higher percentage of SCRs than the control group. Overall, this body of evidence suggests that the somatic marker (indexed by the number of SCRs) underpins the development of hazard perception that, in turns, increases as a function of on-road experience and exposure. Moreover, the study by Tagliabue and Sarlo (2015) suggests that it is possible to train hazard perception in a safe environment, such as a driving simulator. These considerations would partially explain the reported inexperienced drivers' high rate of road crashes. Moreover, they can represent a first step toward the development of a driving training. In the next sections of the chapter, other aspects that influence inexperienced drivers' on-road behavior are discussed.

2.2 Attention

2.2.1 Attention and distraction

The ability to monitor the environment and to detect salient stimuli while engaged in demanding tasks is crucial in everyday life (Posner, 1980; Posner, Snyder, & Davidson, 1980). Nevertheless, bottom-up and top-down factors influence this ability (Folk, Remington, & Johnston, 1992; Posner, 1980; Posner *et al*., 1980; Yantis & Jonides, 1990) and their interaction contributes in determining the focus of attention while monitoring the environment.

Goal-directed behaviors (such as driving) require a high attentional focus on task-relevant stimuli while ignoring distractors. Lavie, Hirst, De Fockert, and Viding (2004) proposed a load theory of attention to explain the process of distractor rejection, that would depend on the level of cognitive and perceptual load involved in the current task. The model includes two routes: The first is based on a passive perceptual mechanism (early attentional mechanism) that prevents perception of distractors when the cognitive system is under high perceptual load. The second (late selection mechanism) allows to ignore distractors perceived in situations of low cognitive and perceptual load (Lavie *et al*., 2004). Corbetta and Shulman (2002) proposed a model that includes two connected neural networks: A goal-directed, and a stimulus-driven attentional system. The goal-directed system is guided by top-down expectations, such as stimulus localization, timing, or content, but interacts with the stimulus-driven system. Thus, on the one hand, top-down processes influence the salience of stimuli and, on the other hand, stimulus-driven processes draw attention toward events that are potentially significant for the ongoing behavior.

One of the most frequently employed paradigms to study attention and monitoring is the novelty oddball (see Friedman, Cycowicz, & Gaeta, 2001). Three types of stimuli are presented with different probabilities of occurrence (Friedman *et al*., 2001; Squires, Squires, & Hillyard, 1975): Standard, frequent stimuli (*e.g*., 1000 Hz pure tones, usually in the 80-90% of trials), infrequent deviant stimuli (*e.g.*, higher pitch tones, in the 5-10% of trials), and a series of out of context, unexpected, and task-irrelevant "novel" stimuli (*e.g.*, a dog bark, again in the 5-10% of trials). In the active version of the paradigm, participants must respond to the deviant stimuli (in this circumstance called target stimuli) and ignoring the others.

The processing of the deviant target stimuli elicits specific event-related potentials (ERPs), such as an amplitude modulation of the N1 and the typical P3 (or P3b), which are thought to represent the involuntary detection of a stimulus that can be modulated by top-down processes such as selective attention (Hink, Hillyard, & Benson, 1978; Horváth, Winkler, & Bendixen, 2008), and the categorization of a task-relevant stimulus among distractors, respectively (Squires *et al.*, 1975). The P3b originates from temporal-parietal regions associated with attentional allocation with a mean latency of about 350 ms (Squires *et al.*, 1975). The N1 has usually a mean latency between 80 and 120 ms from the stimulus onset (Horváth *et al.*, 2008), although latencies from 50 to 200 ms have been reported depending on the physical and temporal features of the stimuli (Näätänen & Picton, 1987). When participants' attention is diverted from the oddball task (*e.g.*, concomitant task) it is possible to measure the mismatch negativity (MMN), with a mean latency around 100-200 ms from the onset of the deviant (Horváth *et al.*, 2008; Näätänen, Gaillard, & Mäntysalo, 1978). The MMN reflects a specific memory-based deviance-detection process derived from the comparison of the incoming stimulus to the neural representations generated on the basis of stimulus regularities (Horváth *et al.*, 2008; Näätänen & Picton, 1987). Finally, the novel event usually elicits the P3a component (Friedman *et al.*, 2001), which has its peak around 220-300 ms post-stimulus in the fronto-central regions (Squires *et al.*, 1975). Recently, research has shown that the P3a may indicate an involuntary high -level process triggered by the need to detect events which might require further processing (Horváth *et al.*, 2008) and not simply the attentional switching *per se*, independently of the meaning and the salience of the stimulus, as it was thought before. Thus, Horváth *et al.* (2008) suggested that the processes indexed by the P3a are activated by significant events.

Another version of the oddball paradigm has been developed to overcome the issues related to the overlap between the concepts of probability and deviance in the novelty oddball, derived from the lower occurrence of the target and novel stimuli as compared with the standard stimuli (Parmentier, Elsley, Andrés, & Barcelóis, 2011). The so-called multi-feature oddball tries to solve this overlap confounding (Pakarinen, Takegata, Rinne, Houtilainen, & Näätänen, 2007) allowing the comparison between the ERPs elicited by different types of deviant stimuli in the same task. The standard stimuli cover generally the 50% of the total trials, whereas the other half of the trials includes the same percentage of different categories of deviant stimuli (differing either in duration, intensity, frequency, or even in many features at the same time). This manipulation allows to assess the effect of the deviance for a given probability of the categories of deviant stimuli, eliciting ERPs with different amplitudes for each category of deviants (Pakarinen *et al*., 2007).

As we will see in the next paragraph, the oddball paradigm can be a useful tool to investigate the attentive deployment during a driving task.

2.2.2 Monitoring and attentional shifting while driving

Driving a vehicle involves several cognitive abilities, such as monitoring the environment, orienting and shifting attention toward potentially salient or dangerous events (Crundall *et al*., 2002; Underwood *et al*., 2003). Michon (1985) made a distinction between the so-called skill-based behaviors (*e.g*., keeping on-road stability), which are automatic and do not involve high levels of cognitive resources, and the rule-based behaviors (*e.g*., maneuvering the vehicle in the traffic), which require a large amount of resources and are more prone to distraction and interference (Wester, Böcker, Volkerts, Verster, & Kenemans, 2008).

To assess the effects of distraction while driving, an auditory novelty oddball was employed in two interrelated studies (Wester *et al*., 2008; Wester, Verster, Volkerts, Böcker, & Kenemans, 2010). In the first (Wester *et al*., 2008), twenty participants were divided into two groups. The first group (active) had to perform a lane-keeping task on a car simulator along a scenario without traffic and to detect at the same time the deviant target stimuli in an oddball paradigm by pressing one button on the steering wheel of the simulator. The second group (passive) had to perform the same driving task while being presented with the oddball paradigm, but without responding to the stimuli. Both groups had a dual-task condition (driving and oddball) and two single-task conditions (oddball alone and driving alone). The EEG activity of all the participants was recorded in each condition. The main results showed that in the passive group no differences in the MMN amplitude were present between the dual- and the single-task conditions. Moreover, only in the passive group the attention was less directed towards the novel stimuli during the dual task (reduction in the amplitude of the P3a)

Finally, the P3b, elicited by the deviant target stimuli in the active group, did not significantly differ in the dual- *vs*. single-task conditions. In a subsequent study, Wester *et al*. (2010) replicated the previous results regarding the MMN and found an effect of the level of alcohol in reducing the amplitude of the P3a during the dual task (driving and oddball) for both the active and the passive group. Finally, the P3b amplitude elicited by the deviant target stimuli linearly decreased with alcohol intake, but only in the dual-task condition. Thus, the authors concluded that two types of mechanisms interact in determining the amplitude of the P3a and P3b components: Top-down, necessary to drive, and bottom-up factors related to attentional capture, which are affected by alcohol.

The interaction between top-down and bottom-up factors represents, as seen, the core of many attentional models, including that by Corbetta and Shulman (2002). The results of the studies by Wester *et al.* (2008; 2010) are also in line with the model by Lavie *et al.* (2014), which states that when the task-related cognitive load increases less attentional resources are available for other concomitant processes, leading to a reduction of distractors' processing. Indeed, the processing of novel sounds was reduced during the more cognitive demanding condition (*i.e.*, the dual task) in both the studies by Wester *et al*. (2008; 2010), as attested by the decrease in the P3a amplitude.

Experience influences the way in which people monitor the environment during a driving task too. For instance, concerning visual attention, a series of studies (Crundall *et al*., 2002; Underwood *et al*., 2003; Tagliabue *et al*., 2013) demonstrated that experienced drivers have a peculiar pattern of gaze behavior and are more able to detect peripheral stimuli under demanding conditions as compared with inexperienced drivers; However, after a learning phase on a simulated moped driving, also inexperienced drivers improved their ability to detect peripheral stimuli. Thus, experience seems to modulate the time course of attentional allocation, suggesting the presence of different attentional strategies in participants with higher on-road experience. The novice drivers' difficulties in detecting peripheral stimuli may stem from the large amount of resources that they need to employ in the driving task itself, whereas more experienced drivers may have more free resources to monitor the environment and to detect peripheral stimuli.

2.3 Personality traits and personality profiles

Besides attention and hazard perception, other factors that contribute to influencing inexperienced drivers' on-road behaviors are personality traits, which are of crucial importance when considering the "problem young driver" (Scott-Parker *et al*., 2013). Indeed, as also suggested by Lucidi, Giannini, Sgalla, Mallia, Devoto, and Reichmann (2010), although young drivers as a group are more likely to drive dangerously, not all young drivers show high-risk tendencies. Thus, the attempt to identify the most at-risk subgroups of inexperienced drivers needs to include the assessment of personality traits and the understanding of their relationships with risky driving behaviors.

2.3.1 Personality traits linked to risky driving behaviors

Among personality traits, the multidimensional construct of sensation seeking (SS) is consistently linked to driving behavior. SS is a personality trait defined by the predisposition to seek novel, varied, exciting and intense sensations and experiences (Zuckerman, 1994). In his systematic review, Jonah (1997) found consistent associations between high SS and risky driving behaviors in the majority of the articles included in his work. The associations were steady across cultures, stronger for males, and tended to decline with age (Jonah, 1997). Although SS in general seems to account for up to 15% of variance in risky driving, its thrill seeking (TS) subdimension is the most related to on-road risky behaviors (Jonah, 1997). Moreover, considering that SS is an unstable trait, with higher levels during adolescence and youth (Zuckerman, 1971), it represents a predictor of reckless driving, especially among young road users (Dahlen, Martin, Ragan, & Kuhlmanet, 2005; Jonah, 1997). For instance, among adolescents, high levels of SS are linked with driving under the influence of substances, exceeding the speed limit and racing other vehicles (Arnett, Offer, & Fine, 1997). Moreover, SS is a predictor of teenagers' aggressive driving and driving anger (Dahlen *et al*., 2005). To explain these relationships, Jonah (1997) proposed two alternative hypotheses. The first states that the so-called sensation seekers do not perceive the real level of risk of on-road situations because of overconfidence in their driving skills. The second is that sensation seekers normally perceive the risk but accept it to experience thrill. If negative outcomes do not occur, the level of perceived risk furtherly decreases, leading the sensation seekers to increase their frequency of risky on-road behaviors (Jonah, 1997).

Sensation seeking is also frequently associated with aggressiveness in predicting risky driving behaviors (Arnett *et al*., 1997; Ulleberg & Rundmo, 2003). High levels of aggressiveness (*i.e*., the tendency to act in a verbally and physically aggressive way and to experience anger and frustration; Arnett *et al*., 1997; Ulleberg & Rundmo, 2003), correspond to high frequency of speeding behaviors among teenagers (Arnett *et al*., 1997). Nevertheless, this relation is not completely clear, since some studies (*e.g.*, Ulleberg & Rundmo, 2003) found only an indirect relationship between high aggressiveness and risky on-road behaviors, with a small-to moderate effect size.

Another trait that has been frequently reported as related to driving behaviors is locus of control (LC; Montag & Comrey, 1987; Özkan & Lajunen, 2005; Warner, Özkan, & Lajunen, 2010). It can be defined as a personality trait that reflects the degree to which people perceive events as under their control (internal LC) or under the control of external forces that cannot be managed (external LC). Controversial evidence of links between LC and the increase in the crash rate are reported in the literature. For instance, Montag and Comrey (1987) reported an association between external LC and higher crash rate, whereas Özkan and Lajunen (2005) found a correspondence between internal LC and a higher number of self-reported crashes, violations and errors in a sample of young drivers. Moreover, Warner *et al*. (2010) found a positive relation between internal LC and speeding behavior.

Finally, focusing on adolescents, research has proven that crash rate rises consistently when adolescents are with a peer (Preusser, Ferguson, & Williams, 1998) and that they tend to drive faster and to show higher frequency of aberrant behaviors when carrying peer than adult passengers (Baxter, Manstead, Stradling, Campbell, Reason, & Parker, 1990). These behaviors are included in what Allen and Brown (2008) call direct (proximal) peer influence, occurring when adolescents are driving and carrying their peers as passengers (Baxter *et al*., 1990). However, peer influence can also be expressed indirectly (distal influence; Allen and Brown, 2008). Indeed, the so-called "caravan peers" (*i.e.*, peers driving other vehicles on the road), whose conduct is observed by adolescents, can influence teenagers' driving behavior too (Allen and Brown, 2008). This indirect influence may shape adolescents' norm setting and their beliefs about peers' behavior and, in turn, it may lead to a distortion in risk estimation. This is exactly what a survey about 11,000 teenagers in the United States showed (Ginsburg, Winston, Senserrick, Garcia-España, Kinsman, Quistberg *et* *al*., 2008). Only 15% of the respondents perceived the teenage drivers as inexperienced, although 60% of the sample stated that inexperience heavily affects road safety. Moreover, about a half of the respondents stated that they frequently see their peers involved in on-road risky behaviors. These results indicate that adolescents do not perceive their peers as inexperienced and as potentially dangerous drivers, although they can detect risky driving behaviors among them. Thus, beliefs about peers may affect the development of adolescents' defensive driving strategies (*e.g*., self-regulation, as a function of beliefs about others' driving behaviors), contributing to increase their crash rate. As reported in the next section, all these aspects have been frequently considered jointly by studies that attempted to identify personality profiles with different degrees of crash proneness.

2.3.2 Personality profiles and driving: Models and techniques

In order to identify subgroups of young drivers more prone to on-road risky behaviors and crashes, researchers adopted two main approaches that will be briefly discussed. The first approach consists in assessing the relations between drivers' personality profiles, identified through self-assessment measures, and self-reported driving behaviors. For instance, Ulleberg (2011) carried out a survey of 6,000 Norwegian drivers, between 18 and 23 years old, measuring participants' personality traits (SS, anxiety, altruism, aggressiveness, and normlessness), self-reported angry driving and risky on-road attitudes and behaviors. Through a cluster analysis of the personality variables, six groups were identified. Two of them were considered at risk for road crashes: The first was characterized by high levels of SS and normlessness but low anxiety and altruism. The second included participants with high levels of SS, anxiety, aggressiveness, and angry driving. These two groups also reported the riskiest driving habits and the highest frequency of road crashes. The heterogeneity in the profiles' characteristics led the author to state that young drivers cannot be treated as a homogenous group.

In Italy, Lucidi *et al*. (2010) identified different young drivers' personality profiles and assessed their relationship with self-reported aberrant on-road behaviors, which are theoretically divided into intended and unintended violations, errors, and lapses and usually measured through the Driver Behavior Questionnaire (DBQ; Reason *et al*., 1990). The authors measured a number of personality traits (*e.g*., SS, anger, anxiety, and LC), self-reported aberrant on-road behaviors, and the amount of accident involvement. Three clusters emerged: Risky drivers (high levels of SS, angry driving, normlessness and an external LC), worried drivers (high levels of anxiety and hostility) and careful drivers (high levels of altruism and low levels of anger, hostility, SS, and normlessness).

The participants included in the first group reported the highest crash rate and the riskiest driving attitudes although they perceived themselves as less prone to accidents. Careful drivers showed a reverse profile (low rates of errors, violations, lapses, and crashes) whereas worried drivers were classified as a medium-risk profile, because they reported a rate of lapses comparable to that of risky drivers but overall better attitudes. In a recent study (Lucidi, Mallia, Giannini, Sgalla, Lazuras, Chirico *et al*., 2019) three subtypes of moped-drivers (risky, worried and careful) were identified in a large sample of Italian adolescents by means of a cluster analysis of personality characteristics. The three profiles, besides being comparable to those usually identified among experienced drivers, proved to be different in terms of attitudes toward safe driving, self-reported crashes, and risky driving behaviors.

Overall, these three studies provided evidence about the chance to identify different personality profiles among young road users with specific relations with their self-report crash rate and on-road attitudes. However, the limits of this approach rely on the shortcomings of the self-report measures, that prevent the possibility to draw a direct link to behavioral variables. Moreover, the use of self-report driving measures limits the inclusion of inexperienced drivers in the sample, resulting in the impossibility to discriminate between the role of driving experience and personality traits in determining driving behaviors.

The second research approach tries to overcome these limits using driving simulators to assess the driving behaviors of participants and studying the relations between these objective parameters and the personality profiles. Deery and Fildes (1999) identified five clusters in a sample of adolescents (16–19 years old) on the basis of their personality traits and driving attitudes. The most at-risk group was characterized, by high levels of hostility and SS, and by risky driving attitudes. Then, the authors randomly selected a subsample of participants to test, through a driving simulator, whether individuals with different personality profiles differed in their driving behaviors. The results showed that individuals in the more at-risk cluster were also more prone to the negative effects of workload while driving, had difficulties in facing hazardous scenes, and were less cautious in terms of driving speed.

Following a methodology similar to that employed by Deery and Fildes (1999), Marengo, Settanni, and Vidotto (2012) identified different personality profiles among Italian adolescents (14–15 years old) with various degrees of moped-driving experience. Three clusters emerged: Profile A was characterized by high levels of anxiety and low levels of SS and altruism. Profile B showed high levels of SS and impulsivity and low levels of altruism and anxiety, being considered the most at-risk. Profile C showed high levels of altruism and was characterized by a more internal LC. Then, the authors compared the clusters in terms of self-report and simulated driving measures. Participants' performance was assessed through 12 courses on a moped-driving simulator (Honda Riding Trainer), divided into three sessions. For the performance in each course, a letter score was provided: A (safe performance), B (almost safe), C (near miss), and D (accident). The first measure analyzed was the number of accidents (D score). In addition, the authors developed a safe driving index based on scores A, B, and C. Profile B participants showed the highest rate of self-reported aberrant driving behaviors (*e.g.*, driving under the influence of substances) and had the worst performance on the simulator, that is the highest accident rate and the lowest safe driving index score.

The main contribution of this second research approach to the issue of the identification of different young drivers' subgroups is the demonstration that it is possible to overcome the limits of self-report measures, confirming at the same time some previous evidence. To illustrate, profile B in the study by Marengo *et al*. (2012) was comparable to the "risky drivers" in Lucidi *et al.* (2010; 2019), whereas profile A was comparable to one of the low-risk groups of Ulleberg's (2001) study.

The similarity between the teenagers' clusters identified by Marengo *et al*. (2012) and previous results from samples with different ages suggests the presence of important differences among adolescents in the early stage of driving experience too. However, this approach has one crucial limitation too. Indeed, the identification of profiles as a function of personality traits can lead, as seen, to inconsistent results, probably due to cultural peculiarities and to the instability of some personality traits during the lifespan (*e.g*., SS), weakening the replicability and generalizability of the results. The solution, suggested and tested in the studies reported in the next chapters, may rely on the use of driving simulators and on the identification of *driving* profiles predicted by different configurations of personality traits, instead of *personality* profiles related to driving behaviors. Indeed, to the best of our knowledge, no studies have attempted to revert the approach, by taking advantage of the wide numbers of driving parameters collected by the HRT to identify driving profiles, and by comparing them in terms of personality characteristics at a later stage.

Thus, from all the above-mentioned considerations stemmed the idea to employ the HRT simulator to develop and test a training protocol for hazard perception (Chapter 3), also studying its generalizability to the real road context (Chapter 5), and to try to identify driving profiles, differentially associated to risk-related psychological characteristics (Chapters 4 and 5).

Chapter 3

Hazard perception and attention while driving: Psychophysiological indices

In this chapter, three studies are reported. The first (section 3.1; Tagliabue, Gianfranchi, & Sarlo, 2017) focused on the attempt to understand the psychophysiological mechanisms underpinning hazard anticipation in novice road users during a two-sessions active training on the HRT simulator. The second (section 3.2; Tagliabue, Sarlo, & Gianfranchi, 2019) employed part of the previous sample to compare the behavioral and psychophysiological effects of a three-session active training on the HRT with a three-session passive training (watching video-clips of hazardous scenes) among novice road users. The last (section 3.3; Gianfranchi, Mento, Chierchia, Duma, Sarlo, & Tagliabue, submitted) measured the changes in the attentional ERP components (N1, MMN, P3a) elicited by a multi-feature passive oddball task during a one-session training on the HRT. These changes are supposed to be the consequence of the increase in the amount of available attentional resources, due to the training.

3.1 Training hazard perception in a simulated environment: Skin conductance as a somatic marker

The aim of this study[1] (Tagliabue *et al.*, 2017), that represents the first step toward the definition of a training for hazard perception in the HRT simulator, was to investigate the psychophysiological mechanisms through which hazard perception develops.

[1] An extended version of section 3.1 can be found in Tagliabue, M., Gianfranchi, E., & Sarlo, M. (2017). A first step toward the understanding of implicit learning of hazard anticipation in inexperienced road users through a moped-riding simulator. *Frontiers in Psychology*, *8*, 768.

We started from a body of evidence (already discussed) that can be summarized as follows. Previous studies (Crundall *et al*., 2003; Kinnear *et al*., 2013) pointed out that experienced participants show greater electrodermal activity when facing hazardous scenes than participants with a lower level of driving experience. Moreover, inexperienced drivers showed SCRs both during the viewing of on-road scenarios and while actively riding through them; However, psychophysiological reactivity was greater in the latter case, as shown by Tagliabue and Sarlo (2015). As seen in Chapter 2, in the framework of the SMH (Damasio, 1994), the process of somatic marking should be anticipatory, but evidence on a better anticipatory mechanism of experienced drivers tested by means of viewing of hazardous video-clips is unclear. Thus, we hypothesized that a virtual, active and more naturalistic procedure, in which participants can act to avoid or prevent the hazards, could allow to assess (if present) the change in anticipatory activity of inexperienced road users during a training with the HRT. Indeed, if learning to drive consists of an improvement in hazard perception, meant as the ability to predict incoming dangers in advance and indexed by increases in electrodermal activity, running again the same virtual courses should lead to the development of earlier SCRs, along with a better driving performance.

3.1.1 Material and methods

Participants

Sixteen undergraduate students at the University of Padua (9 F; mean age: 20 years; age range: 19–24 years) were recruited. All participants were novice drivers/riders in that they had held their driver's licenses for no more than 2.5 years (range 5–30 months; mean 12.3 months). We set the inclusion criterion for road exposure to 5,000 km of overall (with both cars and two-wheeled vehicles), as indicated by previous works (*e.g*., Crundall *et al*., 2003; Crundall, 2016; Kinnear *et al*., 2013). Participants were paid 13 euros per session, had normal or corrected-to-normal vision and were naïve to the purpose of the study. The study was approved by the Ethical Committee for the Psychological Research of the University of Padua.

Apparatus and setting

The HRT simulator (see Chapter 1 for more details) was employed for the driving task, using the automatic transmission option. For the electrodermal activity recording, an amplifier system -a Grass Model SCA1 skin conductance coupler, associated with a Grass CP122 AC/DC Strain Gage amplifier (Grass Instrument Co., W. Warwick, RI, USA) – that returned the ongoing values of the electrodermal activity through a display, was placed near a secondary screen that reproduced the virtual environment that the participants were seeing while driving. A video camera recorded the driving performance shown by this monitor, along with participants' electrodermal activity (shown by the display of the amplifier) in an .AVI file.

Procedure

The experiment was divided into two sessions scheduled a week apart. At the beginning of the first session, the participants filled out a questionnaire regarding their driving and riding habits and experience. They also signed an informed consent form. Then, they were invited to sit on the HRT, and two electrodes were placed with K-Y lubricating jelly on the left foot over the abductor hallucis muscle — adjacent to the sole of the foot and midway between the proximal phalanx of the big toe and a point directly beneath the ankle, following the advice of Boucsein, Fowles, Grimnes, Ben-Shakhar, Roth, Dawson, *et al.* (2012). The participants were instructed to drive in the simulated environment following the vocal instructions given by the HRT, avoiding hazards and accidents and respecting the traffic rules. Finally, they were invited to inhale, hold and release their breath for a few seconds to test the reliability of the SC signal recording.

In each session, all participants were administered the same five courses at the HRT on peripheral roads, so as to observe the expected anticipation in psychophysiological responses. As already stated in Chapter 1, each course of the HRT includes 8 hazardous scenes, apart from one course that includes only 7. Here, each participant faced a total of 39 potentially hazardous scenes in each session (four courses with 8 hazardous scenes, one with 7). The courses were administered according to their increasing degree of difficulty (see Miceli *et al*., 2008). An exercitation course without other virtual road users was provided at the beginning of the first session. At the end of each course, a 3-min rest was provided so as to allow the SC to return at the baseline level. Each session lasted about 45 minutes.

Data reduction and design

Concerning the driving performance, we computed the percentage of crashes occurred (number of accidents × 100/total number of the hazardous scenes in each course and session).

To reduce and analyze the psychophysiological data, we extracted the .csv files provided by the simulator in which all the variables of the driving performance of participants were collected for each course at a sampling rate of 30 Hz. Through the analysis of the .AVI files, we coded the values of electrodermal activity at the same sampling rate. In this way we managed to match the electrodermal values with the behavioral variables the simulator provided for each participant in each course and session.

For each of the 39 hazard scenes in each session, we identified the clues of each participant (*i.e.*, the points on the paths, in terms of x/y coordinates, *after* which hazards began to develop). The clues were the same for every participant and were used to split the scenes into two parts: Baseline, (a 5-s time window before the clue, providing the baseline for the electrodermal activity of each participant at that point of the course), and postclue (a 10-s time window during which the hazards develop). A 10-s window was chosen to prevent the overlapping with other physiological changes due to other events present in the complex scenarios, ensuring that the recorded increase in electrodermal activity was linked to the specific hazard.

We computed (for each participant in each scene) the mean baseline value of the electrodermal activity and identified, in the post-clue window, the moment at which an SCR (*i.e.*, an increase from baseline in electrodermal activity greater than 0.05 μS) occurred. Then, we computed the percentage of SCRs of the hazardous scenes in each course and session (number of SCRs × 100/total number).

The onset anticipation of the SCRs was measured in terms of distance on the path after the clue, that is as the absolute difference (irrespectively from the left/right direction of the motion) between the x-or-y-value of the position of the participant on the path (depending on the direction of the movement) when the clue appeared, and the x-or-y- value of the position of the participant when the first SCR occurred. Since, each unitary change corresponds to a shift of approximately 1 m in the virtual scenario, the SCR onset was computed as the distance in meters that the participants covered after the clue, until the appearance of the first SCR.

The dependent variables, analyzed in three subsequent Analyses of Variance (ANOVAs), were the percentage of accidents, percentage of SCRs, and mean onset of SCRs. In all the analyses, we used a 2 × 2 × 5 design with the following independent variables: *Gender* (between factor; two levels), *Session* (within factor, 2 levels) and *Course* (within factor; 5 levels). The *Gender* factor was added to control also for possible effects of gender in learning processes. All the analyses were performed with IBM SPSS Statistics 22 software. The α level was set at 0.05. Post hoc comparisons were corrected with Bonferroni method.

3.1.2 Results

Driving performance

The first ANOVA was carried out on the accident percentages to check for learning in terms of reduction in the number of accidents. Two sources of variance reached significance: *Session*, $F(1,14) = 31.64$, $p < 0.001$, $MSe = 132.85$, and *Course*, $F(4,56) = 6.19$, $p < 0.001$, $MSe = 243.22$. The *Course* × *Gender* interaction was also significant, $F(4,56) = 2.69$, $p < 0.05$, $MSe = 243.22$. Concerning the *Session*, the percentage of crashes was higher in the first than in the second session (M = 26% *vs*. 16%, respectively). Regarding the *Course* effect, after an increase from the first (*i.e.*, the easiest) to the second course ($p < 0.01$ at the post hoc), the percentage of accidents decreased in the third course ($p < 0.01$), remaining stable in the last two courses (Figure 2).

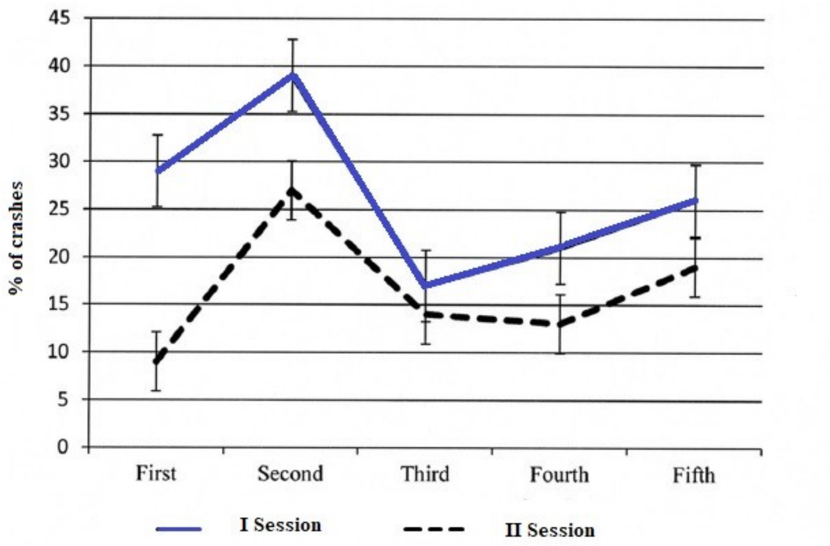

Figure 2: Participants' percentages of crashes as a function of courses (x-axis) and sessions. Error bars represent standard errors. Adapted from Tagliabue *et al*. (2017).

Concerning the *Course* × *Gender* interaction, the post hoc tests resulted significant only for the second and third courses, with males incurring fewer accidents than females ($p < 0.05$). However, in the last two courses, the percentage was comparable between groups.

Skin conductance responses

A second ANOVA with the same design was conducted on the SCR percentages. Again, *Course* and the *Course* × *Gender* interaction reached significance, with $F(4,56) = 2.93$, $p < 0.05$, $MSe = 247.07$, and $F(4,56) = 2.67$, $p < 0.05$, $MSe = 247.07$, respectively. The percentage of SCRs decreased along the courses (Figure 3), with no differences between sessions. Post hoc comparisons showed significant differences between the percentages of the third and fourth courses relative to those of the first course (all the comparisons with $p < 0.05$). The post hoc tests on *Course* × *Gender* interaction failed in revealing significant comparisons. However, the trend seems to show that females had higher SCR percentages in the first two courses, although in the last three courses, mean percentages of SCR were comparable to those shown by males.

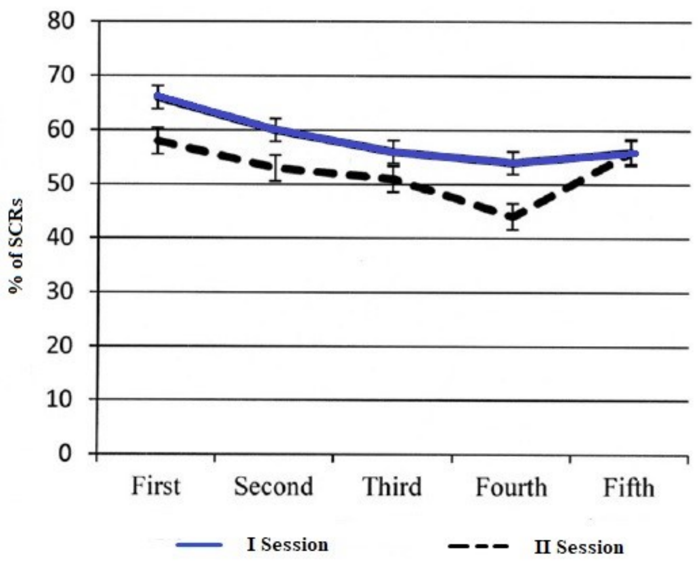

Figure 3: Participants' percentages of SCRs as a function of courses (x-axis) and sessions. Error bars represent standard errors. Adapted from Tagliabue *et al*. (2017).

Finally, we conducted the same ANOVA as before on the SCR onset. Only the factor *Session* reached significance, $F(1,11) = 6.85$, $p < 0.05$, $MSe = 43.51$. However, because in this analysis three participants were discarded because of missing data (*i.e.*, no SCRs in some levels of the factor *Course*), we ran the ANOVA without the *Course* factor (*i.e.*, averaging participants' SCR onset values across courses within each session). Note that this kind of missing data is not due to errors in the procedure or failures in data collection but is strictly related to the phenomena under investigation, in that if no increase in electrodermal activity occurs, then, in that condition, no SCR onset can be calculated. Neither the factor *Session* factor nor the *Session* × *Gender* interaction were significant. However, inspecting the data, the distribution of the second session showed the presence of one outlier. By repeating the ANOVA after having discarded this participant, the significance of the *Session* factor replicated the results of the first ANOVA conducted on mean SCR onset, with $F(1,13) = 7.41$, $p < 0.05$, $MSe = 10.64$. The mean SCR onset (in terms of m after the clue) was 13 m in the first session and 10 m in the second session, thus showing clear anticipation. No other sources of variance reached significance.

3.1.3 Discussion

The present data show an evident effect of learning in the second session, in which participants' crash rate significantly decreases, although, considering the trend in the courses, the decrease is not linear. This result may be due to the joint effect of learning and task difficulty. Indeed, participants seem to begin to ride cautiously in the first course, increasing their accidents in the second course. However, the number of crashes significantly decrease with the progress of the training, despite the increase in the task difficulty, confirming, as well as the general decrease in the accidents' percentage in the second session, the achievement of the learning process. Concerning the differences between males and females, our results suggest the presence of different timings for learning. However, regardless of these slight differences, males and females reach the same learning level during the training.

As for the SCR percentages, the data indicate a decrease as training proceeds. Although this result might appear counterintuitive because of the opposite trends shown in previous studies (*e.g.*, Crundall *et al*., 2003), it is substantially in line with our hypothesis regarding the effectiveness of an active training on the HRT. To illustrate, the results obtained in previous studies showed that high levels of hazard sensitivity were paralleled by a greater frequency of SCRs (Crundall *et al*., 2003). However, this effect was found in a passive detection task, during which participants could not do anything to prevent the hazards. In contrast, in the present study, participants could act in the virtual environment to prevent the development of risky situations. Thus, as participants' ability to ride safely improved, less serious hazards might have occurred, and, consequently, less frequent changes in electrodermal activity might have been recorded, because the reactions of participants' affective appraisal system were less often required.

Beside these considerations, if the participants have learned to ride in such a way that allows them to prevent the occurrence of hazards, this learning should lead to an earlier reaction, as predicted by the SMH framework (Damasio, 1994). Indeed, our results regarding the onset of SCRs seem to confirm this prediction: Participants showed implicit activation in proximity of the hazards approximately 3 m before when they faced the same courses anew during the second session.

The first limitation of the present study is that the retest phase (second session) included the same identical courses of the first. This choice was deliberately made to better assess the changes occurring during learning processes, controlling for possible confounding effects, such as those related to differences in the characteristics of different HRT courses. Moreover, it can be argued that in the second session participants may have recalled some of the hazards from the first-time driving and that the results regarding the second session may reflect a memory effect. Indeed, this is precisely what can be predicted starting from the SMH (Damasio, 1994), as this is the way in which our appraisal system is supposed to work. As shown in the studies with the IGT (Bechara *et al*., 1994; 2000), re-exposure to the same situation is an essential step for the emotional system to develop anticipatory SCRs and, consequently, to learn how to avoid hazards. Finally, another aspect that requires attention is the sample structure. Given that most of the participants were females (frequently underrepresented among motorcyclists), our sample might not represent the target population. However, it is noteworthy that recent evidence suggests that the use of two-wheels motor vehicles is strongly increasing among females (Organization for Economic Co-operation and Development – OECD- International Transport Forum, 2015).

The present study arose within the framework of decision-making in risky contexts (Bechara *et al*., 1994; 2000; Damasio, 1994) and of investigations of the role of hazard perception while driving (*e.g*., Kinnear *et al*., 2013; Tagliabue & Sarlo, 2015; Slovic & Peters, 2006). Specifically, we aimed at focusing on the development of the implicit mechanism of hazard perception. For this purpose, extending the experimental procedure of Tagliabue and Sarlo (2015), we administered the participants with a two-sessions driving training (*i.e*., driving in a virtual environment on the same road courses for two times) to understand how SCRs vary when participants face already experienced conditions. An improvement in the ability to detect and avoid hazards should lead to safer driving behaviors with a consequent reduction in the crash rate but, crucially, with an anticipation of the SCR onset, which would attest a greater readiness of the implicit, unconscious and prompt detection of the presence of hazards. The results were in line with our predictions, thus demonstrating that (1) a two-session training on the HRT leads to an improvement in the driving performance (reduction in the percentages of accidents) when facing the same hazards anew; (2) this effect is a consequence of the improvement in the ability to detect earlier the already-encountered hazards, as attested by the anticipation in the onset of the SCRs during the second session and by the decrement of the percentages of SCRs within the courses and between the sessions. The next step will be the investigation of whether and how this learning (and these psychophysiological effects) generalizes to other situations, such as different courses with different hazards, possibly comparing it with a more classical hazard perception training, such as spotting hazards (for instance, by pressing a key) while viewing hazardous video-clips.

3.2 Comparison between active and passive training in a simulated environment: Differences in skin conductance changes

On the basis of the results of Tagliabue *et al*. (2017), the here reported study[2] (Tagliabue, *et al*., 2019) aimed at investigating three main questions that were still open, specifically whether: (a) the attested anticipation in the SCR onset can generalize to new hazardous situations; (b) the improvement in anticipatory ability underpinning hazard perception differs on the basis of whether the training is active (*i.e*., driving the HRT simulator) or passive (*i.e*., watching a video-clip); and (c) the overall safety in driving performance, besides the crash occurrence, is impacted by active training.

As regards the third point, in the previous studies (Tagliabue & Sarlo, 2015; Tagliabue *et al*., 2017) the improvement in driving performance was measured focusing only on the crashes, neglecting a crucial aspect of road safety, namely avoiding hazard development. Indeed, learners may reach a lower percentage of crashes because they have learned to perform last-minute maneuvers to avoid impending collisions, and not because they have learned to drive safely. Therefore, in the present study we measured also the Evaluation score (see Chapter 1, section 1.6) as an index of driving performance and effectiveness of the training (see aim c).

Moreover, participants in the active group of the study conducted by Tagliabue *et al*. (2017) were asked to complete a third training session (six new courses) to investigate the generalization of learning (see aim a). Finally, a new group of participants matched for age, gender and on-road

[2] An extended version of section 3.2 can be found in Tagliabue, M., Sarlo, M., & Gianfranchi, E. (2019). How can on-road hazard perception and anticipation be improved? Evidence from the body. *Frontiers in Psychology*, *10*, 167.

exposure (passive group) was recruited and involved in a training consisting in watching the video-clips of the performance of the active group, identifying hazards by pressing a key (see aim b). Both behavioral and SC data were collected to assess how participants learned to detect and anticipate hazards.

The experimental work was based on three hypotheses: (1) if the effects of learning can be generalized to new scenes, from a behavioral point of view we expect to find in the third session not only a reduction in the crash rate in the simulated environment, but also a reduction in near misses frequency and an increase in the safe behaviors; (2) if the improvement of hazard perception is triggered by the active training more effectively than by watching a video-clip, the passive group would show a different pattern of psychophysiological activation, as compared to the active group; (3) if the anticipatory ability acquired in the first two sessions by the participants that actively drive the simulator really generalize to new road scenes, we expected SCR onset to occur earlier in the third session than in the first. However, as predicted by the SMH (Damasio 1994), if re-exposure is necessary for the development of anticipatory SCRs, then in the third session the SCR onset should be similar to that of the first session.

3.2.1 Material and methods

Participants

Thirty-eight undergraduates at the University of Padua were included in the present study. However, data from one male participant of the passive group were excluded due to SC recording issues. Consequently, the data from the matched participant of the active group were also eliminated. The final sample included 36 participants (18 F; mean age 19.47 years; range 18–24 years). The active group was the same of the study conducted by Tagliabue *et al*. (2017), with the inclusion of three new participants. The other 18 participants, assigned to the passive condition, were all naive to this set of studies.

All participants were novice drivers/riders; They held a driver's license for no longer than 3 years (range 5–36 months; mean 9.8 months). Nine students had a riding license, but reported road exposures under 5,000 km. The active and passive groups were balanced for age (mean age = 19.88 and 19.05 years, respectively) and gender (9 males and 9 females in each group). All participants had normal or corrected-to-normal vision. They were paid €39 for their participation. The study was approved by the Ethical Committee for the Psychological Research of the University of Padua.

Setting and apparatus

We employed the same tools (HRT, amplifier, and video camera) and setting as already described for the study of Tagliabue *et al.* (2017). See section 4.1.1 for more details.

Procedure

For both groups, the procedure included three sessions scheduled one week apart. Each session lasted approximately 45–60 min. At the beginning of the first session, after signing informed consent forms, each participant completed a questionnaire collecting data related to their driving and riding experience. Then, electrodes for SC recording were attached following the same procedure employed in Tagliabue *et al*. (2017).

The participants in the active group were instructed on how to use the HRT controls and then performed an exercitation course in the virtual environment. For more details regarding their task, see section 3.1.1. Note that here a third session has been added, in which six new courses were administered (48 scenes). Thus, overall, each participant in the active group faced 126 hazardous scenes. Again, the order of administration of the courses was fixed for each session, from the easiest to the most difficult (Miceli *et al*., 2008).

The passive group had to watch a simulation of the courses (five in the first two sessions and six in the third one) undertaken by an anonymous HRT rider and identify hazards, by pressing a button on the handlebar of the simulator. This detection task had the purpose of ensuring constant attention to the video. Given the purpose of the task, no responses were recorded. Each participant in the passive group, matched to a same-gender participant in the active group, watched the replay of the performance of his/her paired participant in the corresponding sessions.

At the end of each course, a 3-min rest was scheduled for participants in both groups to allow SC to return to the baseline level.

Data reduction and design

As seen, the HRT provides an evaluation of participants' riding safety for each scene. Possible scores range from A (totally safe driving behavior) to D (crash), depending on the safety of the driving performance. Each hazardous scene received a score. Thus, we computed the percentage of accidents occurred in the active group (number of accidents × 100/total number of scenes). Moreover, we computed the percentages of scores (A, B, C, or D, computed for each score as number of scenes with that score × 100/ total number of scenes) attributed to each scene.

As in the previous study (Tagliabue *et al*., 2017), the coding procedure for the electrodermal activity was based on the videos recorded by the camera (in which the electrodermal values displayed by the amplifier and the performance on the HRT were synchronized) and on the .csv files, provided by the simulator (see section 4.1.1). As in the previous work (Tagliabue *et al*., 2017), for each hazardous scene, we identified a clue and we focused on two temporal windows: A 5-s baseline pre-clue window and a 10-s post-clue window. Note that each participant was provided with the same clues, both in the experimental and the control group. Again, following the methodology employed in Tagliabue *et al.* (2017), the mean level of the electrodermal activity in the baseline window was computed. Then, an SCR was detected as the first increase in amplitude of at least 0.05 μS (Boucsein *et al*., 2012; Tagliabue and Sarlo, 2015; Tagliabue *et al*., 2017) in the post-clue window, with respect to the baseline. The time of SCR onset was then converted into its corresponding position on the path, with a change in the coordinates' value corresponding again approximately to 1 m. The coding procedure was the same for all the participants in every session since each participant in the control group watched the performance of his or her matched experimental participant, thereby being exposed to the same scenes.

The SCR percentages were computed for each category of scene (A, B, C, and D), as the proportion of SCRs detected over the total scene in each category. In courses in which one kind of scene (A, B, C, or D) did not occur, depending on the participant performance in the active group, the missing data were replaced by the mean for the same risk category. The onset anticipation of the SCRs was measured in terms of distance on the path after the clue, as in the previous study (Tagliabue *et al*., 2017).

To investigate the effectiveness of the three-sessions training, we conducted two different analyses. The first was an ANOVA on the accidents' percentages with *Session* as a within-subject factor (three levels). The second was a deeper analysis of the participants' driving performance, conducted through a multivariate analysis of variance (MANOVA) on the percentages of score (A, B, C, or D, computed for each score as the number of scenes with that score × 100/ total number of scenes) attributed to each scene with *Session* as a within-subject three-level factor.

To investigate whether SCRs paralleled improvements in driving performance, and to compare psychophysiological reactivity of the active and passive group, two analyses were carried out on the percentage of SCRs.

The first was a 2 × 3 ANOVA conducted on the overall SCR percentages (independently of the degree of risk of the scenes) and was aimed at comparing the present results with those obtained by Tagliabue *et al.* (2017), with *Group* (active *vs*. passive) as a between-subject factor and *Session* as a within-subject factor. The second was a 2 × 3 × 4 ANOVA with *Group* as a between-subject factor, and *Session* and *Risk degree* (A, B, C, and D) as within-subject factors. This second ANOVA was run to provide more information through a deep analysis of the pattern of psychophysiological changes as a function of the risk degree.

Finally, to test the impact of the three-session training on the anticipatory ability, we analyzed the SCR onset via a $2 \times 3 \times 4$ ANOVA with *Group* as a between-subject factor, and *Session* and *Risk degree* as within-participants factors.

The α level was set at 0.05. Post hoc comparisons were run with Bonferroni correction.

3.2.2 Results

Driving Performance

As regards the driving performance, in the ANOVA on the percentage of accidents made by the active group, the factor *Session* resulted significant ($F(2,34) = 107.46$, $p < 0.001$, $\eta_p^2 = 0.86$), with participants having 28% of accidents in the first session, 16% in the second, and 4% in the third. All comparisons were significant at the post hoc tests (all the comparisons with $p < 0.05$), confirming that performance improved both in the second (including the same courses as the first session) and in the third session, when participants had to face new, different courses.

In the MANOVA on the percentage of scene's scores, the factor *Session* resulted significant (Wilks' $\lambda = 0.027$, $F(6,12) = 71.64$, $p < 0.001$, $\eta_p^2 = 0.97$). At the univariate level, the factor *Session* was significant for each score: $F(2,34) = 99.85$, $p < 0.001$, $\eta_p^2 = 0.86$, for the A score; $F(2,34) = 5.01$, $p < 0.05$, $\eta_p^2 = 0.23$, for the B score; $F(2,34) = 37.26$, $p < 0.001$, $\eta_p^2 = 0.69$, for the C score; $F(2,34) = 103.08$, $p < 0.001$, $\eta_p^2 = 0.86$, for the D score. As shown by the post hoc, the percentage of A scores (*i.e.*, totally safe performance) significantly increased from the first to the second (M = 15 *vs*. 26%), and from the second to the third session (M = 47%). On the other hand, the percentage of D scores (*i.e.*, accidents) decreased along the three sessions (M = 28, 16, and 4%, respectively) (see Figure 4). As for the B and C scores, no differences were found between the first two sessions. However, in the third session, the percentage of both scores significantly differed from both the previous sessions, in that the percentages of B scores significantly increased in the last session, whereas those of C scores significantly decreased.

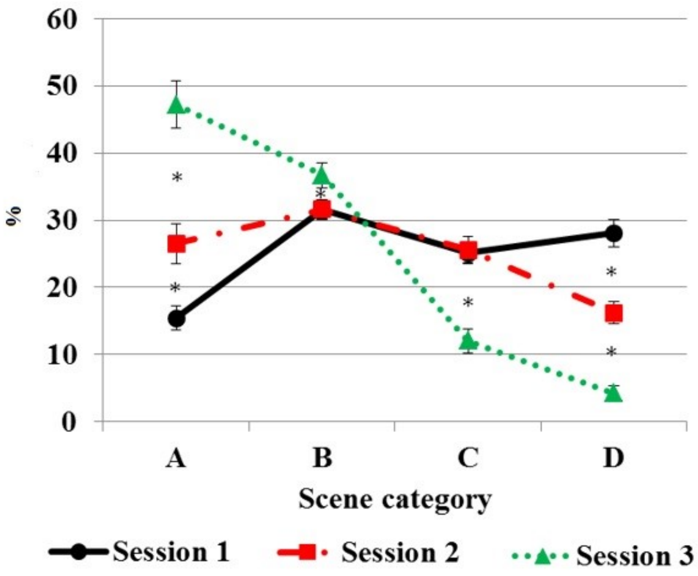

Figure 4: Percentages of the scores obtained in the three sessions. Error bars represent standard errors. Asterisks indicate significant differences at the post hoc tests. Adapted from Tagliabue *et al*. (2019).

Skin conductance responses

In the first of the ANOVAs on the SCR percentages, the factor *Group* and the factor *Session* were significant ($F(1,34) = 17.60$, $p < 0.001$, $\eta_p^2 = 0.34$, and $F(2,68) = 7.47$, $p < 0.01$, $\eta_p^2 = 0.18$, respectively). Participants in the active group showed a higher percentage of SCRs than participants in the passive group (M = 52 *vs*. 27%, respectively). Moreover, the SCR percentages decreased from 47% in the first session to 39% in the second, and 34% in the third. The post hoc indicated that the percentages in the first session were significantly higher ($p < 0.05$) than those in both second and third sessions.

In the second ANOVA on the SCR percentages, the factors *Group* and *Risk degree* resulted significant ($F(1,34) = 20.01$, $p < 0.001$, $\eta_p^2 = 0.37$, and $F(3,102) = 53.22$, $p < 0.001$, $\eta_p^2 = 0.61$, respectively). In addition to the significant difference between the two groups, with the active group showing higher SCR percentages than passive group as in the previous analysis, SCR percentages increased as the scenes became increasingly risky (32% in A scenes, 35% in B scenes, 46% in C scenes, and 60% in D scenes). The SCR percentages in A and B scenes did not differ at the post hoc tests, but both were significantly different from SCR percentages in C ($p < 0.05$) and D scenes ($p < 0.05$) *i.e.*, near misses and crashes. The SCR percentages in C and D scenes were also significantly different from each other ($p < 0.05$).

Moreover, the *Group* × *Risk degree* and the *Session* × *Risk degree* interactions resulted significant ($F(3,102) = 6.01$, $p < 0.01$, $\eta_p^2 = 0.15$, and $F(6,204) = 2.85$, $p < 0.05$, $\eta_p^2 = 0.08$, respectively).

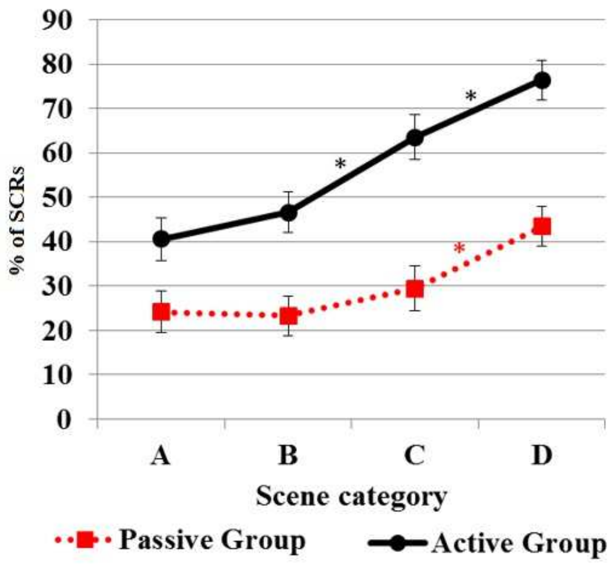

Figure 5: Percentages of SCRs in each scene category, depending on the degree of risk. Error bars represent standard errors. Asterisks indicate the significant differences at the post hoc. The differences between groups are significant in each scene category. Adapted from Tagliabue *et al*. (2019).

Concerning the *Group* × *Risk degree* interaction (Figure 5), the post hoc confirmed the significant difference between groups in each scene category. Additionally, in the active group, the SCR percentages increased in both C and D scenes. This would indicate that, differently from the passive group, the psychophysiological reactivity paralleled the rise in the degree of risk. As for the *Session* × *Risk degree* interaction (Figure 6), comparable SCR percentages were elicited in A and B scenes, but C and D scenes elicited significantly lower SCR percentages ($p < 0.05$). Moreover, SCR percentages in C and D scenes did not differ from each other. In the last two sessions, the SCR percentages were higher in C than in B scenes ($p < 0.05$) and higher in D than in C scenes ($p < 0.05$).

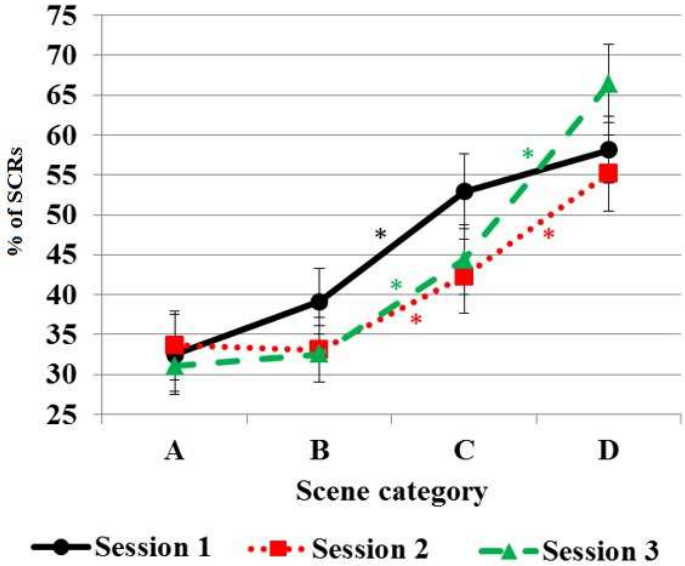

Figure 6: Percentages of SCRs in each scene category, depending on the degree of risk developed. Error bars represent standard errors. Asterisks indicate the significant differences at the post hoc. Adapted from Tagliabue *et al*. (2019).

Finally, in the ANOVA on the SCR onset, only *Session* reached significance ($F(2,68) = 11.94$, $p < 0.001$, $\eta_p^2 = 0.26$). The post hoc showed that the onset in the second session (10 m) was significantly different from the SCR onset in both the first (14 m) and third (16 m) sessions ($p < 0.05$). The SCR onset of the first and third sessions did not differ significantly.

3.2.3 Discussion

The results of the present study confirm previous evidence that showed an improvement in performance as the training on the HRT progresses (Tagliabue & Sarlo, 2015; Tagliabue *et al*., 2017). The decrease of crash rate along the three sessions confirms that the active training on the HRT led not only to a contextual learning (*i.e*., better performance when the same course is administered, as in the first two sessions), but also to a generalization of learning, at least at the behavioral level, as shown by the improvement recorded also in the third session, in which six new courses were faced. The improvement is clear both in the reduction in the accident rate and in the overall performance safety. Indeed, participants who actively drove the simulator showed safer driving behaviors in the third session, attested by the significant reduction in C and D scores.

From a psychophysiological perspective, SCRs to incoming hazards should indicate the activation of an implicit mechanism responsible for hazard perception, in line with the dual processing system model (Slovic *et al*., 2006). Thus, the reduction in the SCR percentage during the training might appear counterintuitive, given that as the training progresses participants should have learned to effectively react to hazards. As explained in Tagliabue *et al*. (2017), this effect may be a consequence of the reduction in the rate of near misses and crashes produced by the training itself. This reduction would make the reaction of the implicit system less necessary, resulting in a decrease in the SCR percentages. The finding of a significant decrease in the C and D scenes (*i.e*., near misses and accidents) in the third session (new courses) supports this explanation.

Moreover, the present data confirm the results of Tagliabue and Sarlo (2015), in that they replicate the results concerning greater emotional involvement in the active group (higher SCR percentages than the passive group). The analysis of the changes in SCR percentages as a function of the scenes' degree of risk indicates the presence of differences between groups even in the modulation of psychophysiological reactivity. Indeed, differently of the active group, the participants in the passive condition showed an increase in the SCR percentages only in scenes where an accident occurred, possibly indicating a failure to discriminate among different levels of hazard. Note that this is an important result since, while driving, the correct "categorization" of the risk degree might simplify the selection of the most appropriate behavior. Moreover, as illustrated in Figure 5 and 6, an accurate discrimination between no- and low-risk scenes (*i.e*., scenes scored A and B, *vs*. C), shown by the psychophysiological reactivity, emerged in the second session and was preserved through the third session. This can be interpreted as an improvement in the ability of transferring the acquired learning to the new scenes, providing a first evidence of generalization in the simulated environment.

Conversely, about anticipatory ability, our data did not confirm the generalization of this ability to new, different scenes, neither in the active nor in the passive group, as shown by the unchanged SCR onset between the first and third session. Nevertheless, the anticipation of the SCR onset in the second session, when the participants faced the same hazardous scenes (as in Tagliabue *et al*., 2017) was confirmed, independently of the group (active or passive). These findings, if confirmed, provide important information about the necessary features for road safety training, as they show the need of a training aimed at exposing novice road users to as many different hazards as possible, taking advantage of the safer conditions of training guaranteed by a simulated environment to increase the likelihood that the users will recognize such hazards in advance once they will be on real road. Moreover, the absence of differences in the anticipatory ability between groups indicate that both the passive and the active training produce comparable results in term of hazard anticipation. However, the results regarding the percentage of SCRs indicate that an active training is more effective in enhancing the ability of discriminating between different levels of on-road hazard than a passive, more classical, training.

To summarize, the participants in the active group showed an improvement in the driving performance (learning), that seemed to generalize to new scenes, as shown by the scores obtained in both the second and third sessions. Moreover, they learned to discriminate among different degrees of risk via their implicit system and to generalize this achievement to the third session. By contrast, the anticipatory ability, in terms of SCR onset, did not generalize to scenes not previously faced. On the other hand, passive training seemed to involve to a lesser extent the implicit system of hazard detection and recognition.

Thus, the main implications of the present research are that (a) experiencing adverse consequences in simulated road scenarios yields an improvement in the ability of early identification of risk when it is faced anew; (b) the psychophysiological correlates of this ability indicate that actively driving in a simulated scenario is more effective than passive viewing of risky video-clips; (c) the anticipatory skill develops on the basis of previous experience, as predicted by the dual model of decision making (Slovic *et al.*, 2006) and by the SMH (Bechara et al., 2000; Damasio, 1994).

The present results, basically, indicate a need to develop training protocols that allow future motorcyclists to be exposed to the largest possible sample of road hazards before starting to drive on the real road, confronting each potential road risk. Thus, every effort to map the greatest possible variations of the most common accidents dynamics so as to replicate them via the simulators, must be firmly supported, such as the development of learning protocols able to enhance the likelihood of transferring the same anticipatory ability to unexperienced scenes, so as to promote generalization to several categories of hazards.

The principal limitation of the present research is related to the generalizability of the observed effects to real on-road conditions. Although the use of simulators for driving assessment and training is spreading in several Countries, the pros and cons of this strategy, as seen, remain controversial and need to be tested in terms of generalizability to the real road. Nevertheless, the present data suggest that, although during the training of novice road users the generalization is evident from a behavioral perspective, implicit learning requires previous exposure to each specific scenario. Therefore, any attempt to improve training programs, by monitoring and mapping conditions under which accidents occur and by exposing novice road users to the largest possible variety of risky scenarios, should be firmly encouraged.

3.3 Electrophysiological correlates of attentional monitoring during a driving training in a simulated environment

As seen, an active driving training on the HRT improves hazard perception, leading to a safer driving performance and to the development of somatic markers processes as a function of the risk degree of the simulated scenarios. However, as previous evidence pointed out (Tagliabue *et al*., 2013; Wester *et al*., 2008; 2010), driving safely also requires a high level of attentional resources. The amount of available resources influences the ability to monitor the environment and detect changes. Moreover, both top-down and bottom-up factors seem to interact in determining the amount of resources available for attentional monitoring (Corbetta & Shulman, 2002; Lavie *et al*., 2014). Because attentional monitoring has a key role in everyday life and in driving itself, we believe it is important to consider also possible changes in attentional deployment consequent to a driving training, going beyond the previous focus of the literature on the identification of differences between stimulus processing in single and dual tasks conditions applied to the driving context (see Wester *et al*., 2008; Wester *et al*., 2010).

Thus, we extended the design of Wester *et al.* (2008), who administered both passive and active oddball paradigms during a simple driving task (see Chapter 2). In the present study (Gianfranchi *et al*., submitted), we wanted to assess the changes in attentional monitoring during a driving task in a simulated environment by recording participants' EEG activity while administering an auditory, passive multi-feature oddball paradigm. The aims of the study were: (1) To assess changes in attentional monitoring, as indexed by changes in its electrophysiological correlates (N1, MMN, and P3a amplitudes) during a driving task performed simultaneously to an auditory passive oddball task, as a function of the experience acquired in performing the driving task; (2) To investigate whether these changes show any modulation depending on the salience of the oddball stimuli (*i.e.*, their degree of consistency with the experimental driving context).

Concerning the first aim, we hypothesized that, as long as the training proceeds, more attentional resources should be available to monitor the experimental environment. Thus, the processing of the deviant oddball stimuli should increase and should be reflected by an increase in the amplitudes of the attentional ERP components at the end of the training.

To reach the second aim, we decided to employ three categories of deviant stimuli in the multi-feature oddball paradigm: Deviant pure tones, human sounds, and traffic-related sounds. We reasoned that the increased processing of the three categories of deviant stimuli should show a modulation depending on their salience. Specifically, traffic-related sounds should elicit a larger P3a in the second block, indicating that the cognitive system is more prone to identifying events whose meaning is related to the driving context, although task-irrelevant.

Finally, to control for differences between an active, more demanding task (*i.e.*, driving) and a passive, less demanding one, we introduced a control condition (*i.e.*, watching a video-clip of a course on the simulator while being presented with the passive oddball task) at the beginning of the procedure. This also allows the comparison of our results with those of previous studies (*e.g.*, Wester *et al.*, 2008). We assumed that, all things being equal (*i.e.*, being exposed to a passive oddball task), driving requires a higher amount of resources than watching a video-clip with hazardous scenes. Therefore, we expected to find higher amplitudes of the ERP components triggered by the deviant stimuli (N1, MMN, and P3a) in the control as compared to the driving condition. This result would be in line with previous studies attesting an amplitude modulation of these components as a function of the amount of resources needed to face different types of dual tasks, although unrelated to driving context (for a review, see Kok, 1997).

3.3.1. Material and methods

Participants

Twenty-four students from the University of Padua were enrolled (12 M; mean age: 24.5 years; range: 19-28). All of them had held a valid driver's license for at least 18 months and declared on-road exposure of at least 5,000 km/year. All the participants reported no history of neurological, neuropsychiatric disorders or drug consumption, and they all had normal or corrected-to-normal vision. The participants were not paid for their participation.

Written informed consent was obtained from all participants before the beginning of the experimental procedure. The project was approved by the Ethical Committee for Psychological Research of the University of Padua.

EEG recording and apparatus

The EEG signal was recorded through a Micromed SD MRI 64 system (Micromed/Treviso, Italy), amplified and digitized with a sampling frequency of 512 Hz. A 32-channel Electro-cap was used, in accordance with the 10-20 International System (Jasper, 1958), with Ag/AgCl electrodes referenced to the linked mastoids. All electrode impedances were less than 10 KΩ and balanced.

The Micromed SD MRI 64 system was connected to a PC Intel® Core™ i5-2500 with a Windows 7 Professional operating system, which presented the oddball stimuli using the E-prime task-presentation software (2.0.10 version, Psychology Software Tools). Two speakers were placed on a table behind the participants, approximately 40 cm away. These speakers reproduced the oddball stimuli presented by the E-prime software with a fixed intensity of 75 dB (see Wester *et al.*, 2008). We decided to use external speakers to allow the participants to listen to the instructions that the HRT gave on the path to follow and to the oddball stimuli at the same time.

Stimuli and procedure

At the beginning of the experiment, the participants were invited to sit on the HRT simulator, where the EEG cap was applied. The participants completed a practice course on the simulator (3 minutes without traffic and with no instructions on the path to follow). Then, the experimental procedure started.

As a first step, all the participants sat on the HRT and watched the same video-clip of a first-person course on the simulator (control condition, 4 minutes in duration). Then, the driving task began. The six courses used for the task were divided according to their degree of difficulty, to create two comparable blocks. In each block, three courses were ordered following the possible permutations, obtaining six sequences per block. The order of administration of the blocks was counterbalanced between subjects. The final number of possible sequences was 12, each administered to two participants (one male and one female).

During both the first and second block of the driving task, the participants had to drive along three simulated courses on main roads with traffic, following the same instructions described in section 3.1.1. Each course lasted approximately 4 minutes, depending on the speed and on the driving style of each participant. Thus, each block lasted for about 12 minutes. Both after the control condition and the first block, a 3-min rest was scheduled to prevent the effect of tiredness and workload. Then, the participants had to face the last three courses (second block) on the simulator.

In each phase of the procedure (*i.e.*, control condition, first block, and second block), the participants were presented with a multi-feature oddball paradigm (Pakarinen *et al.*, 2007). The standard tones (pure tones of 1000 Hz, 100 ms in duration) were presented in 64% of the trials. Three types of deviant stimuli (500 ms in duration) were used: Deviant pure tones (1300 Hz), human sounds (*i.e.*, a cough, adult's laugh, child's laugh, a scream) and traffic-related sounds (*i.e.*, a hard brake, a horn, a tire screech, a bicycle bell). Each category of deviants was presented in 12% of the trials. The stimuli were taken from the International Affective Digitalized Sounds database (IADS-2 for European Portuguese; Soares, Pinheiro, Costa, Frade, Comesaña, & Pureza, 2013). All stimuli were matched for loudness. The inter-stimulus interval (ISI) was 500 ms. Thus, the duration of a trial was 600 ms in the case of standard stimuli and 1 s in the case of deviants.

Seven lists of stimuli (one for the control condition and six for the courses on the HRT) were built up, pseudo randomizing them on the basis of the advice by Pakarinen *et al.* (2007). In an array of three deviants, each deviant category was presented once, and two successive deviants were always of different categories. The presentation of the oddball task was synchronized with the beginning of the control video-clip or, in the driving condition, with the beginning of the course on the simulator. The seven lists were always presented in the same order. Importantly, the participants were explicitly instructed to ignore the oddball stimuli and to focus on their main task (*i.e.*, watching the video-clip during the control condition, and driving the simulator during the two blocks of the driving task).

Behavioral analysis

We evaluated the driving performance in terms of Evaluation score (*i.e.*, the A, B, C, and D scores) and accidents, assigning a value of 1 to the A score, 2 to B, 3 to C, and 4 to D, so that a higher score corresponded with a worse performance. Moreover, for each course, we computed the percentages of hazardous scenes in which participants avoided the crash (number of scenes without crash*100/total number of scenes).

Then, we ran two 2 × 3 separate repeated-measures analyses of variance (ANOVAs) on the percentages of avoided accidents and on the mean evaluation scores (*i.e.*, means of the scores obtained for each scene of the simulator, from 1 to 4), with *Block* and *Course* as within-subject factors. The α level was set at 0.05. All the behavioral analyses were conducted with the IBM SPSS 23 statistical software package.

EEG preprocessing and ERP analysis

All EEG recordings were processed offline using EEGLAB v 14.1.1b, a MATLAB toolbox (Delorme & Makeig, 2004). First, the signal was bandpass filtered between 0.1 and 30 Hz. Then, it was divided into epochs starting 100 ms before and ending 600 ms after stimulus onset. EEG epochs were baseline corrected over the 100-ms interval preceding the stimulus onset.

We applied a multistep procedure for artifact handling. Epochs were visually inspected to interpolate bad channels and to detect and eliminate large artifacts (*e.g.*, large head movements) by rejecting the contaminated epochs. Then, an independent component analysis (ICA; Delorme, Sejnowski, & Makeig, 2007) was run on large artifact-free data. All independent components were visually inspected, and, according to their morphology and scalp distribution, only those related to eye blinks and eye movements were discarded. The remaining components were then projected back to the electrode space, obtaining uncontaminated EEG epochs. As a final step, we further applied an artifact rejection procedure aimed at removing all epochs with an amplitude voltage exceeding a threshold of $\pm$ 100 μV.

The total number of artifact-free trials per experimental phase was as follows: Control condition (M = 313.79; SD = 20.56), first block (M = 809.71; SD = 48.41), and second block (M = 804.08; SD = 82.04). The number of accepted trials did not significantly differ between the two experimental blocks ($t(23) = 0.504$, $p > 0.05$).

Average ERP waveforms were computed separately for each stimulus condition (standard tones, deviant pure tones, human sounds, and traffic-related sounds) and experimental phase (control condition, first block, and second block) for each electrode site. The MMN was computed as the difference in waves obtained by subtracting the average ERP of the standard tones from the average ERP of each deviant category (deviant pure tones, human sounds, and traffic-related sounds).

The analysis was focused on specific temporal windows concerning the target ERP components, namely, the N1 (120-200 ms), MMN (120-200 ms) and P3a (200-350 ms). The temporal windows were selected on the basis of the visual inspection of the average signal in each condition (in the case of the MMN on the differential waves), so as to include the peak. Then, we applied a whole-brain analysis (30 electrodes) to the targeted ERP components (N1, MMN, P3a), using a two-tailed t-test (α = 0.05) permutation approach (1000 Monte-Carlo permutations). To control for the 1st-type family error in both the spatial and temporal dimensions, the false discovery rate (FDR) correction for multiple comparisons was applied (Benjamini & Hochberg, 1995; Lage-Castellanos, Martìnez-Montes, Hernanzed-Cabrera, & Galàn, 2009). The critical t-score for each contrast was ± 2.07.

The ERP components were statistically compared in pairs of experimental conditions. All the statistical analyses were performed via the Brainstorm Software (v3.4 – Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011). To compare our results with those of previous literature (Wester *et al*., 2008), we collapsed the two driving blocks conditions to create a single condition called driving condition. Then, we contrasted ERP activity in the control *vs*. driving condition, at first regardless of the category of the deviant stimuli, and then by separately inspecting the ERP components elicited by each category of deviant (deviant pure tones, human sounds, and traffic-related sounds). Subsequent comparisons were run contrasting the first *vs*. the second block of driving for each deviant stimulus.

3.3.2 Results

Driving performance

In the ANOVA performed on the percentages of avoided accidents, only the *Block* main effect reached significance ($F(1, 23) = 5.23$, $p < 0.05$, $\eta_p^2 = 0.19$). The participants showed a significant increase in the percentages of avoided accidents from the first to the second block of courses (M = 90.8% *vs.* 93.8% respectively). The *Block* main effect was also significant in the ANOVA performed on the mean Evaluation scores ($F(1, 23) = 25.64$, $p < 0.001$, $\eta_p^2 = 0.53$). This result shows that participants reported significantly better (*i.e.*, lower) Evaluation scores in the second block, going from a mean score of 1.91 *vs.* 1.71 in the first three courses and in the last three respectively, thus testifying to the effectiveness of the training. No other sources of variance reached significance.

Event-related potentials

As for the ERP data, in the first contrast between the control and the driving condition (considering all deviants collapsed), the permutation analysis confirmed our hypothesis regarding the N1 and P3a components. Larger negativity was shown during the control as compared to the driving condition between 120 and 150 ms after the stimulus onset ($p < 0.01$). This effect was observed in a cluster of electrodes covering mainly centro-parietal scalp locations. Moreover, greater positivity was shown between 250 and 300 ms ($p < 0.01$), this modulation occurring in a cluster of fronto-central electrodes (Figure 7). Regarding the MMN, no significant differences were found between the two conditions.

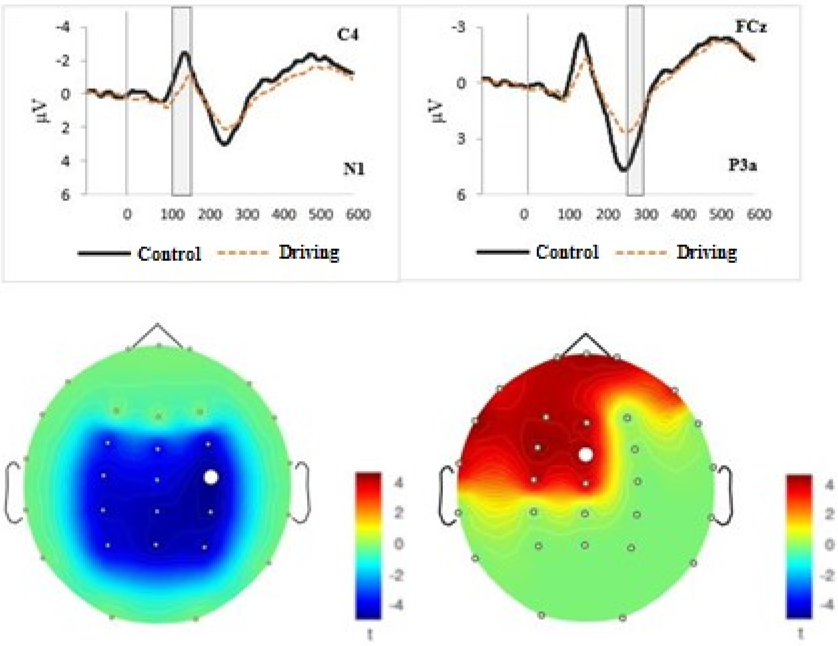

Figure 7: ERP activity regardless of the category of the deviants in the control and in the driving condition. The upper panels show the ERP activity in the electrode with the greatest effect among those exceeding the critical t-score threshold for statistical significance. Significant groups of electrodes are shown below each waveform plot for the first (left) and the second (right) significant time window (both highlighted), corresponding to the N1 and P3a components, respectively. The white dots represent the electrode with the largest amplitude modulation. Adapted from Gianfranchi *et al*. (submitted).

In the second analysis, the differences in the ERPs elicited by the three categories of deviants in the control *vs*. driving condition were considered. Using the same statistical procedure as in the first analysis, we compared the ERP components elicited by each category of deviant stimuli in the control condition and in the two collapsed blocks of the driving condition (Figure 8). Concerning the deviant pure tones, no significant differences were found between the control and the driving condition. Regarding the traffic-related sounds, greater negativity was observed for the N1 during the control than the driving condition in the 131- to 160-ms time window ($p < 0.01$), mostly in centro-parietal electrodes, and for the MMN in the 141- to 160-ms time window ($p < 0.05$) at central electrodes. Moreover, the P3a showed a higher amplitude during the control than the driving condition in the 200- to 300-ms time window ($p < 0.01$) at centro-frontal scalp locations. Finally, concerning the human sounds, the P3a showed a greater positivity in the control as compared to the driving condition in the 250- to 300-ms time window ($p < 0.05$) in a group of centro-frontally distributed electrodes.

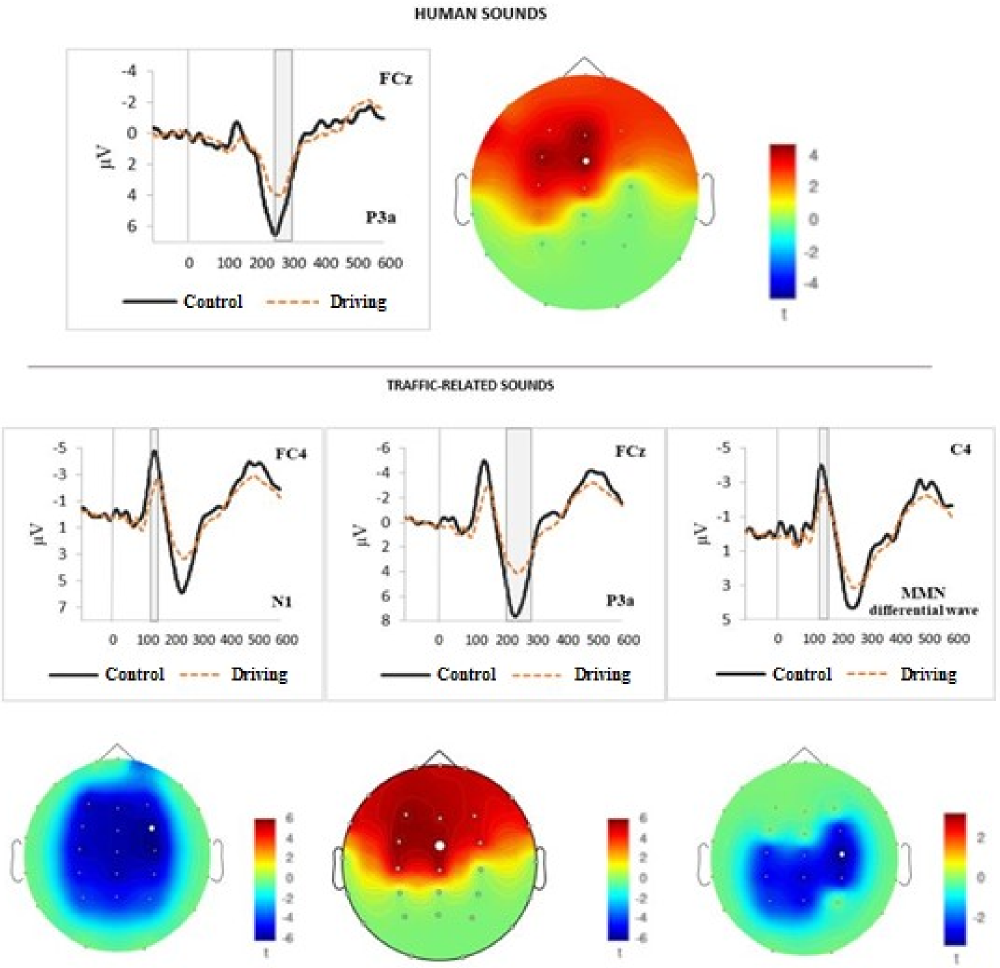

Figure 8: ERP activity for the human sounds and the traffic-related sounds in the control and driving conditions. For the human sounds (upper panels), the left panel shows the ERP activity in the electrode with the greatest effect among those exceeding the critical t-score threshold for statistical significance. The significant group of electrodes is shown on the right of the waveform plot for the significant time window, corresponding to the P3a component. For the traffic-related sounds (center and bottom panels), the upper panels display the ERP activity in the electrode showing the greatest effect among those exceeding the critical t-score threshold for statistical significance. Significant groups of electrodes are shown below each waveform plot for the first (left), second (center), and third significant time window (all highlighted), corresponding to the N1, P3a, and MMN (differential wave) components, respectively. The white dots represent the electrode showing the largest amplitude modulation. Adapted from Gianfranchi *et al.* (submitted).

We finally compared the ERPs elicited in the first *vs.* the second driving block separately for each category of deviant. Regarding the deviant pure tones, greater negativity was shown for the MMN in the second block as compared to the first. The significant time window was between 160 and 189 ms ($p < 0.05$), involving a group of centro-parietal electrodes (Figure 9). Neither the N1 nor the P3a showed significant differences for the deviant pure tones between the two blocks. Concerning the traffic-related sounds, greater positivity was found for the P3a in the second block, between 275 and 285 ms ($p < 0.05$) in fronto-central scalp locations (Figure 9). No significant differences were found between blocks in the ERPs elicited by the human sounds.

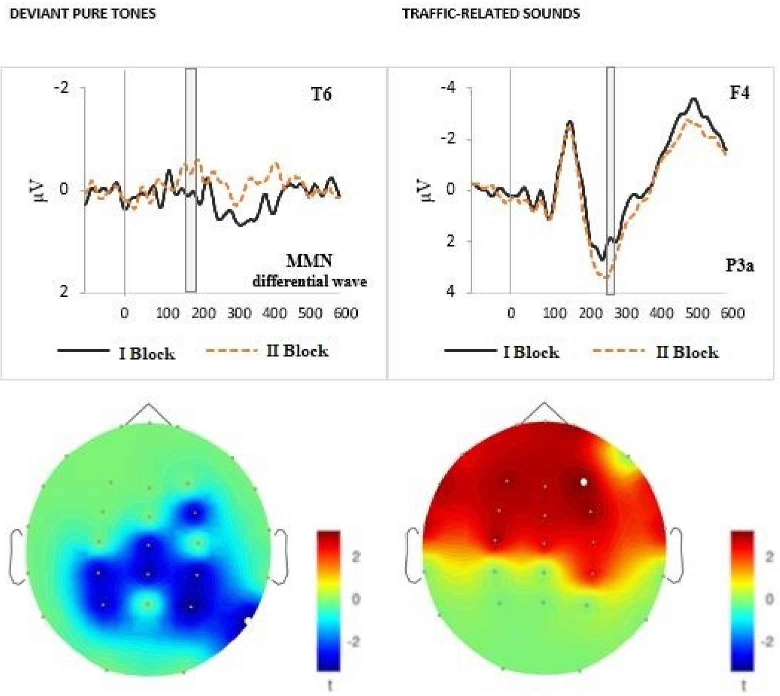

Figure 9: ERP activity for the deviant pure tones (left panels) and the traffic-related sounds (right panels) in the two blocks of the driving condition. The upper panels show the ERP activity in the electrode with the greatest effect among those exceeding the critical t-score threshold for statistical significance. Significant groups of electrodes are shown below each waveform plot for the first (left) and second (right) significant time window (both highlighted), corresponding to the MMN component (differential wave) for the deviant pure tones and the P3a component for the traffic-related sounds. The white dots represent the electrodes showing the largest amplitude modulation. Adapted from Gianfranchi *et al*. (submitted).

3.3.3 Discussion

Considering that bottom-up and top-down aspects influence the attentional monitoring of the environment (*e.g.*, Corbetta & Shulman, 2002; Folk *et al*., 1992) and that both the amount of available resources and stimulus salience contribute in determining the ability to identify and process infrequent stimuli (Corbetta & Shulman, 2002; Lavie *et al*., 2004), the present study aimed at identifying changes in the electrophysiological correlates of attentional monitoring and shifting as a function of learning during a complex task, such as driving in a simulated environment. Starting from previous evidence (Crundall *et al*., 2002; Tagliabue *et al*., 2013; Underwood *et al*., 2003), we hypothesized that learning to drive safely on the HRT would lead not only to an increase in the amount of available attentional resources, but also to a change in the way they are distributed across different types of environmental stimuli. The changes should be paralleled by modifications in the ERPs and could be modulated by the salience of the stimuli, meaning the degree of consistency with the experimental (driving) context.

We expected to find a behavioral improvement in the driving performance from the first to the second block of courses. The results confirm that the participants showed a safer driving performance, avoiding a higher percentage of crashes and having better evaluation scores in the second block of courses on the simulator. This is in line with previous studies (Tagliabue & Sarlo, 2015; Tagliabue *et al*., 2017; Tagliabue *et al*., 2019; Vidotto *et al*., 2011) and indicates that the participants gained experience through the task, learning to drive more safely. This improvement should result in a higher amount of available attentional resources for monitoring salient stimuli. Therefore, it is possible to predict an increase in the amplitude of the ERPs elicited by the deviant oddball stimuli during the second block of courses.

First, we compared the ERPs in the control and driving conditions, independently of the category of deviants. This allowed us to assess the differences in the attentional ERPs elicited by the oddball stimuli during a passive simple task (watching a video-clip) compared with an active and more demanding task (driving a simulator). As expected, this first analysis shows that the amplitudes of the N1 and P3a components elicited by the deviant oddball stimuli were significantly greater during the control condition than during driving. However, for the MMN no differences were present. These results are in line with the hypothesis that the participants had fewer resources to employ in monitoring the experimental environment when involved in a driving task. This is supported by the evidence of reduced neural resources allocated to stimuli processing. Specifically, the reduction in the amplitude of the N1 to the deviant stimuli suggests that participants devoted less neural activity to overall detection of potentially relevant information. Coherently, the P3a modulation (low amplitude) during the driving simulation suggests that the cognitive system was less prone to switching attention toward potentially relevant stimuli.

In other words, in both early and late computational stages triggered by the deviant stimuli, neural activity was greater during an effortless dual task (*i.e.*, passively watching a video-clip while being presented with oddball stimuli) than during a high-demanding dual task (*i.e.*, driving while being presented with oddball stimuli). This is also in line with the findings of Wester *et al*. (2008), who reported a reduction in the P3a in the driving task but no effects on the MMN in an active dual-task condition (lane keeping on a simulator and passive novelty oddball) as compared to a passive single-task condition (passive novelty oddball alone). Our results indicate that the same effects can be found comparing two dual tasks (one active and one passive) that require different degrees of cognitive load. Thus, we can infer that during the driving task, participants had fewer resources for both stimulus detection (N1) and attentional shifting toward distracting stimuli (P3a), regardless of their features. This is also in line with the theory of Lavie *et al*. (2004), which postulates that distraction, defined as non-relevant stimuli processing, is reduced when cognitive load increases.

The subsequent comparisons between the ERPs recorded during the control and the driving condition were focused on the presence of category-specific differences between the three types of deviant stimuli and the standard one. These analyses show that a reduction in the N1 during the driving task was present only for the traffic-related sounds, along with a reduction of the MMN. Finally, the P3a exhibited a smaller amplitude in the driving condition for both the traffic-related and the human sounds. No differences were found for the deviant pure tones, suggesting that they were processed in the same way in the two experimental conditions.

Overall, these results support the hypothesis that an active dual task needs more attentional resources to be performed than a passive one. During driving, participants generally exhibit both a reduced detection of the traffic-related task-irrelevant sounds (reduction in the N1) and a reduced sensitivity to change detection (reduction in the MMN). In addition, they are less efficient in shifting attention toward meaningful stimuli, as attested by the reduced P3a for both traffic-related and human sounds.

To summarize, first, the traffic-related sounds seem to trigger stimulus detection, change detection, and attentional switching because of their consistency with the experimental driving context (*i.e.*, salience). Second, the three components related to the three processes just mentioned are modulated by the cognitive load: In the control condition, the amplitudes of the N1, MMN, and P3a components were larger due to the greater amount of resources available. Third, this cognitive-load modulatory effect on the P3a component was evident for the human sounds too. This is in line with the findings of Horváth *et al.* (2008), indicating that the P3a is sensitive to social or emotional stimuli that would be processed to a greater extent in the control condition, when more resources are available.

Finally, concerning the learning effects, we expected to find an amplitude increase in the second block of driving, maybe modulated by the salience of the deviant stimuli. Our results show that in the second block of simulated courses, the deviant pure tones elicited a larger MMN, whereas the traffic-related sounds elicited a larger P3a, confirming that, after participants acquired experience with the task and learned how to drive more safely, more attentional resources become available to process the oddball deviant stimuli. Specifically, the increase in the amplitude of the

MMN for the deviant pure tones and of the P3a for the traffic-related sounds suggests that the participants detected more easily the change in these categories of stimuli.

The finding of the increase in the MMN in the second block only for the deviant pure tones might be surprising. One explanation may derive from the structure of the driving training administered to the participants. Previous studies (Tagliabue *et al*., 2017; Tagliabue *et al*., 2019; Vidotto *et al*., 2011) showed that complete driving training on this simulator is achieved during at least two sessions with multiple courses. Thus, it is possible to hypothesize that a single session including only six courses does not allow enough free resources for monitoring, detecting changes, and shifting attention. The consequence would be a "strategic" use of the resources, mirrored by better change detection for the less complex stimuli (*i.e*., MMN for the deviant pure tones), an increase in the readiness to shift attention toward the more complex and more salient stimuli (*i.e*., P3a for the traffic-related sounds), and an unchanged processing of the complex but less salient stimuli (*i.e*., human sounds). Employing more courses (or more sessions) on the HRT might lead to an increase in the amplitude of other components too.

Thus, the main limitation of the present study lies in the number of courses on the HRT simulator, which might not be enough to achieve complete driving training, allowing participants to have enough attentional resources to deeply monitor the experimental environment. Moreover, the variety of driving experience of the participants needs to be controlled in future research, and a wider sample should be employed. The practical implications of the present findings, besides confirming the need to administer multi-session driving trainings on the HRT, range from the identification of potential differences in the attentional monitoring among several categories of drivers (*e.g*., patients with attention-deficit/hyperactivity disorder, elderly people, recidivists) to the development of training and evaluative protocols of attention while driving, based on EEG activity.

Chapter 4

The identification of driving profiles in a simulated environment

The second branch of the present Ph.D. project deals with the development of an assessment method to identify different *driving* profiles (instead of the most common identification of *personality* profiles) among drivers with different degrees of on-road exposure, through the HRT simulator. In this chapter, two studies are reported: The first (section 4.1; Gianfranchi, Spoto, & Tagliabue, 2017a) reversed for the first time the classical approach of the identification of personality profiles, attempting to identify driving profiles based on the performance at the HRT. The two emerged profiles resulted differently related to self-reported aberrant driving behaviors. The second study (section 4.2; Gianfranchi, Tagliabue, Spoto, & Vidotto, 2017b) applied the same methodology on a different sample, attempting also to assess the links among the identified driving profiles, SS, and decision-making ability, as measured by the IGT (Bechara *et al*, 1994; 2000).

4.1 The identification of driving profiles and their relationship with aberrant on-road behaviors

The present study[3] (Gianfranchi *et al*., 2017a) arises from the need to find an integrated methodology for assessing novice road users' abilities, so as to overcome the limits of the classical approach, which consists in the identification of personality profiles through self-report measures and in the assessment of relations between these profiles and self-reported driving behaviors and attitudes. Conversely, in the present study participants' performance was monitored through the HRT simulator and recorded in terms of driving variables that were used to cluster participants in

[3] An extended version of section 4.1 can be found in Gianfranchi, E., Spoto, A., & Tagliabue, M. (2017). Risk profiles in novice road users: relation between moped riding simulator performance, on-road aberrant behaviors and dangerous driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *49*, 132-144.

groups with different driving profiles. These profiles were subsequently compared in terms of differences in the scores at the Driver Behavior Questionnaire (DBQ; Reason *et al*., 1990) and the Dula Dangerous Driving Index (3DI; Dula & Ballard, 2003). We expected to find various clusters differing in driving performance and representing specific on-road risk levels. We also expected to find relations between these clusters and the DBQ and 3DI scores. If our expectations were met, they would support the implementation of an integrated approach (*i.e.*, detailed records of simulated driving performance and self-reported measures) as a suitable tool for identifying driving profiles among novice road users.

4.1.1 Material and methods

Participants

Ninety-one students in psychology at the University of Padua were recruited (64 F; mean age: 19.61 years; range: 18-24 years). Eighty-six had a car driving license and 5 only had a moped riding license. Fifty-two participants (57.14% of the sample) were first-time riders (*i.e.*, they had never ridden a two-wheeled vehicle before). Of the 39 participants who had previously ridden a moped or motorcycle, none spent more than 2h/week driving motor vehicles. Mileage over the previous year was also collected: Sixty-six students declared less than 10,000 km, and 25 declared more than 10,000 km. All participants were paid 25 euros for their participation in the study. They all had normal or corrected-to-normal vision.

Tools

The HRT simulator was employed for the assessment of driving performance. Moreover, two self-report measures were used to assess participants' predisposition to commit on-road errors and violations, and to drive unsafely: The DBQ and the 3DI questionnaires (Dula & Ballard, 2003; Reason *et al*., 1990).

Given that our sample included some participants who had never driven a car, all 28 original items of the 3DI were used for this study, whereas the original 50-items version of the DBQ (Reason *et al*., 1990) were adapted - following an approach already reported in the literature (*e.g*., Steg & Van Brussel, 2009). Essentially, DBQ items relating to car driving (*e.g*., ''Lock yourself out of the car with the keys still inside") were omitted. The final number of items in the DBQ was 42. The DBQ subscales were: Slips and Lapses (15 items); Errors (8 items); Intended Violations (16 items); Unintended Violations (3 items); and the total score. The Slips and Lapses subscale represent attention and memory failures (Lucidi *et al*., 2010) considered quite harmless, with no influence on driving safety (Steg & Van Brussel, 2009), with items such as ''Fail to notice someone stepping out from behind a bus or parked vehicle until it is nearly too late". The Errors subscale consists of items such as ''Plan your route badly, so that you meet traffic congestion you could have avoided". Such errors are defined as planned action failures (Lucidi *et al*., 2010). Violations are generally defined as deviations from proper behavior that also imply a strong motivational component (Lucidi *et al*., 2010). In particular, the Intended Violations subscale refers to deliberate violations of traffic legislation (''Deliberately disregard the speed limits late at night or very early in the morning"), whereas the Unintended Violations subscale concerns violations due to being distracted (''Lost in thought or distracted, you fail to notice someone waiting at a zebra crossing or a pelican crossing light that has just turned red"). Each item needs an answer on a six-point Likert scale, indicating the frequency of each behavior, from 0 = never, to 5 = nearly all the time (Reason *et al*., 1990). The total score ranged from 0 to 210.

The 3DI includes 28 items on a five-point Likert scale ranging between A (scored as 1) and E (scored as 5) to indicate the frequency of the investigated behaviors. The items are grouped into three subscales (Aggressive Driving, Risky Driving, Negative Emotional Driving) and the total score. The Aggressive Driving subscale (7 items) concerns all types of behavior exhibited by a driver intended to damage other road users physically or psychologically (‘‘When someone cuts me off, I feel I should punish him/her”). Risky Driving (12 items) refers to risky behaviors not originally intended to harm people, although potentially harmful (‘‘I will illegally pass a car that is going too slowly”). The Negative Emotional Driving subscale (9 items) refers to emotions and cognitions not intended to harm, such as anger, frustration or irritation (‘‘I lose my temper when driving”; Dula & Ballard, 2003). The total score ranges from 28 to 140.

An Italian version of the DBQ is currently available (Smorti & Guarnieri, 2016), but it was published after the completion of data collection for the present study, and it was based on a reduced, 27-item version of the questionnaire (Martinussen, Lajunen, Møller, & Özkan, 2013), whereas we adopted the original version of the DBQ to ensure the exploration of a wider range of behaviors. Both the DBQ and the 3DI were translated into Italian using the back-translation method, with the support of a bilingual collaborator.

Procedure

The whole procedure consisted in two sessions. The sample was randomly split into two groups to control for the effects of the order of the tasks in the two sessions. One group (N = 47) rode the HRT during the first session and answered the questionnaires during the second session a few days later; The other group (N = 44) was administered with the opposite order. At the beginning of the first session, all participants were told about the procedures of the study (but not about its specific aims) and signed an informed consent form. The task on the HRT (duration: Around 45 minutes) consisted of one practice course and five test courses on secondary roads. Participants were given practical instructions on how to use the simulator and asked to ride along the courses as if they were on a real road, following the instructions received from the HRT speaker and trying to avoid collisions. They rested for about three minutes after completing each course to avoid any fatigue effect. The study was approved by the Ethical Committee for Psychological Research of the University of Padua.

Data coding and analyses

Participants' driving performance was assessed along all five test courses using the output .csv files automatically generated by the simulator. As we will see, the analysis of the driving profiles was conducted in depth, through a wide number of indices, for two reasons. First, all the variables related to a participant's driving performance on the simulator needed to be examined as a first, innovative step towards the development of a complete driving assessment protocol. Second, it has been demonstrated that speed and acceleration on the simulator relate to self-reported driving behavior (*e.g.*, Lucidi *et al*., 2010), but other indices might deserve consideration as well.

Thus, eighteen indices were extracted for all the risky scenes in the five courses: The final scores represent the mean values of all these indices along the courses. The indices included the means and standard deviations of throttle opening (measured as a percentage), front and rear pressure on the brakes (measured in kg), speed (measured in km/h), on-road instability (measured as horizontal deviations from the right side of the road), number of braking actions, accidents, and sections of the course where participants exceeded the speed limit. For the speed limit, we also considered the time spent over the speed limit (the sum of the frames over the speed limit), and the mean and maximum overspeeding values. As overall performance indicators, we also extracted the number of accidents that occurred along the courses and the mean Evaluation score, computed assigning a score of 1 to all the scenes evaluated as A, 2 to B, 3 to C and 4 to D, so as that higher values corresponds to a worse, more risk-taking performance. Aberrant and dangerous driving behaviors were assessed by means of the above-described DBQ and 3DI subscales and total scores.

A cluster analysis was run, using the HRT parameters as grouping variables, employing the Ward's hierarchical clustering method with the squared Euclidean distance measures. The standardized scores (Z-scores) of the grouping variables were used. A K-means cluster analysis was run specifying the number of clusters as emerged by the Ward's method in order to select the most appropriate cluster solution for our sample. The chi-square test was used on the frequencies of participants' mileage in the clusters to check whether the cluster partition was affected by mileage. As a final step, a binary logistic regression was run to clarify which variables really discriminate between the two clusters. The HRT indices were first inspected using a bivariate Pearson's correlation test to seek any correlations among them and prune the pairs of variables with correlations higher than 0.50, which could cause collinearity. Then, the binary logistic regression was run with the HRT indices as independent variables and the cluster solution as the dependent variable, to explore the relative importance of the variables selected. A stepwise backward procedure was used.

In order to evaluate the relations between participants' riding profiles and their DBQ and 3DI scores, two separate 2 × 2 × 2 × 2 Multivariate analyses of variance (MANOVAs) with *Group* (*i.e.*, the cluster solution), *Gender*; *Driving Exposure* (2 levels: Less than *vs*. more than 10,000 km/year), and *Order* of the tasks (2 levels: Questionnaires completed before *vs*. after using the HRT) were run on all the subscales and the totals of the DBQ and 3DI. Analyses were performed using the IBM SPSS 22 statistical package. The post hoc analyses were run using Bonferroni's correction.

4.1.2 Results

The driving profiles

From the inspection of the merging coefficients and of the dendrogram, two clusters emerged. One included 65 participants (46 F; mean age: 19.61 years), whereas the second included 26 participants (18 F; mean age: 19.49 years). The chi-square test confirmed that the two clusters were comparable in terms of annual mileage ($\chi^2(1) = 0.006$, $p = 0.941$), with 72.3% *vs.* 73% of participants with an annual mileage of less than 10,000 km in the first and in the second cluster respectively. As shown in Figure 10, the clusters exhibited two opposite trends on the HRT indices, indicating opposite driving behaviors.

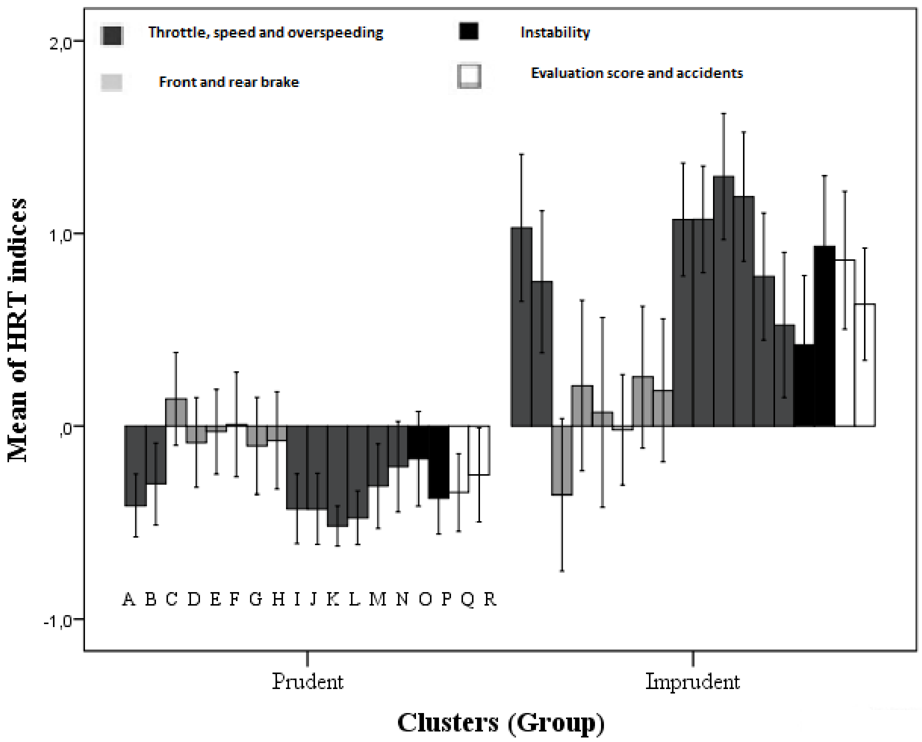

Figure 10: Trend of the two clusters on the 18 indices of the HRT (Z-scores). The variables are: Mean of throttle opening (A) and its Standard deviation (B); Number of breakings with the front brake (C); Mean (D) and Standard deviation (E) of its pressure; Number of breakings with the rear brake (F); Mean (G) and Standard deviation (H) of its pressure; Mean (I) and Standard deviation (J) of speed; Time spent over the speed limit (K); Number (L), Mean (M), and the highest value of overspeeding (N); Mean (O) and Standard Deviation (P) of on-road instability; Number of accidents (Q); Evaluation score (R). Vertical bars represent standard errors. Adapted from Gianfranchi *et al.* (2017a).

The predictors (Z-scores) included in the subsequent binary logistic regression (after inspecting for correlations) to clarify which variables really discriminate between the two clusters were Number of front and rear braking, Mean of pressure of both brakes, Mean of speed, Mean of overspeeding, Mean of on-road instability, Highest value of overspeeding, and Evaluation score.

The final model (χ^2 (4) = 67.61, $p < 0.001$) is reported in Table 1. Overall, 89% of the 91 participants were correctly classified by the regression model. The variables relating to speed and overspeeding significantly predicted inclusion in the second cluster, meaning that higher scores for these variables were more likely in the second cluster. Based on these results, participants in Cluster 1 have been defined as ‘‘Prudent drivers,” while those in Cluster 2 have been labeled as ‘‘Imprudent drivers”.

Variables	β	Wald χ^2(df)	p value
Mean of pressure of the front brake	−0.15	2.94 (1)	0.086
Mean of speed	0.56	12.93 (1)	<0.001
Mean of overspeedings	0.18	3.82 (1)	0.050
Highest value of overspeedings	0.17	4.27 (1)	0.039
Intercept	−2.60	13.89 (1)	<0.001

Table 1: The final regression model. Adapted from Gianfranchi *et al*. (2017a).

Driving profiles and aberrant driving behaviors (DBQ)

The multivariate-level results of the first MANOVA on the DBQ subscales and total score show that, among the principal effects, *Gender*, *Driving Exposure* and *Order* reached significance (Wilk's λ = 0.74, $F(4,73) = 6.32$, $p < 0.001$, $\eta_p^2 = 0.26$, Wilk's λ = 0.88, $F(4,73) = 2.53$, $p < 0.05$, $\eta_p^2 = 0.12$, and Wilk's λ = 0.85, $F(4,73) = 3.24$, $p < 0.01$, $\eta_p^2 = 0.15$ respectively). Concerning the two-way interactions, the *Group* × *Order* and *Gender* × *Driving Exposure* reached significance at the multivariate level (Wilk's λ = 0.87, $F(4,73) = 2.84$, $p < 0.05$, $\eta_p^2 = 0.14$, and Wilk's λ = 0.85, $F(4,73) = 3.24$, $p < 0.01$, $\eta_p^2 = 0.20$ respectively). Finally, the three-way interaction *Group* × *Gender* × *Driving Exposure* was significant (Wilk's λ = 0.82, $F(4,73) = 4.10$, $p < 0.01$, $\eta_p^2 = 0.18$).

For the sake of simplicity, the univariate results will be listed separately for each subscale.

For the Intended Violations subscale, only the *Gender* factor was significant ($F(1,76) = 5.93$, $p < 0.05$, $\eta_p^2 = 0.07$), with males scoring higher than females (M = 13.04 *vs*. 8.62). For the Errors subscale, the main effects of both *Order* and *Gender* reached significance ($F(1,76) = 10.33$, p <

0.01, $\eta_p^2 = 0.12$, and $F(1,76) = 10.40$, $p < 0.01$, $\eta_p^2 = 0.12$, respectively). Overall, participants who answered the questionnaire before driving the simulator scored higher than those who rode the simulator first (M = 7.23 *vs*. 6.73), and females scored higher than males (M = 7.64 *vs*. 5.44). The *Group* × *Order* interaction was significant too ($F(1,76) = 10.81$, $p < 0.01$, $\eta_p^2 = 0.38$). Imprudent drivers who answered the questionnaire after using the simulator scored lower for Errors than Imprudent drivers who answered the questionnaire first (M = 6.11 *vs*. 10.75; $p < 0.01$ at the post hoc comparisons). The *Gender* × *Driving Exposure* interaction was also significant ($F(1,76) = 15.82$, $p < 0.001$, $\eta_p^2 = 0.17$). Males with a higher mileage scored lower for Errors than females with the same mileage or males with a lower mileage. The significance of the three-way *Group* × *Gender* × *Driving Exposure* interaction ($F(1,76) = 11.61$, $p < 0.01$, $\eta_p^2 = 0.13$) indicated that the abovementioned pattern of results only occurred in Imprudent drivers. Indeed, Imprudent males with higher on-road exposure scored lower for Errors than Imprudent females with the same exposure (M = 3.40 *vs*. 11.00; $p < 0.01$ at post hoc), and lower than Imprudent males with a lower exposure (M = 3.40 *vs*. 10.50; $p < 0.01$).

For the Slips and Lapses subscale, no main factors reached significance. However, the *Group* × *Order* interaction reached significance ($F(1,76) = 5.39$, $p < 0.05$, $\eta_p^2 = 0.07$). Among participants who answered the questionnaire before using the simulator, Prudent drivers scored lower than Imprudent (M = 11.82 *vs*. 18.60; $p < 0.05$). The *Group* × *Gender* × *Driving Exposure* interaction was significant as well ($F(1,76) = 5.63$, $p < 0.05$, $\eta_p^2 = 0.07$). As for Errors, among Imprudent drivers, males who drove more than 10,000 km/year scored lower for Slips and Lapses than Imprudent females with the same exposure (M = 10.60 *vs*. 17.00; $p < 0.05$ at post hoc tests) and than Imprudent males with a lower exposure (M = 10.60 *vs*. 20.50; $p < 0.05$).

As regards the Unintended Violations subscale, only the *Group* × *Gender* × *Driving Exposure* interaction reached significance ($F(1,76) = 7.04$, $p < 0.05$, $\eta_p^2 = 0.09$), and the trends were the same as for Errors, and for Slips and Lapses. All the significant differences concerned the

Imprudent drivers: Imprudent males with a higher on-road exposure scored lower than Imprudent females with the same exposure ($M = 2.80$ *vs*. 7.00; $p < 0.01$), and lower than Imprudent males with a lower exposure ($M = 2.80$ *vs*. 6.25; $p < 0.05$).

Concerning the total DBQ score, no main source of variance was significant. However, the *Gender* × *Driving Exposure* ($F(1,76) = 5.82$, $p < 0.05$, $\eta_p^2 = 0.07$) and *Group* × *Order* ($F(1,76) = 5.44$, $p < 0.05$, $\eta_p^2 = 0.07$) interactions were significant. The means of the *Group* × *Order* interaction resembled those seen for the Slips and Lapses subscale, indicating that Prudent drivers who answered the questionnaire before using the simulator scored lower than Imprudent drivers completing the two sessions in the same order ($M = 31.95$ *vs*. 49.30; $p < 0.05$), while no significant differences emerged between participants who performed the simulator task first. The three-way *Group* × *Gender* × *Driving Exposure* interaction reached significance for the total scores too ($F(1,76) = 6.11$, $p < 0.05$, $\eta_p^2 = 0.07$). The lower scores for the Imprudent males with a higher driving exposure than for the Imprudent males with a lower driving exposure ($M = 33.40$ *vs.* 53.50; $p < 0.05$) were confirmed in the total scores as well.

Driving profiles and dangerous driving (3DI)

The multivariate-level results of the second MANOVA on the 3DI subscales and total score show that, among the principal effects, only *Driving Exposure* reached significance (Wilk's $\lambda = 0.90$, $F(3,74) = 2.83$, $p < 0.05$, $\eta_p^2 = 0.10$). Among the two-way interactions, *Group* × *Driving Exposure*, *Group* × *Order*, and *Gender* × *Order* were significant (Wilk's $\lambda = 0.88$, $F(3,74) = 3.28$, $p < 0.05$, $\eta_p^2 = 0.12$, Wilk's $\lambda = 0.89$, $F(3,74) = 2.99$, $p < 0.05$, $\eta_p^2 = 0.11$, and Wilk's $\lambda = 0.89$, $F(3,74) = 3.14$, $p < 0.05$, $\eta_p^2 = 0.11$ respectively). Finally, the three-way interaction *Group* × *Gender* × *Driving Exposure* reached significance (Wilk's $\lambda = 0.90$, $F(3,74) = 2.82$, $p < 0.05$, $\eta_p^2 = 0.10$), whereas the *Gender* × *Driving Exposure* × *Order* interaction resulted marginally significant at the multivariate level (Wilk's $\lambda = 0.91$, $F(3,74) = 2.58$, $p = 0.06$, $\eta_p^2 = 0.10$).

The univariate results will be listed separately for each subscale. Moreover, two results that reached significance at the univariate level but not at the multivariate one will be discussed because of their theoretical relevance, *i.e.* the *Gender* effect for the Risky Driving subscale and the *Group* × *Gender* × *Order* interaction for the Aggressive Driving.

No main source of variance or interactions were significant for the Negative Emotional Driving scores.

For the Risky Driving subscale, the main effect of the *Gender* factor reached significance only at the univariate level ($F(1,76) = 4.74$, $p < 0.05$, $\eta_p^2 = 0.06$), with higher scores for males than for females ($M = 20.22$ *vs*. 17.09). Moreover, the *Group* × *Order* interaction reached significance ($F(1,76) = 4.48$, $p < 0.05$, $\eta_p^2 = 0.06$). The Imprudent drivers who answered the questionnaires after using the HRT tended to report more risky driving behaviors than the Imprudent drivers who answered the questionnaires first, even though the difference was only marginally significant (Figure 11).

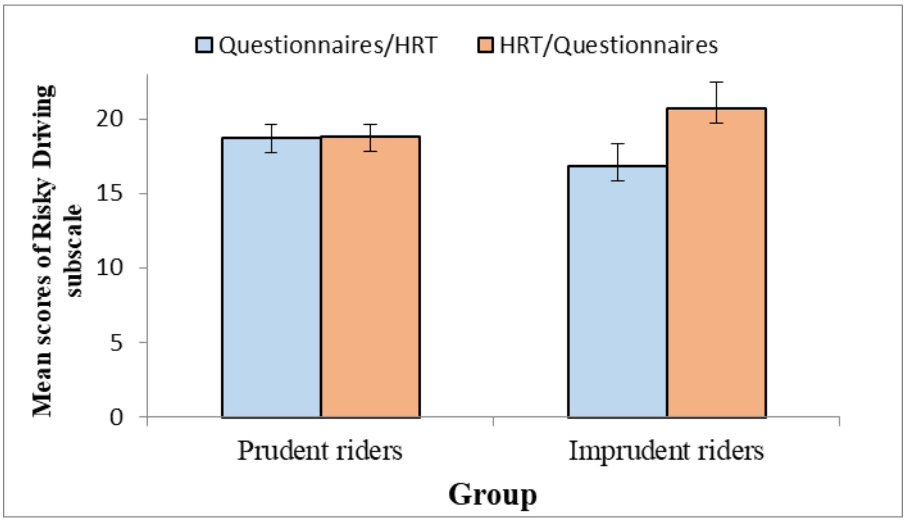

Figure 11: The *Group* × *Order* interaction for Risky Driving subscale in the 3DI questionnaire. Vertical bars represent standard errors. Adapted from Gianfranchi *et al*. (2017a).

For the Aggressive Driving subscale, the *Driving Exposure* factor reached significance ($F(1,76) = 7.82$, $p < 0.01$, $\eta_p^2 = 0.09$): The scores were higher for participants with a higher mileage

(M = 12.44 *vs*. 11.21). However, *Driving Exposure* interacted significantly with *Group* ($F(1,76) = 8.95$, $p < 0.01$, $\eta_p^2 = 0.11$): Imprudent drivers who drove more than 10,000 km/year scored higher than Imprudent drivers with a lower mileage (M = 14.67 *vs*. 10.25; $p < 0.05$). The *Gender* × *Order* ($F(1,76) = 4.27$, $p < 0.05$, $\eta_p^2 = 0.05$) and *Group* × *Gender* × *Order* ($F(1,76) = 4.57$, $p < 0.05$, $\eta_p^2 = 0.06$) interactions reached significance at the univariate level for this subscale too. Though none of the post hoc comparisons reached significance and the three-way interaction was not significant at the multivariate level, the trends seem to indicate that, among participants who answered the questionnaires first, Imprudent males scored lowest and Imprudent females scored highest. Concerning the *Gender* × *Driving Exposure* × *Order* interaction ($F(1,76) = 7.30$, $p < 0.01$, $\eta_p^2 = 0.08$), among the females who answered the questionnaires first, those who drove more than 10,000 km/year scored higher than those with a lower mileage (M = 15.00 *vs*. 10.62; $p < 0.05$).

Finally, concerning the total 3DI score, *Driving Exposure* reached significance ($F(1,76) = 5.50$, $p < 0.05$, $\eta_p^2 = 0.07$), replicating the result obtained for Aggressive Driving, in that participants with a higher mileage reported higher scores (M = 55.88 *vs*. 51.74). As shown by the *Group* × *Driving Exposure* interaction ($F(1,76) = 7.11$, $p < 0.01$, $\eta_p^2 = 0.09$), this is true only for the Imprudent group: Imprudent drivers who drove more than 10,000 km/year scored higher than Imprudent drivers who drove less than 10,000 km/year (M = 62.13 *vs*. 49.80, $p < 0.05$). Finally, the *Gender* × *Driving Exposure* × *Order* interaction ($F(1,76) = 4.30$, $p < 0.05$, $\eta_p^2 = 0.05$) should be considered with caution due to small sample size in some conditions that may be the cause of the marginality of the result at the multivariate level. The trends indicated that males who used the simulator first and with a high mileage had higher total scores than males who used the simulator first but had a low mileage (M = 62.33 *vs*. 52.87). However, no comparisons resulted significant at post hoc tests.

4.1.3 Discussion

The present study represents a first step toward the development of an assessment protocol of novice road users' driving profile by means of the HRT simulator. The main aim of the present study was the identification of driving profiles in a sample of novice road users, with various degrees of on-road exposure, through the detailed assessment of their driving style. Moreover, reversing the classical methodology that aimed to identify personality profiles differently characterized by specific driving behaviors, we attempted to study the association between participants' driving profiles and their scores in two questionnaires regarding aberrant and risky on-road behaviors. We expected to find different clusters based on the performance with the simulator, and thus to discriminate between different driving styles associated with specific patterns of DBQ and 3DI scores.

As regards the identification of driving styles, the cluster analysis on participants' driving performance revealed two groups with opposite patterns: Prudent and Imprudent drivers. A binary logistic regression confirmed that driving at higher speed and exceeding the speed limit significantly predicted the inclusion in the Imprudent profile. Participants' behavior was assessed on multiple driving indices to explore the potential of the HRT as a driving assessment tool. The high correlations found among the indices considered have two important implications. First, they confirm the vehicle response validity, which corresponds to the internal validity (Shechtman, *et al.*, 2009), of the HRT. Second, this collinearity ensures a good flexibility in the choice of variables that can be included to assess driving behavior. Consequently, it is possible to conclude that the HRT can be potentially employed to identify driving profiles on the basis of just a few predictors.

Concerning the relations between driving profiles, DBQ and 3DI, our findings replicate reports in the literature of gender-related and experience-related differences. Indeed, in the DBQ, females declared more driving errors than males, whereas males scored higher on the Intended Violations subscale (see Reason *et al.*, 1990) and tended to score higher on the Risky Driving subscale of the 3DI (Dula & Ballard, 2003). These gender effects are supposed to be a consequence of a self-attributed gender stereotype (Dula & Ballard, 2003). The mileage seems to play a role in the Aggressive Driving and 3DI total scores, which rise as the annual mileage increases, as reported also by Richer and Bergeron (2012).

Interestingly, although no differences were found between the two clusters *per se*, the *Group* factor interacted with the *Order*, showing opposite trends for the DBQ and 3DI subscales. Specifically, among the Imprudent drivers, those who answered the questionnaires after using the simulator scored lower for Slips and Lapses, for Errors, and in their total score in the DBQ. However, they scored higher on the Risky Driving subscale of the 3DI. Taken together, these findings suggest that the questionnaires investigate different dimensions of risky driving behaviors. Indeed, the results regarding the underestimation of likelihood of errors and lapses among imprudent drivers who used the HRT before answering the questionnaire, are in line with the results of a recent study (Vidotto, Tagliabue, & Tira, 2015), in which participants' self-evaluations of their own driving safety were worse after the HRT training. This effect should be consequence of participants' improvement in general hazard awareness, due to their previous exposure to the simulator. The reversal of this effect for DBQ subscales might be linked to the different construct behind the questionnaire: Indeed, Imprudent drivers might only see risky driving (but not errors/lapses) as potentially dangerous.

The *Group* factor significantly interacted also with the *Driving Exposure*, but only for the 3DI total and the Aggressive Driving subscale: Among the Imprudent drivers, those with a higher mileage also scored higher on these subscales, suggesting that aggressive and dangerous driving behaviors are connected not only to driving exposure but also to driving style. Imprudent drivers may be more influenced by the effects of driving exposure (such as stress, tension, need to keep high attentional levels) because of their driving style, and this might lead, in turn, to a greater predisposition to behave more aggressively and dangerously on the road.

With respect to the role of *Gender*, its interaction with *Driving Exposure* indicates that mileage has a different effect on males and females in terms of their reported scores of Errors. Specifically, males with a high on-road exposure reported fewer driving errors. This is in line with the literature (Reason *et al*., 1990) indicating that males tend to report fewer driving errors. In addition, de Winter and Dodou (2010) noted that the Errors score on the DBQ tends to decrease with driving experience. In our sample, this effect seems stronger for males than for females, maybe as the result of an overestimation of more experienced male drivers' own capability, linked to a gender stereotype. This is confirmed by the significantly higher score on the Intended Violations scale of males in our sample. Moreover, these trends are also modulated by driving style, in that the effects just described have been observed only in Imprudent drivers also for Slips and Lapses, and for Unintended Violations. Moreover, although it needs cautious consideration due to the failure in reaching significance at the multivariate level, the different trends found between males and females in terms of Risky Driving may suggest that males are more prone to report risky driving behaviors than females. This result, although in line with the literature (*e.g*., Jonah 1997; Jonah *et al*., 2001) will need further consideration in the next studies.

As regards Aggressive Driving subscale, in the Imprudent cluster, males who answered the questionnaires before using the HRT tended to report less aggressive behaviors, while females reported more aggressive behaviors. This effect is inconsistent with previous reports of males usually scoring higher overall in the 3DI (Dula, Adams, Miesner, & Leonard, 2010). Future studies need to employ a larger and better stratified sample size to help clarifying this point. Indeed, the small sample size and the low number of participants included in some conditions prevent us from drawing firm conclusions on this point.

However, some conclusions can be drawn based on all the other results. First, to the best of our knowledge, this is the first study in which participants were clustered on the basis of a large number of variables extracted from participants' performance with a simulator, rather than on the basis of multiple psychological dimensions. Our results show that this methodology can be used to assess driving styles thoroughly, differentiating between safe and risky behaviors, as indicated by our cluster analysis and binary logistic regression. The fact that more clusters are often described in the literature (*e.g.*, Lucidi *et al.*, 2010; Marengo *et al.*, 2012) may be a consequence of differences in sample sizes. Second, the fact that gender and driving exposure were found to be associated with aberrant driving behaviors and dangerous driving, as measured by the DBQ and 3DI, confirms the need to consider these variables as well when assessing an individual's driving abilities. This is a further proof of the need of an integrated methodology, allowing to assess at the same time multiple behavioral and psychological variables.

Some limitations need to be considered. First, the sample size, although generally adequate, prevented us from drawing firm conclusions on the three-way interactions. Second, although we translated and adapted the items using the back-translation method, and the results in terms of gender and driving experience replicate previous findings, a factorial validation of the questionnaires' structure might help to clarify further aspects of our findings possibly associated with cultural variables (that we did not consider). Moreover, further investigations should consider pruning some of the simulator parameters, as suggested by the results of the regression. An important contribution of this study comes from its focus on the characteristics of novice road users. Indeed, most self-report instruments for assessing driving abilities and risky behaviors are intended for experienced drivers/riders. This is a crucial limitation that needs to be overcome, in order to develop specific assessment and training protocols. The relations found between the simulator indices (and the clusters identified), aberrant driving behaviors, and dangerous driving further support the value of VR, jointly with self-report measures, for the assessment and training of drivers' abilities in road accident prevention programs.

4.2 Sensation seeking and decision-making ability: Their influence on driving behaviors

As seen in the previous section, the HRT can be successfully employed as a tool to identify different driving styles among novice road users. Moreover, these styles seem to be differently related to self-report measures of aberrant and risky driving behaviors (Gianfranchi *et al*., 2017a). The same methodology was applied in the study reported in the present section[4] (Gianfranchi *et al*., 2017b). Here, we attempted to assess the relationships between the identified driving styles and two aspects that proved to be crucial when it comes to road safety: Sensation seeking (SS), and decision-making ability as measured by the IGT (Bechara *et al*., 1994; 2000).

The general aim of the present study was to provide a more integrated framework of the relationships among SS, decision-making ability, and driving behaviors by exploring their mutual relations. Starting from the previous evidence that driving profile is characterized by several aspects concurring in determining the quality of performance, and that, as we saw, driving behaviors can be clustered so as to identify different driving profiles, we reasoned that it could be interesting and, above all, useful to investigate whether different driving profiles (as measured by the HRT simulator) were characterized also by different levels of SS and by differences in the decision-making ability.

The first prediction we made was that different driving profiles should be linked to different levels of SS. In particular, based on the literature (*e.g*., Jonah, 1997), we expected the involvement of the thrill-seeking dimension of SS. Then, in the next step, we reasoned that, as reported by Jonah (1997), "sensation seeking may account for only ca. 10–15% of the variance in risky driving" (Jonah, 1997; p. 660). This means that considering just SS level is not enough when assessing risky

[4] An extended version of section 4.2 can be found in Gianfranchi, E., Tagliabue, M., Spoto, A., & Vidotto, G. (2017). Sensation seeking, non-contextual decision making, and driving abilities as measured through a moped simulator. *Frontiers in Psychology*, *8*, 2126.

driving behaviors. Therefore, we reasoned that, among the wide variety of risky driving predictors, decision-making ability is one of the most interesting.

Decision-making ability and hazard perception are deeply related. As seen, the somatic marker mechanisms underpin hazard perception (Tagliabue & Sarlo, 2015; Tagliabue *et al*., 2017; Tagliabue *et al*., 2019) and influences the decision-making ability in risky or ambiguous contexts (Bechara *et al*., 1994; 200; Damasio, 1994). Moreover, although sometimes controversial, previous evidence suggest that worse decision-making skills correspond to riskier on-road behaviors, assessed either with self-report tools or with driving simulators (*e.g*., French, West, Elander, & Wilding, 1993; Farah, Yechiam, Bekhor, Toledo, & Polus, 2008). On the other hand, it seems that SS (particularly its thrill-seeking dimension) may have a role in decision-making tasks, such as the IGT (Buelow & Suhr, 2013). Thus, our second prediction was the presence of a correlation between thrill seeking and the performance at the IGT. The last step was the investigation of the combined effects of SS and decision-making on driving parameters. Thus, the third prediction was to find modulatory influences of sensation-seeking levels and quality of decision-making (as measured by the IGT) on risky behaviors.

4.2.1 Material and methods

Participants

One hundred and thirty-one students (89 F; mean age: 19.78 years; range: 18-29 years) at the University of Padua took part in the study. Both license types and mileage were assessed, although only 111 participants answered the questions regarding these factors. Specifically, 60 reported a mileage of less than 5,000 km/year, while 51 declared a mileage of more than 5,000 km/year. Overall, 53 participants drove a car, 33 drove both a car and a two-wheeled vehicle, and 5 participants rode only a moped. All of the students were paid 25 euros for their participation and had normal or corrected-to-normal vision.

Tools

The driving performance was assessed with the HRT simulator.

To measure the SS, we employed the Italian version of Zuckerman's Sensation Seeking Scale V (SSS V; Galeazzi *et al*., 1993; Zuckerman, 1994), including 40 forced choices that investigate various behaviors considered representative of high or low levels of SS. The questionnaire includes four subscales with 10 items each: Thrill and Adventure Seeking (TAS – "I would like to try parachute jumping"), Disinhibition (DIS – "I often like to get high, drinking liquor or smoking marijuana"), Boredom Susceptibility (BS – "I get bored seeing the same old faces"), and Experience Seeking (ES – "I like to try new foods that I have never tasted before"). Each scale has a maximum score of 10 and the total score ranges from 0 to 40.

To test participants' decision-making ability, we administered a computerized version of the IGT (Bechara *et al*.; 1994; 2000). The participants started with an imaginary monetary amount of 2,000 euros and were told that their goal was to maximize their profit, by picking a card at time from one of the four decks displayed on the PC screen (intel Core i3 PC, Windows 7 operating system, 1920 × 1080 resolution). As in the original task, each card had a monetary win, but some cards could have also a loss, whose amount depended on the decks: The decks were either disadvantageous (decks A and B) or advantageous (decks C and D. There were no time limits. The participants were not informed of the total number of choices they had to make (100) or of the differences between the decks.

Procedure

All the participants filled in an informed consent form and were told about the study procedures before the beginning of the experiment. The procedure consisted of two sessions scheduled a few days apart from each other. All the participants filled in the SSS V and a questionnaire regarding driving experience and exposure. Then, the sample was randomly split into two groups to counterbalance the order of the tasks. One group (N = 66) drove the HRT during the first session of the experiment and performed the IGT during the second session, while the other group (N = 65) performed the tasks in the opposite order. The driving task (about 45 minutes) included five test courses on secondary roads, preceded by one practice course. The transmission was set to automatic. The participants were instructed on how to use the controls to ride the simulator. They were told to ride along the courses as if they were on the road, respecting the traffic rules and avoiding collisions, following the instructions of the HRT regarding the path to follow. After each course, a rest period of 3 min was scheduled to prevent fatigue. The study was approved by the Ethical Committee for the Psychological Research of the University of Padova.

Data coding and analyses

The driving performance was monitored through the .csv files extracted from the HRT. The same 18 indices already described in Gianfranchi *et al*. (2017a) were computed (see section 4.1.2).

For the SSS V, the original scoring instructions were followed.

Concerning the IGT, the participants were expected to start choosing randomly from the four decks, slowly shifting their choices toward the advantageous decks as the task progressed (see Bechara *et al*., 1994; 2000). Following the literature, we divided the task into 5 blocks of 20 trials each to assess the participants' behavioral changes during the IGT (Bechara *et al*., 1999, 2000; Bechara & Martin, 2004). For each block, we first computed the so-called Net Score (Bechara *et al*., 1999; Bechara & Martin, 2004), obtained by subtracting the sum of the disadvantageous selections from that of the advantageous ones. Higher scores correspond to a higher number of choices from the advantageous decks, indicating better decision-making ability. Then, as an overall measure of performance in the IGT, we computed an index that we called the Goodness of Decision-Making (GDM) score, computed as the mean Net Score obtained by participants in the last three blocks, since the last 40–60 choices refer to decision-making under risk (Bechara & Martin, 2004; Buelow & Suhr, 2013). Participants who reached GDM scores higher than the median of the sample were labelled as good decision makers, whereas the other participants as bad decision makers.

Overall, the statistical analyses were divided into two steps. The first dealt with identifying the driving profiles through a cluster analysis (following the same steps of Gianfranchi *et al*., 2017a). The second focused on the relationships among these profiles, SS, and decision-making ability. Concerning this second step, three directions were followed. First, in line with our first prediction, we tested differences in SS and decision-making ability as depending on the different driving profiles identified through a MANOVA on the dependent variables SSS V subscales and GDM score. Then, Spearman correlations were computed between GDM and SSS V (second prediction). Finally, to test our third prediction, we carried out a second MANOVA with the significantly correlated scores as independent variables, to identify any influence of SS and decision-making ability on all the indices of the driving performance. The analyses were conducted using the IBM SPSS 23 statistical package, with α level set at 0.05 and Bonferroni correction for the post hoc tests.

4.2.2 Results

The driving profiles

We performed a cluster analysis on the 18 HRT indices following the same procedure described in Gianfranchi *et al*. (2017a) and reported in section 4.1.1. The final solution identified three clusters with different driving patterns (Figure 12).

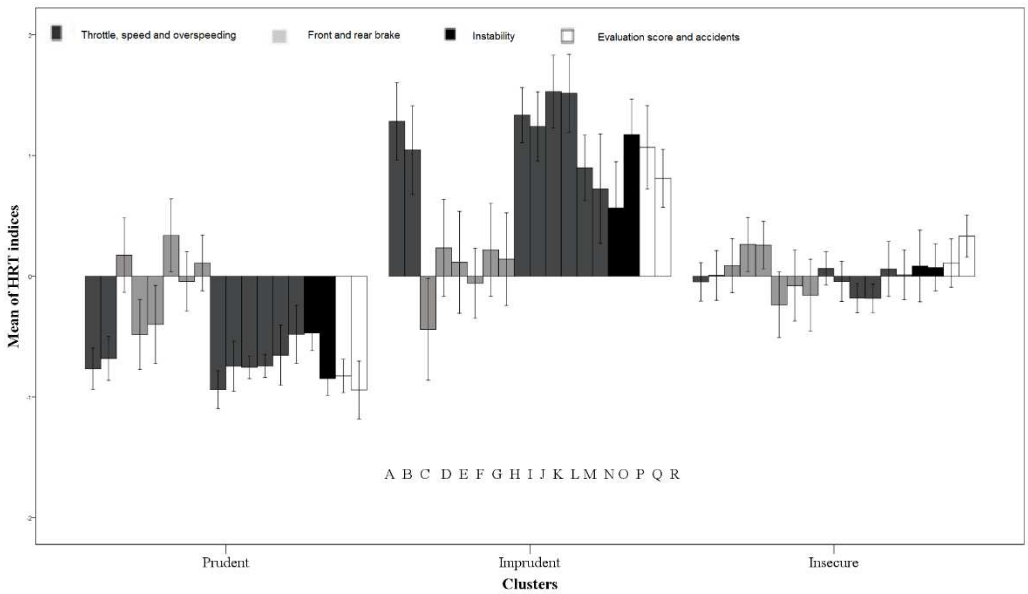

Figure 12: Trends of the three clusters on the 18 HRT indices. The indices are reported for all three clusters in the order displayed by the letters on the bottom of the panel. The indices are: Mean of the throttle opening (A) and its Standard deviation (B), Number of breakings with the front brake (C), Mean (D), and Standard deviation (E) of front brake pressure, Number of brakings with the rear brake (F), Mean (G), and Standard deviation (H) of rear brake pressure, Mean (I) and Standard deviation (J) of speed, Time spent over the speed limit (K), Number (L), Mean (M), and the highest value (N) of overspeeding, Mean (O) and Standard Deviation (P) of on-road instability, Number of accidents (Q) and Evaluation score (R). Vertical bars represent standard errors. Adapted from Gianfranchi *et al*. (2017b).

The first cluster, named "Prudent," included 45 participants (31 F; mean age: 19.45 years). Among them, 23 had an annual mileage of less than 5,000 km. The second cluster, composed by 29 students, was labelled "Imprudent" (16 F; mean age: 20.07 years). Twelve of them declared an annual mileage of less than 5,000 km. The third cluster, labeled “Insecure,” included 57 participants (mean age: 19.9 years old; 42 F). Twenty-five students declared less than 5,000 km/year. Although the missing information regarding the on-road exposure for 20 participants, a chi-square test confirmed that there were no significant differences between the clusters in terms of annual mileage (χ^2 (2) = 0.603, p = 0.740). Each cluster showed a peculiar trend in HRT performance, with a clear opposite trend between the first and second clusters, while the third can be described as intermediate. Specifically, the second cluster is characterized by the highest values in almost all the indices, indicating less safe behavior (*e.g.*, higher speed and acceleration rate). “Prudent drivers” reported very low values for speed, acceleration, and accidents and the best (*i.e.*, the lowest) Evaluation score, whereas the “Insecure drivers” seemed to be characterized by an intermediate behavioral pattern compared to that of the others. In addition, “Insecure drivers” showed some elements of insecurity, such as a trend of braking suddenly, even though their speed was lower compared with that of the Imprudent participants.

Driving profiles, sensation seeking and decision-making

After the identification of the driving profiles, the next step was to compare the HRT performance with the scores on the SSS V and on the IGT. First, we run a 2 × 2 × 2 MANOVA on the four subscales of the SSS V and on the GDM score, with three between factors: *Group* (*i.e.*, the clusters; 3 levels), *Gender*, and *Road Exposure* (*i.e.*, above or below the cutoff value of 5,000 km/year; 2 levels). The analysis was run on 111 subjects because of the missing data regarding the annual mileage. The multivariate results revealed significant effects for the *Gender* (Wilks’ λ =0.84, $F(5,95) = 3.64$, $p < 0.01$, $\eta_p^2 = 0.16$) and for the interaction *Group* × *Gender* (Wilks’ $\lambda = 0.82$, $F(10,190) = 2.02$, $p < 0.05$, $\eta_p^2 = 0.10$). Univariate results indicated that *Gender* reached significance for the BS subscale ($F(1,99) = 8.76$, $p < 0.01$, $\eta_p^2 = 0.08$). Males reported higher BS scores than females (M = 3.96 *vs*. 2.81). Moreover, at the univariate level, the interaction *Group* ×

Gender reached significance for the TAS subscale ($F(2,99) = 4.45$, $p < 0.05$, $\eta_p^2 = 0.08$) and for the DIS subscale ($F(2,99) = 4.91$, $p < 0.01$, $\eta_p^2 = 0.09$). Insecure males showed lower TAS scores than Insecure females (M = 4.60 *vs*. 6.70; $p < 0.05$ at the post hoc comparisons) and Prudent males (M = 7.90; $p < 0.01$). On the other hand, Prudent males showed higher DIS scores than Prudent females (M = 5.87 *vs*. 4.26, $p < 0.05$), and Prudent females showed lower DIS scores than Insecure females (M = 4.26 *vs*. 5.60; $p < 0.05$).

Concerning the relations among driving profiles, SS and IGT, we found a significant negative correlation (Table 2) between TAS and GDM score ($r = -0.24$, $p < 0.01$), indicating that high thrill and adventure seekers also showed worse decision-making abilities.

SSS V	GDM score	*p*-values
TAS	**- 0.24**	**0.006**
ES	0.09	0.282
DIS	−0.03	0.696
BS	−0.03	0.702
Total score	−0.09	0.284

Table 2: Spearman correlations between SSS V scales and GDM score. In bold the significant correlation. Adapted from Gianfranchi *et al*. (2017b).

Thus, based on the TAS scores, we identified high and low thrill and adventure seekers as the participants who got lower or higher TAS scores respectively, in comparison with the median of our sample. Analogously, we labeled the participants who got greater or lower GDM scores than the median of the sample as good decision makers and bad decision makers respectively.

As a final step, to explore the combined influence of SS and decision- making on driving performance, we carried out a 2 × 2 MANOVA on all 18 HRT variables with two between participant factors: *TAS* (2 levels) and *GDM* (2 levels). At the multivariate level only the interaction *TAS* × *GDM* reached significance (Wilks' $\lambda = 0.76$, $F(18,110) = 1.89$, $p < 0.05$, $\eta_p^2 = 0.24$).

Univariate results indicated that the interaction *TAS* × *GDM* was significant for Number of accidents and mean Evaluation scores ($F(1,127) = 4.06$, $p < 0.05$, $\eta_p^2 = 0.03$, and $F(1,127) = 5.03$, $p < 0.05$, $\eta_p^2 = 0.04$, respectively). As depicted in Figure 13, high thrill and adventure seekers who were also bad decision makers had significantly more accidents and showed riskier riding simulator performance than high thrill and adventure seekers who were good decision makers.

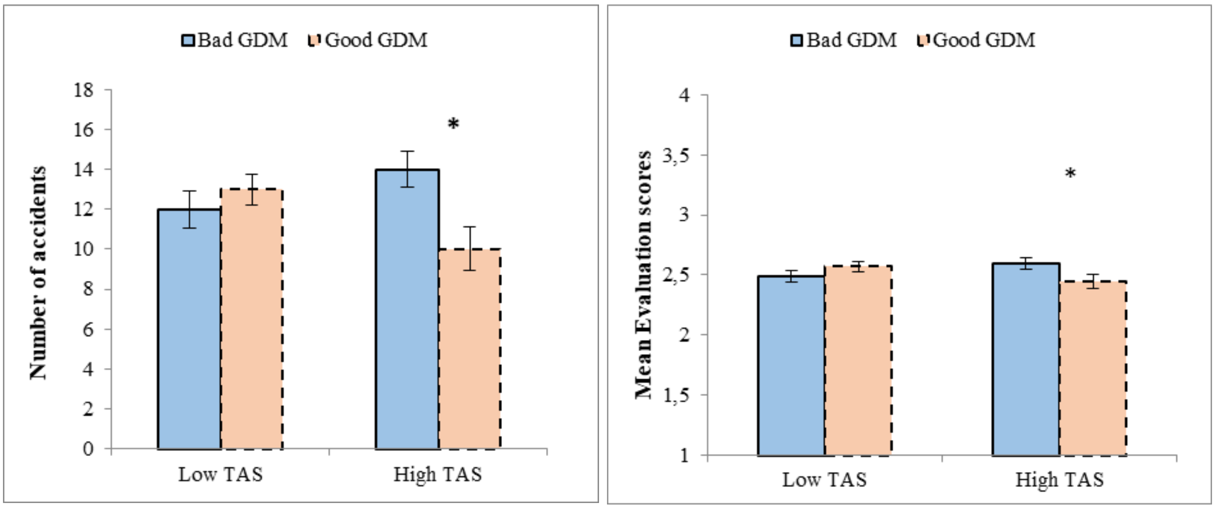

Figure 13: Participants' number of accidents (left panel) and Evaluation scores (right panel) as a function of TAS level and GDM. The asterisks indicate significant post hoc comparisons ($p < 0.05$) with Bonferroni correction. The Evaluation scores ranged from 1 (safe performance) to 4 (accidents), so as that higher scores correspond to worse performance. Vertical bars represent standard errors. Adapted from Gianfranchi *et al.* (2017b)

4.2.3 Discussion

The present study confirmed the possibility of employing the HRT simulator as a tool to assess driving abilities among young participants with different degrees of road exposure. A cluster analysis was run on the indices collected by the HRT, identifying three clusters corresponding to different driving profiles, namely Prudent, Imprudent, and Insecure. Furthermore, in line with our first prediction, differences emerged among the clusters in terms of SS and gender. Specifically, Insecure and Prudent groups show modulation by gender on TAS and DIS dimensions of SS.

Concerning our second prediction, links between SS and decision-making ability were also confirmed, as attested by the significant correlation between TAS and GDM scores. Finally, the results confirmed the third prediction too, in that the effect of SS on driving performance resulted indirect and modulated by decision-making ability.

Gender differences in terms of SS were already reported by Zuckerman (1994), with males usually scoring higher on the TAS and DIS subscales. Our data only partially confirm previous results, showing higher DIS among males, but only in the Prudent group. Conversely, in the Insecure group, females are more thrill and adventure seekers than males. Interestingly, the fact that Prudent females are less disinhibited than Insecure females may indicate that DIS has a detrimental effect on riding performance in females, whereas the fact that Prudent males scored higher at TAS than Insecure males might suggest that thrill seeking does not necessarily represent a disadvantageous trait, at least among males.

These data indicate that SS *per se* does not necessarily leads to risky driving behaviors. Indeed, the results show that bad decision makers with higher TAS scores are more prone to accidents and have generally less safe driving performance on the HRT. In other words, high TAS participants are also risky road users only if they are simultaneously bad decision makers in the IGT. As already noted by Jonah (1997), high sensation seekers may perceive less risk in a variety of on-road situations, or they may have a correct risk perception but accept the risk to experience the thrill and to maintain their level of arousal (Zuckerman, Bone, Neary, Mangelsdorff, & Brustman, 1972). However, the present result is in favor of a further alternative explanation: High sensation seekers can be split into two sub-categories depending on their decision-making ability. Those with good decision-making skills do not behave dangerously on-road, whereas those with poor performance in a decision-making task do. In this framework, general decision-making skills and SS are independent factors that contribute in affecting driving performance.

The clusters here identified are comparable to those reported by Lucidi *et al*. (2010) and Marengo *et al*. (2012). Indeed, both studies reported the presence of three groups of young drivers in their samples, clustered on the basis of many psychological variables. For instance, Marengo *et al.* (2012) named the clusters "Risky," "Worried," and "Safe" drivers, respectively. The three groups showed different behavioral patterns in the HRT in terms of accidents and performance safety. The present study, partially replicating these findings, followed a different and deeper approach, in that we clustered the participants based on their driving performance, focusing on a wide number of driving variables.

Some limitations and future perspectives need to be considered. First, although the comparability among the three driving profiles in terms of annual mileage suggests that driving style is a relatively independent behavior modality, on the other hand, this may seem counterintuitive, as practice has been proved to improve hazard perception and driving performance. However, practice develops over the years, whereas our sample included young-aged participants. Moreover, 20 participants did not provide information regarding their on-road exposure. Considering the importance of this aspect, future research should investigate whether and how exposure, age and training may interact in determining driving behaviors, also involving samples with different age ranges. Moreover, one might wonder whether the relation of SS and decision-making ability with driving behaviors assessed through a simulator really transfers to the real on-road context. Several studies found a link between these dimensions and real on-road behaviors (e.g., Jonah, 1997; Sagberg, Selpi, Bianchi Piccinini, & Engström, 2015). Thus, it is possible to predict that the implications of both SS and decision-making ability in driving behaviors as revealed in the present work might be replicable also when considering real road. However, more research is needed to both replicate this methodology on different aged samples and to assess the generalizability of our findings to real on-road context. This is precisely the aim of the following studies.

Chapter 5

The development of a tool to assess and train specific driving profiles among inexperienced adolescents

As seen in previous chapters, it is possible to employ the HRT simulator with novice drivers so as to train their hazard perception (Tagliabue & Sarlo, 2015; Tagliabue *et al*., 2017; Tagliabue *et al*., 2019) and to identify different driving profiles, related to specific patterns of self-reported aberrant on-road behaviors, SS, and decision-making (Gianfranchi *et al*., 2017a; 2017b). However, a further step of the research is needed to assess the generalizability of previous results to both real on-road contexts and to participants with different age and on-road exposure. Moreover, attempting to control on-road exposure by selecting a totally inexperienced sample would be useful because it would allow to disentangle the effects of on-road exposure from those of personality traits in influencing driving profiles.

In the present chapter, the data of two studies are reported. The first (section 5.1; Gianfranchi *et al*., 2018) aimed at identifying "embryonic" driving profiles in a sample of totally inexperienced adolescents that were trained and tested along two sessions on the HRT. Moreover, different patterns of personality traits were found to predict the inclusion in the driving profiles. The second study (section 5.2) presents the preliminary and unpublished results of a 12-months follow up on the same sample of adolescents regarding bike crashes rate and near misses in the real world, also comparing these data with those of a control group that did not participate in the training on the HRT and that was matched to the trained group for gender, age, and on-road exposure.

5.1 The influence of personality traits on "embryonic" driving profiles

Starting from previous results (Gianfranchi *et al*., 2017a; 2017b), we thought that applying the same methodology of these studies to a sample composed by inexperienced adolescents would lead to the chance of identifying "embryonic" driving profiles, that is, driving profiles not yet influenced by driving experience and that can be the future target of *ad hoc* trainings, before starting to drive in the real world. Moreover, the identification of specific personality patterns that predict the inclusion among one or the other driving profile would confirm the role of personality traits in influencing driving behaviors, also disentangling the effects of on-road exposure and personality on the driving behavior. Finally, we decided to assess the adolescents' beliefs about their peers' on-road risky behaviors too, since previous studies (e.g., Lucidi *et al*., 2010; Marengo *et al*., 2012) neglected the crucial influence of these aspects over the adolescents' driving behavior. Thus, the aims of the here reported study[5] (Gianfranchi *et al*., 2018) were (1) the identification of different profiles of simulated moped-driving in adolescents with no on-road experience and (2) the assessment of the relations between the driving profiles and personality traits and beliefs about their peers' on-road risky behaviors.

5.1.1 Material and methods

Participants

One hundred and fourteen adolescents (55 F; mean age: 14.85 years; range: 13–19 years) enrolled in High Schools of Padua, took part in the study. All of them had no on-road driving or riding experience, but they all used bicycles, with most of them (60%) declaring they ride a bike several times a week or each day. All of the participants had corrected or correct-to-normal vision. They were not paid for their participation. Written informed consent was obtained by all the participants and, for the participants under the age of 18, also by their parents. The project was approved by the Ethical Committee for the Psychological Research of the University of Padova.

[5] An extended version of section 5.1 can be found in Gianfranchi, E., Tagliabue, M., & Vidotto, G. (2018). Personality traits and beliefs about peers' on-road behaviors as predictors of adolescents' moped-riding profiles. *Frontiers in Psychology*, *9*, 2483.

Tools

Driving behavior was assessed by means of the HRT simulator.

All of the participants filled in a battery of questionnaires aimed at assessing their personality traits and their beliefs about peers' on-road behaviors.

Specifically, SS was assessed through the Sensation Seeking Facets measure from the International Personality Item Pool (Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002). It includes 30 items divided into three subscales (10 items each; score range 10-50 for each subscale and 30-150 for the total) aimed at measuring different aspects of SS, from here onwards called thrill-seeking (TS) so as to be consistent with the labels and the contents of the subscales. The subscales are Dangerous TS (*i.e.*, the seeking of dangerous activities; "I might enjoy a free fall from an airplane"), Impulsive TS (*i.e.*, the tendency to be impulsive and unpredictable; "I am unpredictable, people never know what I am going to say"), and Calculated TS (*i.e.*, the willingness to take calculated risks or to face the most common fears; "I would love to explore strange places"). The items are scored on a 5-point Likert scale ranging from 1 (strongly agree) to 5 (strongly disagree).

Participants' LC was assessed through two self-report measures. The first is the driving locus of control scale of the Italian Cognitive Behavioral Assessment (CBA BG; Vidotto, Sica, & Baldo, 1995), which includes 27 items (*e.g.*, "Prudence does not matter to avoiding traffic accidents") on a 5-point Likert scale (from strongly agree to strongly disagree) and a total score range of 27- 135. The second tool is the multidimensional Traffic Locus of Control scale (T-LOC; Özkan & Lajunen, 2005), that aims at discriminating between different dimensions of specific on-road LC. It is composed of four subscales (Others, Self, Vehicles and Environment, and Fate) in which participants must rate whether a crash can result from different types of circumstances (*e.g.*, "Other drivers' risk-taking," "Bad weather or lighting conditions", "My own risk-taking"). The items are on a 5-point Likert scale (from not at all possible to highly possible).

The New–Buss questionnaire (N–B; Gidron, Davidson, & Ilia, 2001), an 8-item self-report tool, was used to assess participants' aggressiveness. The questionnaire is the brief version of the Buss–Perry Aggression Questionnaire (Buss & Perry, 1992), and each of the four subscales that compose the tool includes two items of the original scale. The subscales are Verbal Aggression ("I can't help getting into arguments when people disagree with me"), Anger ("Sometimes, I fly off the handle for no good reason"), Physical Aggression ("Given enough provocation, I may hit another person"), and Hostility ("I sometimes feel that people are laughing at me behind my back"). All the items are on a 5-point Likert scale (from extremely uncharacteristic of me to extremely characteristic of me) with a score range of 2-10 for each scale.

Finally, participants' beliefs about peers' on-road behaviors were assessed through the three subscales (Risky driving, RD; Aggressive Driving, AD; Negative emotions while driving NE) composing the 3DI (Dula & Ballard, 2003). Note that, as seen in the previous chapter, the original 3DI questionnaire does not assess the behavior of the peers. However, given that participants in our sample could not answer to its items on the basis of their own driving experience, they were asked to rate the frequency of the behaviors described in the items observed among their peers. Indeed, although developed to assess experienced drivers' dangerous actions, the 3DI items refer to behaviors that most people can judge as dangerous or inappropriate (*e.g.*, "I will weave in and out of slower traffic" or "I verbally insult drivers who annoy me"), even without proper driving or driving experience.

Procedure

The procedure included two experimental sessions that were scheduled a few days apart from each other. At the beginning of the first session, all the participants filled in the questionnaires. Then, they were invited to sit on the HRT simulator, where an experimenter illustrated the driving controls and gave all the necessary information regarding the task. Participants were told to ride along the virtual paths as safely as they could, trying to avoid accidents. The HRT was set with moped controls, daylight conditions and automatic transmission so as to prevent any bias derived from driving inexperience. During the first session, participants faced five courses on secondary roads (39 hazardous scenes), preceded by a practice course of 3 min. Six courses on main roads were faced during the second session (48 hazardous scenes): These six courses were employed to test the presence of differences in terms of driving profiles among participants, after that all of them have had the same amount of experience on the simulator (first session). In both sessions the courses were ordered following their degree of difficulty. Before starting the practice, all of the participants were asked about their knowledge on the main road rules and signals (*e.g.*, traffic lights and stop signs), and all of them proved to be aware enough of the main rules and signals.

Data coding and analyses

For the questionnaires, the original scoring instructions were followed.

Concerning driving performance, as in previous works (Gianfranchi *et al.*, 2017a; 2017b) we extracted 18 indices from participants' performance on the HRT, focusing on the second session. The indices were computed only on the courses of the second session. Indeed, we thought that, because our participants were all inexperienced, a proper driving profile could emerge only after a minimum experience with the virtual environment and the driving task.

The statistical analyses were divided into two main steps. After the inspection of the self-report measures (descriptive statistics, Cronbach's alpha and correlations), the first main step was aimed at identifying the driving profiles among the participants in the second session through a cluster analysis. Then, we assessed differences between clusters in terms of risky behaviors through a MANOVA on the percentages of A, B, C, and D scores obtained during the second session, with *Group* (*i.e.*, the identified driving profiles) as the between-subjects factor. Moreover, in order to rule out that the effects observed are due to differences in learning or driving skills already present before the test procedure in the second session, an identical MANOVA was carried out on the A, B, C, D scores of the first session. Post hoc analyses using Bonferroni's correction were conducted, with α set at 0.05.

The second main step aimed at identifying the psychological predictors of the inclusion in the driving profiles. Thus, we ran a multinomial logistic regression with the cluster solution as the dependent variable and the scores from the questionnaires as the predictors. All the analyses were performed with the IBM SPSS 23 statistical software package.

5.1.2 Results

As a preliminary step, following the procedure adopted by Marengo *et al*. (2012) descriptive and reliability statistics (Cronbach's alpha) were calculated for the employed scales, along with Pearson's correlations (Table 3). The correlation coefficients show the presence of significant links among personality traits and between personality traits and beliefs about peers' on-road behavior.

Cronbach's alpha levels ranged from moderate (> 0.50, for some of the scales with a low number of items) to high (> 0.70) except for the subscales Verbal Aggression and Hostility of the N–B questionnaire. However, this last result is not surprising because the N–B scales include only two items each. Thus, following Briggs and Cheek's (1986) suggestion to focus on the inter-item correlation instead of on the Cronbach's alpha in the case of scales with a low number of items, we calculated the inter-item correlations for each N–B scale. The coefficients are 0.26 for Verbal Aggression, 0.58 for Physical Aggression, 0.25 for Hostility, and 0.53 for Anger. Inter-item correlation coefficients higher than 0.20 are considered optimal (Briggs and Cheek, 1986).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	α	Mean (S.D.)	Range
1	Dangerous TS	-															.81	26.1 (7.3)	11-50
2	Impulsive TS	**.73****	-														.89	27.7 (8)	11-48
3	Calculated TS	**.57****	**.52****	-													.69	36.7 (5.8)	22-50
4	CBA BG	.39**	**.57****	.24*	-												.72	69.2 (9.6)	51-91
5	T – LOC Others	-.18	-.16	.02	-.22*	-											.53	20.2 (2.1)	14-25
6	T – LOC Self	-.11	-.02	.11	-.17	**.59****	-										.63	19.7 (2.6)	10-25
7	T – LOC VE	.15	.27**	.21*	.13	.25**	.13	-									.55	15.5 (2)	10-20
8	T – LOC Fate	.32**	.45**	.15	.44**	-.03	-.07	.30**	-								.82	6.7 (2.6)	3-15
9	N-B Verbal Aggr.	.16	.31**	.14	.33**	-.09	-.13	.11	.20*	-							.41	5.4 (1.9)	2-10
10	N–B Anger	.19*	.40**	.07	.32**	-.07	-.01	.10	.16	**.59****	-						.70	5.2 (2)	2-10
11	N-B Physical Aggr.	.26**	.37**	.25**	.29**	-.08	.02	.15	.17	.44**	**.53****	-					.73	5.6 (2.3)	2-10
12	N-B Hostility	.10	.23*	.01	.16	.04	-.07	.25**	.07	.41**	.42**	.30**	-				.40	5.6 (1.8)	2-10
13	3DI AD	.21*	.22*	.09	.31**	-.19*	-.16	.07	.12	.31**	.31**	.29**	.14	-			.74	12.9 (4.1)	7-32
14	3DI NE	.09	.27**	.09	.43**	.01	.08	.11	.10	.43**	.38**	.31**	.20*	**.64****	-		.73	22.7 (5.1)	12-38
15	3DI RD	.37**	.39**	.38**	.48**	-.22*	-.14	.05	.18	.27**	.19*	.26**	.07	**.62****	**.52****	-	.81	21.2 (6.6)	12-47

Table 3: Pearson's correlations, Cronbach's alpha and descriptive statistics for scales of each questionnaire. Double and single asterisks indicate $p < 0.01$ and $p < 0.05$, respectively. In bold the correlations with $r > 0.5$. Adapted from Gianfranchi *et al*. (2018).

The driving profiles

The next step was the identification of the "embryonic" driving profiles through a cluster analysis with the 18 HRT indices of the second session used as grouping variables, applying a hierarchical clustering procedure (Ward's method clustering with the squared Euclidean distance on Z-scored variables). Three clusters emerged, with different driving patterns.

As depicted in Figure 14, the profiles reported specific trends on the HRT indices. The first profile, labeled "Imprudent" (21 participants; 6 F; mean age: 14.90 years), shows an unsafe behavioral pattern, with the highest values in almost all the driving indices (*e.g.*, speed, throttle opening, and Evaluation score). The second profile shows an opposite trend as compared to the Imprudent profile, with low values in all the driving indices and high rates of prevented accidents. Thus, we labeled this profile "Prudent" (47 participants; 30 F; mean age: 14.89 years). Finally, the third cluster, which in Gianfranchi *et al*. (2017b) was labeled "Insecure", shows a mixed pattern, with an overall safe performance, but with elements that can be potentially dangerous, such as the proneness to hard brake. This last cluster includes 46 participants (19 F) with a mean age of 14.78 years. Although the profiles are homogenous in terms of age, a chi-squared test showed significant differences in terms of gender ($\chi^2(2) = 8.71$, $p < 0.05$): Females are predominant in the Prudent cluster (30 F *vs*. 17 M), whereas males are predominant in the Imprudent (6 F *vs*. 15 M).

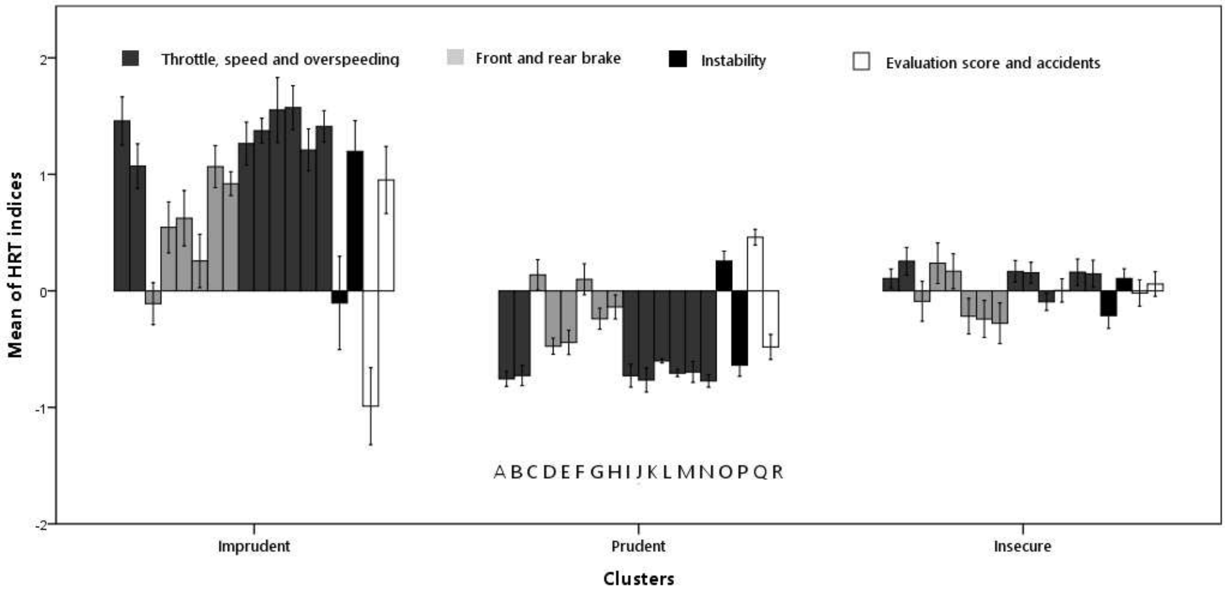

Figure 14: Mean Z-scores of the 18 HRT indices in the three clusters. The indices are listed in the order displayed by the letters on the bottom of the panel for each cluster, as follows: the mean of the throttle opening (A) and its SD (B); number of times using the front brake (C); mean (D) and SD (E) of front brake pressure; number of times using the rear brake (F); mean (G) and SD (H) of rear brake pressure; mean (I) and SD of speed (J); time spent over the speed limit (K); number (L), mean (M), and the highest value (N) of speeding; mean (O) and standard deviation (P) of on-road instability; number of prevented accidents (Q); and mean Evaluation score (R). Vertical bars represent SE. Adapted from Gianfranchi *et al*. (2018).

In order to better understand the differences among the identified driving profiles in terms of risky behaviors, a MANOVA was run on the percentages of A, B, C, and D scores of the second session (calculated over the total of the scenes) with the driving profiles as the between independent variable (3 levels). At the multivariate level, the results show that the three profiles are significantly different (Wilks' λ= 0.69, $F(6, 218) = 7.43$, $p < 0.001$, $\eta_p^2 = 0.17$). Univariate results indicate that significant differences are present in the percentages of each score ($F(2,111) = 16.64$, $p < 0.001$, and $\eta_p^2= 0.23$ for A score; $F(2, 111) = 9.99$, $p < 0.001$, and $\eta_p^2= 0.15$ for B score; $F(2, 111) = 6.58$, $p < 0.01$, and $\eta_p^2= 0.11$ for C score; and $F(2, 111) = 17.74$, $p < 0.001$, and $\eta_p^2= 0.24$ for D score). As depicted in Figure 15, Imprudent drivers showed a less safe performance, with the lowest percentages of A scores (M = 54.3%) than both Prudent (M = 74.6%, $p < 0.001$ at the post hoc) and Insecure drivers (M = 66%, $p < 0.01$). Imprudent drivers showed also the highest percentages of C and D scores: C scores were 9.5% in Imprudent participants *vs*. 4.9% in Prudent ($p < 0.001$) and 6.2% in Insecure ones ($p < 0.05$); D scores were 5.6% in Imprudent drivers *vs*. 0.8% in Prudent ($p < 0.001$) and 2.4% in Insecure drivers ($p < 0.001$). On the other hand, Insecure drivers obtained lower percentages in A scores than Prudent drivers (66% *vs*. 74.6%, $p<0.01$) but higher than Imprudent drivers (66 *vs*. 54.3%, $p < 0.01$), and higher D percentages than Prudent drivers (2.4 *vs*. 0.8%, $p < 0.05$) but lower than Imprudent participants (2.4% *vs*. 5.6%, $p < 0.001$). Finally, they did not differ from Prudent drivers in terms of C scores and from Imprudent drivers in terms of B scores.

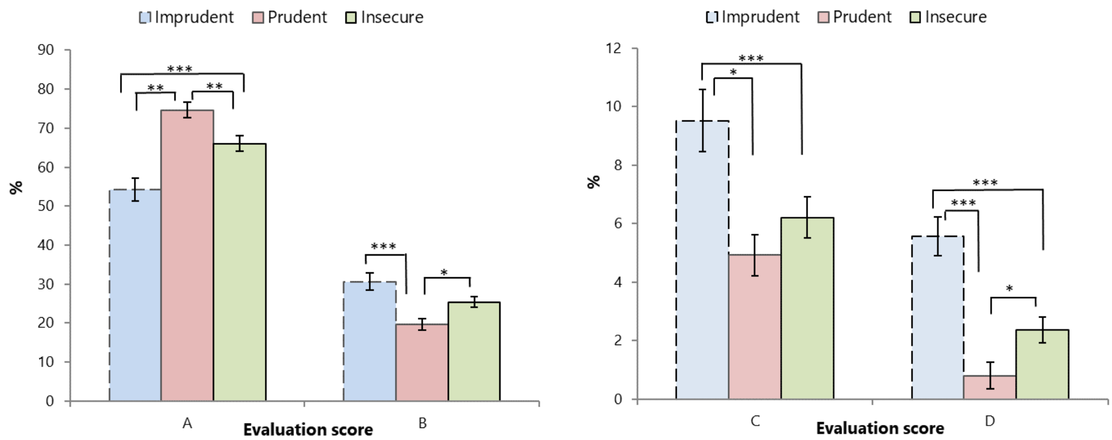

Figure 15: Differences in evaluation scores (percentage) among the clusters. Vertical bars represent standard errors. Asterisks indicate significant differences in the post hoc comparisons with Bonferroni correction: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. Adapted from Gianfranchi *et al*. (2018).

Overall, we can conclude that Imprudent participants showed a less safe driving performance, with high percentages of scenes with crashes (D), near misses (C), and almost safe behaviors (B), reporting at the same time the lowest frequency of totally safe scenes (A). Prudent drivers showed the opposite pattern, but they did not differ from Insecure drivers in terms of near misses (C). Finally, participants in the Insecure cluster reported similar B percentages to those of the Imprudent cluster, testifying that Insecure drivers' performances, although overall better than those of Imprudent drivers, included a significant amount of not totally safe scenes (*e.g.*, hard braking or disrespecting safe distance). An identical MANOVA on the A, B, C, D scores obtained during the first session was carried out. Here, the factor *Group* did not result significant at the multivariate level ($p = 0.111$, $\eta_p^2 = 0.06$), thus allowing us to rule out that the effects just described are due to differences in learning or driving skills already present in the first session.

Driving profiles, personality traits and beliefs

A multinomial logistic regression (stepwise backward method) on the cluster solution as the dependent variable and the scores on all the questionnaires' scales as predictors was run to identify patterns of personality traits and beliefs that can predict inclusion in the driving profiles.

The final model ($\chi^2(18) = 44.99$, $p<0.001$), explained 33% of the variance (Cox and Snell's Pseudo $R^2 = 0.33$) with a classification accuracy of 60.5%. Seven predictors were significant in the final model (Table 4); that is, two dimensions of TS (Dangerous TS and Impulsive TS), two measures of LC (CBA BG and T-LOC Fate subscale), two dimensions of aggressiveness (N–B Anger and Verbal Aggression), and beliefs about peers' risky driving behaviors (3DI RD).

Likelihood Ratio Test

	χ^2	Df	p values
Dangerous TS	10.69	2	.005
Impulsive TS	7.67	2	.022
Calculated TS	5.35	2	.069
CBA BG	8.58	2	.014
T-LOC Fate	8.91	2	.012
N-B Anger	7.75	2	.021
N-B Verbal aggression	7.29	2	.026
3DI RD	6.40	2	.041
3DI AD	5.00	2	.082
Intercept	16.73	2	.000

Table 4: Likelihood ratio test of the final regression model. Adapted from Gianfranchi *et al*. (2018).

The regression coefficients reported at the top of Table 5 show that the likelihood of being included among Imprudent drivers with respect to Prudent and Insecure profiles was increased by lower scores on the 3DI Risky Driving scale ($p < 0.05$) and on the T-LOC Fate scale ($p < 0.05$ compared with Prudent drivers and $p < 0.01$ compared with Insecure drivers) but by higher scores at the CBA BG ($p < 0.01$ with respect to Prudent participants and $p < 0.05$ with respect to Insecure participants). Moreover, higher scores on the Dangerous TS and N–B Verbal Aggression scales play a significant role ($p < 0.05$) in discriminating between Imprudent and Prudent profiles.

When the Insecure profile is used as the reference category (bottom of Table 5), the coefficients show that higher scores on the N–B Anger scale predict inclusion in the Insecure profile, compared to the other two ($p < 0.05$). Moreover, reporting high scores on the Dangerous TS scale ($p < 0.05$) but, at the same time, low scores on the Impulsive TS scale ($p < 0.05$) increased the risk of being included in the Insecure profile, compared to the Prudent one.

Summarizing (Figure 16), factors such as an external locus of control, the underestimation of fate among the causes of crashes and of the frequency of peers' on-road risky behaviors played a critical role in discriminating Imprudent drivers from the other profiles. Moreover, participants with high levels of verbal aggression had a higher likelihood of being included in the Imprudent profile than in the Prudent profile. A low tendency to seek dangerous situations raised the probability of being included in the Prudent profile. Finally, inclusion in the Insecure profile was predicted by high levels of anger, whereas low levels of impulsivity played a role in discriminating between Insecure and Prudent drivers.

Reference category: Imprudent									
Prudent	**Beta**	**χ^2**	**Df**	**p value**	**Insecure riders**	**Beta**	**χ^2**	**Df**	**p value**
Dangerous TS	-0.18	6.48	1	**.011**	Dangerous TS	-0.05	0.43	1	.512
Impulsive TS	0.09	1.82	1	.177	Impulsive TS	-0.06	0.69	1	.406
Calculated TS	-0.11	2.01	1	.157	Calculated TS	-0.17	4.81	1	**.028**
CBA BG	-0.12	7.38	1	**.007**	CBA BG	-0.10	4.83	1	**.028**
T-LOC Fate	0.30	3.87	1	**.049**	T-LOC Fate	0.42	7.58	1	**.006**
N-B Anger	0.14	0.51	1	.475	N-B Anger	0.50	5.71	1	**.017**
N-B Verbal Aggr.	-0.53	6.41	1	**.011**	N-B Verbal Aggr.	-0.37	3.28	1	.070
3DI RD	0.17	4.61	1	**.032**	3DI RD	0.18	4.79	1	**.029**
3DI AD	-0.06	0.42	1	.518	3DI AD	-0.20	3.79	1	.052
Intercept	*13.51*	*11.19*	*1*	*.001*	*Intercept*	*12.94*	*10.55*	*1*	*.001*
Reference category: Insecure									
Imprudent	**Beta**	**χ^2**	**Df**	**p value**	**Prudent riders**	**Beta**	**χ^2**	**Df**	**p value**
Dangerous TS	0.05	0.43	1	.512	Dangerous TS	-0.13	6.35	1	**.012**
Impulsive TS	0.06	0.69	1	.406	Impulsive TS	0.15	6.67	1	**.010**
Calculated TS	0.17	4.81	1	**.028**	Calculated TS	0.62	1.45	1	.229
CBA BG	0.10	4.83	1	**.028**	CBA BG	-0.02	0.46	1	.499
T-LOC Fate	-0.42	7.58	1	**.006**	T-LOC Fate	-0.12	1.29	1	.257
N-B Anger	-0.50	5.71	1	**.017**	N-B Anger	-0.36	4.33	1	**.038**
N-B Verbal Aggr.	0.37	3.28	1	.070	N-B Verbal Aggr.	-0.15	0.84	1	.359
3DI RD	-0.18	4.79	1	**.029**	3DI RD	-0.01	0.02	1	.881
3DI AD	0.20	3.79	1	.052	3DI AD	0.14	2.96	1	.086
Intercept	*-12.94*	*10.55*	*1*	*.001*	*Intercept*	*0.56*	*0.06*	*1*	*.803*

Table 5: Parameter estimates of the multinomial regression with Imprudent (top of the table) and Insecure (bottom of the table) profiles as reference categories. In bold the significant coefficients. Adapted from Gianfranchi *et al*. (2018).

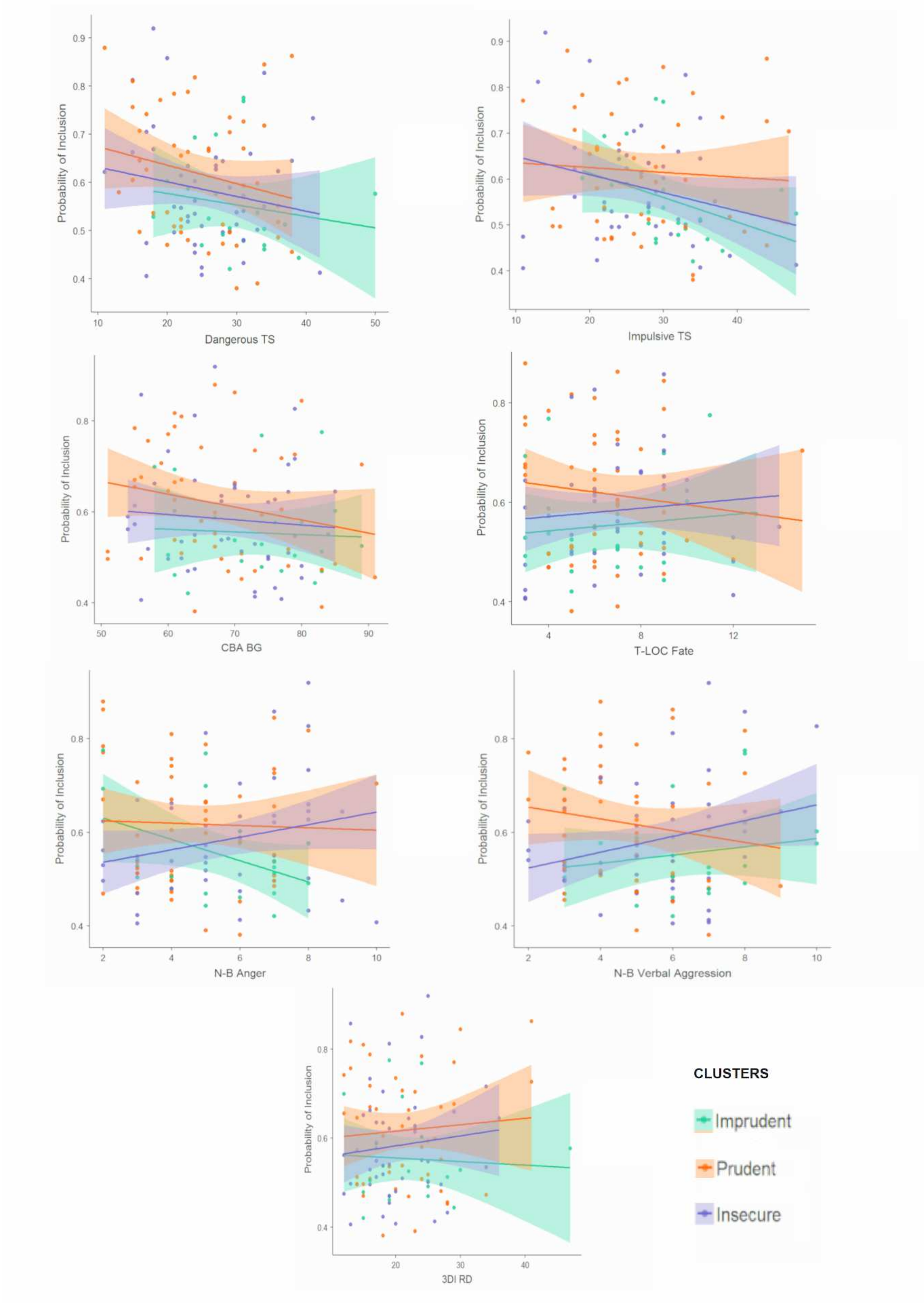

Figure 16: Probability of inclusion in the three profiles for the scores of each significant predictor. Shaded areas represent standard errors. Adapted from Gianfranchi *et al*. (2018).

5.1.3 Discussion

The present study aimed at applying the procedure of identification of driving profiles to a sample of inexperienced adolescents. The idea was that "embryonic" driving profiles should emerge after one session of driving on the HRT simulator and that these profiles should resemble those identified among novice drivers in the previous studies. Thus, a cluster analysis performed on 18 indices of the second experimental session on the HRT simulator allowed the identification of three moped-driving profiles in the present sample: Imprudent, Prudent and Insecure drivers.

The profiles showed different and specific driving patterns. The Imprudent drivers exhibited the most unsafe pattern, with high speed and acceleration levels, high frequency of speeding behavior, and high rates of accidents and instability. The Prudent profile showed the opposite trend, whereas the Insecure drivers had intermediate characteristics.

Moreover, a significant difference emerged among the profiles in terms of gender. The Prudent profile is mostly composed of females, whereas the Imprudent is mostly composed of males. Many studies (for a brief review see Oltedal & Rundmo, 2006) have proved that males are more prone to the effects of SS and to showing risky driving behaviors, also among adolescents (Oltedal & Rundmo, 2006; Marengo *et al*., 2012). Further analyses of the present data confirmed significant differences among the profiles in terms of risky behaviors. Indeed, Imprudent participants reported the lowest percentage of safe scenes and the highest of near misses and crashes, conversely to the Prudent drivers. Insecure drivers had overall a mid-range performance, with a near misses rate comparable to that of the Prudent participants but, at the same time, lower rates of safe scenes and higher percentages of almost safe scenes; these last were comparable to those of the Imprudent profile.

Importantly, the three clusters here identified resemble those already reported in Gianfranchi *et al.* (2017b) and those found by Lucidi *et al*. (2010) and Marengo *et al.* (2012). The similarity between the present driving profiles and those already identified with the same methodology in a group of novice drivers (Gianfranchi *et al*., 2017b) confirms the replicability and generalizability of this assessment method to different types of populations. Moreover, the present study highlights the presence of a certain degree of consistency in the characteristics of driving profiles among samples with different age and on-road exposure, pointing out that it is possible to find inter-individual differences in driving behaviors also among adolescents with no on-road experience, thus stressing the role of personality traits and beliefs.

On the other hand, as regards the comparison with the results by Lucidi *et al*. (2010) and Marengo *et al*. (2012), the present study confirms and even extends their main findings by three more aspects. First, it aims to categorize different profiles based on a quantitative evaluation of their performance on the simulator. Second, it considers personality traits and beliefs as predictors of the profiles so as to find a direct relation between them. Finally, the use of the questionnaire subscales allows us to deeply assess, when present, the relation between personality traits, beliefs and driving performance.

As for the role of the predictors, our results are in line with the previous literature. SS (especially dangerous TS) and LC seem to play key roles in the predictions of participants' driving profiles, with high levels of SS and an external LC being associated with an increase in the risk of an imprudent behavior (Lucidi *et al*., 2010; Marengo *et al*.,2012). It is worth noting that lower scores on the Fate scale of the T-LOC predicted inclusion among Imprudent drivers in our sample. Although attributing the causes of crashes to coincidence or fate may be interpreted as an index of external LC, considering the importance of uncontrollable factors for the crash risk may have a role in developing defensive driving strategies, which in turn may lead to more cautious behavior. Low levels of impulsivity and high levels of anger increased the risk of showing an insecure driving style among adolescents in our sample. Being less impulsive, although frequently associated with cautious behavior (Marengo *et al*., 2012), might also lead to difficulties in self-regulation of driving behaviors when a quick decision is required to face impending hazardous scenarios. This might explain how, in the present research, low levels of impulsivity are associated with insecure but not imprudent behaviors. However, further research is needed to support this conclusion.

Concerning the role of anger, Dahlen, Edwards, Tubré, Zyphur, and Warren (2012) tested a theoretical model of associations between different personality traits, aggressive driving and driving outcomes in a sample including only adult drivers. Their results showed the existence of a positive relationship between low emotional stability (including anger, depression, and anxiety) and aggressive driving, which, as a result, led to more on-road violations, near misses and crashes. In our sample, anger has proved to be predictive of the inclusion in the Insecure profile. At the same time, Insecure drivers showed more reckless behaviors than Prudent drivers, as attested by the lower frequency of safe scenes (A scores) and the higher frequency of both almost safe scenes (B scores) and crashes (D scores). These results are in line with the conclusions by Dahlen *et al.* (2012) as to road violations and crashes, indicating that higher levels of anger may represent a risk factor for less cautious driving behaviors. However, the result related to near misses has not been replicated. This discrepancy may be due to differences in age and experience of the involved samples or in the adopted questionnaires. Nevertheless, our study confirmed the key role of anger in predicting driving behaviors among adolescents.

Finally, underestimating peers' on-road risky behaviors increased the risk of showing imprudent behavior on the HRT, with significantly higher percentages of crashes and near misses. Indeed, a correct estimation of others' potentially hazardous behavior is crucial to preventing crashes and it represents the basis of the development of hazard perception and defensive driving strategies.

A first limitation of the present study is represented by the limited number of questionnaires used to assess personality variables. Indeed, risky driving is influenced by many variables, among which even impulsivity or risk proneness play a prominent role (Megías, Cándido, Maldonado, & Catena, 2018). Moreover, an important limitation is related to the generalizability of the results to real on-road behaviors. Indeed, although the identification of profiles based on participants' performances in a simulated environment rather than on self-report measures represents an advancement in the assessment methods of driving behaviors, there is still controversial evidence on the ecological validity of the simulators. Thus, a further and necessary step will be a follow up on self-reported data and real on-road performance, so as to have the chance to develop effective *ad hoc* training protocols that may provide a crucial contribution to preventing road crashes.

5.2 A first evaluation of the real-world effects of a driving training in a simulated environment

In the present section, the unpublished results of a 12-months follow up study (about real world on-road behaviors) on the sample of inexperienced adolescents involved in the previous study (Gianfranchi *et al*., 2018) are presented. Moreover, the preliminary results of the comparison between these data and the same collected on a control group are reported. The aim of the present study is twofold. First, the follow up data are necessary to verify the generalizability of the results of the two-session HRT training to the real world. Specifically, we expected to find an improvement (in terms of safety) in the driving performance during the training on the HRT and a generalization of this effect to real on-road behaviors, with a reduction of the bike crashes in the year following the training. Second, comparing the crash rate and the frequency of on-road near misses (*i.e.*, the occurrence of a situation in which the participant almost crashed with another vehicle or road user; Sanders, 2015) of the trained group with those of a control group (matched for age, gender, and on-road exposure) that has not been involved yet in the training allows to exclude the role of age and experience in the improvement of on-road performance.

5.2.1 Material and methods

Participants

Of the original sample composed by 114 participants, 101 accepted to take part to the follow up phase of the project (48 F; mean age: 15.94 years; range: 14-20 years).

Of them, 19 participants were classified in the previous study as Imprudent drivers, 41 as Prudent, and 41 as Insecure. Three participants took the car driving license after the training, whereas three took the moped-driving license. All of the participants had correct or correct-to-normal vision.

As regards the control group, to date we have collected the data of 26 participants (19 F; mean age: 17.23 years; age range: 14-19 years) that were matched to 26 participants of the training group on the basis of gender, age and on-road exposure. Specifically, both the training and the control group declared a mean frequency of bike use of one day a week. Moreover, three participants per group hold the car driving license since a minimum of two months, declaring a total mileage of less than 5,000 km. None of the participants was paid for the involvement in the study. All of them were enrolled in the High Schools of Padua. Written informed consent was obtained by all the participants and, for those under the age of 18, also by their parents. The project was approved by the Ethical Committee for the Psychological Research of the University of Padua.

Tools and procedure

As seen in section 5.1.1., the participants in the training group were first involved in a two-session training on the HRT. During the first session they faced five courses on secondary roads for a total of 39 hazardous scenes, whereas during the second session (scheduled some days apart) they drove along six courses on main roads, facing 48 more hazardous scenes. For more details see section 5.1.1.

Concerning the follow up, after the explanations regarding the aims and the procedure of this phase to both parents and adolescents, participants who accepted to participate in the follow up were contacted 12 months after the end of their second session on the HRT. The follow up consisted in a brief structured phone interview (about 10 minutes in duration) in which the experimenter asked to the participants details regarding their licenses, on-road experience, crash rate and frequency of near misses in the last 12 months after the training on the HRT. The questions were developed on the basis of our research need. Specifically, the first questions regarded the holding of any riding or driving license (type, months since the obtainment, km in the last year). Then, the use of bikes was investigated, by asking the frequency of use during the last 12 months. After that, the experimenter asked questions regarding the frequency of crashes, either with car, moped, motorcycle and bike in the last year and in the previous week. All the questions regarding the frequency of any behavior were scored on a 5-points scale from 0 to 4.

The last part of the interview concerned the so-called on-road near misses: The experimenter read a list of 21 situations in which a near miss could have occurred asking to the participant, for each situation, to say whether it has ever happened to him/her and which vehicle (car, moped/motorcycle, bike) he/she was driving. The situations are divided into "Caused by me" (9 items; "I almost had an accident because another road user cut me off") and "Caused by other road users or other situations" (12 items; "I almost had an accident after I did a sudden maneuver").

The last questions regarded a qualitative evaluation of the training experienced on the HRT simulator ("How useful seem to you the training experienced on the simulator?"; from "totally useless" – 0 - to "extremely useful" - 4) and the recall of any situation happened in the real world that could resemble one or more hazardous scenes experienced on the HRT ("Have you had the chance to see on the real road any of the scenes experienced on the simulator?"; "Could you please describe it/them?"). The participants were informed that the experimenter was taking notes of their answers. Moreover, they were instructed to answer sincerely, to take the time they needed to think, and to tell whether any of the questions was inappropriate, having the right not to answer. For the control group, the same interview procedure was followed. However, no questions regarding the training on the HRT were asked. Importantly, due to ethical reasons, each participant in the control group was invited, if wanted, to take part to the training after the completion of the interview.

Data coding and analyses

As regards the driving performance on the HRT, we calculated the percentage of avoided accidents in each session (n of scenes without accidents*100/n of total scenes) and the percentages of scenes in which participants have scored A, B, and C (n of scenes*100/n of total scenes, for each score). Given the driving inexperience of the sample at the time of the training, concerning the follow up evaluation, the only measure that could be compared with the previous data was the frequency of self-reported bike crashes. Finally, to compare the real on-road behavior between the trained and the control group, we extracted the frequency of bike crashes in the two groups both in the last week and in the last year, and, as a final index, the frequency of near misses in the last week. Moreover, we also calculated the proportion of recognized "Caused by me" (n of scenes/9), of "Caused by others/other situations" (n of scenes/12) and the total proportion (n of scenes/21) of near misses.

Thus, the analyses were divided into three steps. A first evaluation of the effectiveness of the training was done, through two separate 3 × 2 repeated measure ANOVAs on the percentage of avoided crashes on the percentages of scenes scored A, B, and C respectively. In both the analyses, *Driving profile* (between-subjects factor; 3 levels: Imprudent, Prudent and Insecure drivers, as labeled in the previous study) and *Session* (within factor; 2 levels: First and second session) were used as the independent variables.

Second, a 3 × 2 repeated measure analysis of covariance (ANCOVA) was run on the mean frequency of bike crashes on the real road with *Driving profile* (between-subjects factor; 3 levels) and *Training* (within-subjects factor; 2 levels: 12 months before training *vs*. 12 months after) controlling for the *Bike use* between the two tested periods expressed as percentage change (follow up frequency - initial frequency/initial frequency*100). Moreover, the evaluation that participants gave to the training was explored through descriptive statistics.

Third, a preliminary analysis was run on 52 participants (26 participants of the control group and their matched 26 trained participants) so as to explore the presence of differences in the crash rate and in the near misses frequency between the trained and the control group. Thus, one univariate ANOVA was run on the frequency of bike crashes in the last year, with *Group* (between-subjects factor; 2 levels: Trained and untrained participants) as predictor. Moreover, a MANOVA with *Group* as predictor was run on the frequency of bike crashes and on the frequency of near misses in the last week. Finally, an identical MANOVA was run also on the "Caused by me", on the "Caused by others/other situations" and on the total proportions of near misses.

All the analyses were performed with the IBM SPSS 25 statistical software package. Post hoc comparisons were run with Bonferroni correction and α level set at 0.05.

5.2.2 Results

In the ANOVA on the percentage of avoided crashes, both *Driving profile* and *Session* reached significance ($F(2,98) = 15.62$, $p < 0.001$, $\eta_p^2 = 0.24$ and $F(1,98) = 664.79$, $p < 0.001$, $\eta_p^2 = 0.87$ respectively). The three profiles were all different in terms of avoided crashes ($p < 0.01$), with Prudent drivers avoiding the crashes in 88% of the hazardous scenes, Imprudent drivers in 80% and Insecure drivers in 85%. However, the safety of the driving performance significantly improved in the second session, independently of the driving profile (M = 73% *vs*. 95% in the first and second session respectively).

As regards the MANOVA on the A, B, C percentages, at the multivariate level both *Driving profile* and *Session* were significant (Wilks' $\lambda = 0.73$, $F(6,192) = 5.52$, $p < 0.001$, $\eta_p^2 = 0.15$ and

Wilks' $\lambda = 0.08$, $F(3, 96) = 351.27$, $p < 0.001$, $\eta_p^2 = 0.92$ respectively). Moreover, the *Driving profile* × *Session* interaction reached significance too (Wilks' $\lambda = 0.72$, $F(6,192) = 5.68$, $p < 0.001$, $\eta_p^2 = 0.15$). At the univariate level, *Driving profile* was significant for A and B scores ($F(2,98) = 9.38$, $p < 0.001$, $\eta_p^2 = 0.16$ and $F(2,98) = 5.60$, $p < 0.01$, $\eta_p^2 = 0.10$ respectively), indicating that both Imprudent (M = 40%, $p < 0.001$) and Insecure drivers (M = 46%, $p < 0.05$) reported less A scores than Prudent drivers (54%). Concerning the B score, Prudent drivers reported B scores in 24% of the hazardous scenes, differing both from Imprudent (M = 30%, $p < 0.05$) and Insecure drivers (M = 29%, $p < 0.05$). The *Session* factor was significant for all the three dependents, ($F(1,98) = 821,17$, $p < 0.001$, $\eta_p^2 = 0.82$ for A score, $F(1,98) = 22.39$, $p < 0.001$, $\eta_p^2 = 0.19$ for B score and $F(1,98) = 169.33$, $p < 0.001$, $\eta_p^2 = 0.63$ for C score). While A scores increased in the second session (M = 28% *vs*. 65%), the other scores decreased (M = 31% *vs*. 25% for B, and M = 19% *vs*. 7% for C score). Finally, *Driving profile* × *Session* interaction (Figure 17) was significant for A and B scores ($F(2,98) = 11.77$, $p < 0.001$, $\eta_p^2 = 0.19$ and $F(2,98) = 10.07$, $p < 0.001$, $\eta_p^2 = 0.17$ respectively).

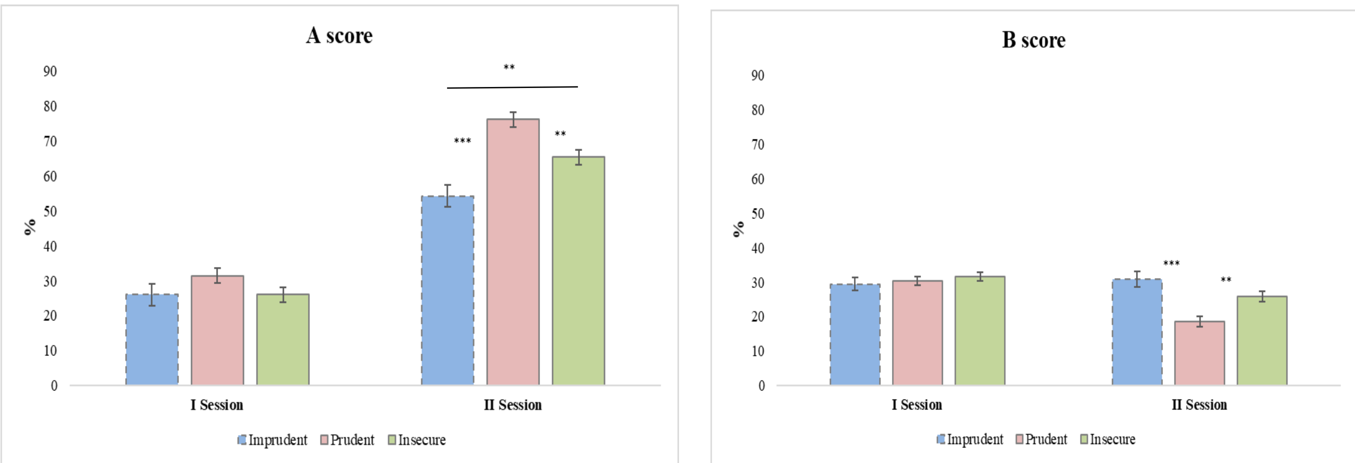

Figure 17: Mean percentages of the A (left panel) and B (right panel) scores in the two sessions, depending on the driving profile. Vertical bars represent standard errors. Asterisks indicate p level, * p < 0.05, ** p < 0.01, *** 0.001.

Although all the driving profiles showed a comparable percentage of A and B scores in the first session and a significant increment in the A scores in the second session ($p < 0.001$ for all the comparisons), Figure 17 shows that in the second session the three profiles significantly differ in terms of percentage of A score, with the Prudent drivers reporting the highest percentage. Moreover, a significant decrement in the second session for the B score is present only for the Prudent ($p < 0.01$) and Insecure drivers ($p < 0.05$). At the same time, as displayed in Figure 17, the Prudent profile significantly differs from the others in terms of percentages of B score.

Moving to the follow up evaluation, in the ANCOVA on the frequency of real road bike crashes only the factor *Training* reached significance ($F(1,97) = 5.75$, $p < 0.05$, $\eta_p^2 = 0.06$). In the 12 months after the completion of the training, controlling for any variation in the frequency of bike use, participants reported a mean frequency of 0.23 bike crashes *vs*. 0.42 of the 12 months preceding the training, independently of the driving profile in which they were included.

Concerning the evaluation of the training made by participants, 46.5% of them considered the training as “quite useful”, and 44.6% as “very useful”. Only one participant declared that the training was totally useless. Moreover, 85.1% of participants declared to have seen on the real road some of the scenes reproduced by the HRT, such as pedestrians crossing the road without paying attention to the traffic or car doors suddenly open.

As regards the comparison between trained and control group, in the univariate ANOVA on the frequency of bike crashes in the last year the *Group* factor was significant ($F(1,50) = 5.07$, $p < 0.05$, $\eta_p^2 = 0.09$). Trained participants reported a lower frequency of bike crashes in the last 12 months compared to the participants in the control group ($M = 0.19$ *vs*. 0.77). In the MANOVA on the frequency of crashes and of near misses in the last week no sources of variance reached significance. Finally, in the last MANOVA on the proportions of near misses, no significant differences were found between groups, although the data seem to suggest a trend of the trained participants, compared to the controls, to report more near misses, both caused by themselves ($M = 0.27$ *vs*. 0.18) and caused by others and other situations ($M = 0.39$ *vs*. 0.30).

5.2.3 Discussion

The data reported in the present section are unpublished and referred to an ongoing study that focuses on two phases necessary for the effectiveness evaluation of a two-session HRT training. The first phase (completed) consists in the follow up of the same sample of trained adolescents involved in the study by Gianfranchi *et al*. (2018), whereas the second step (still ongoing) concerns the comparison of trained participants' follow up data with those of a totally new and untrained control group, matched for age, gender and on-road experience.

Before proceeding with the analysis of the follow up data, an evaluation of the learning effects produced by the HRT training was conducted. Specifically, we assessed the changes in the participants driving behavior between sessions in terms of percentage of avoided accidents and of percentages of A, B, and C scores. The results showed that, although a general improvement in the driving safety is present in the second session, both in terms of avoided crashes and of the frequency of safe (*i.e*., A) and almost safe (*i.e*., B) scores, differences in learning modulated by the driving profile are present. As seen, Prudent drivers generally prevent a higher number of crashes and reported higher percentages of safe scores as compared to both the other two profiles. At the same time, they also reported the lowest percentage of almost safe scores, thus testifying that driving profiles are different in terms of "quality" of the learning.

Indeed, Prudent drivers were the only that not only decreased they crash rate in the second session, but they also increased the number of scenes with safe scores, while decreasing the almost safe one, resulting in a perfectly safe driving performance. On the other hand, Imprudent and Insecure participants decreased less their accident rate in the second session and, at the same time, reported lower percentages of safe scores than Prudent drivers, but higher percentages of almost safe scenes, resulting in a less safe, although improved, driving performance. These results confirm those reported in Gianfranchi *et al*. (2018), shading light also on the difference between the profiles in terms of learning trends and timing. Moreover, they also confirm all the results of the previous studies (Tagliabue *et al*. 2017; 2019; Tagliabue & Sarlo, 2015; Vidotto *et al*., 2011) regarding the improvement in driving safety as a function of improvement in hazard perception, extending these results to a specific population, *i.e*. totally inexperienced adolescents.

Concerning the follow up, our results show that a significant decrease in the frequency of bike crashes is present in the 12 months after the end of the training, controlling for any change in the frequency bike use. Besides confirming the participants' opinion, who for the most part rated the HRT training as quite or very useful, this result is crucial because provides evidence of the generalizability of the effects of the training to real on-road contexts with a vehicle different from that employed for the training. Moreover, the absence of difference in this results in terms of driving profiles allows to conclude that the generalization of the effect occurs independently of the classification made through our driving assessment protocol.

However, this result does not exclude the value of the driving assessment developed with the HRT. Indeed, although a good degree of generalization seems to be present independently of the driving profiles, the assessment is still necessary to identify more at-risk participants, to have the chance to develop *ad hoc* training protocols that can be focused on participants' weak points. Moreover, it is worth noting that the results concerning the generalizability is referred only to bike crashes. Clearly, this is due to the participants' young age that do not allow many of them to drive a motor vehicle. Further studies would be necessary to draw conclusions regarding the generalizability of the effect to motor-vehicles.

As evident, the follow up analysis does not allow us to disentangle the effect of learning from that of age and developmental factors. Thus, the comparison between the trained group and a matched control group gains importance. Indeed, matching the trained group in terms of gender, age and on-road exposure (bike, two-wheeled vehicles, and car) allows to control for these main confounding variables. Our preliminary results concern a sub-sample of the trained participants (N=26), compared to the same number of control (untrained and matched) participants. A significant difference between groups in terms of bike crashes frequency emerged: Trained participants reported a lower frequency of bike crashes in the last 12 months compared to control participants, whereas no differences were present in the frequency of accidents and near misses in the last week.

Moreover, probably due to the dimension of the sample, no differences emerged in terms of proportion of reported near misses, although the raw data seem to indicate a trend for the trained group to report higher proportions of near misses. Overall, these preliminary analyses confirm that the decrement in the bike crash rate is probably due to the HRT training and not to confounding variables. Moreover, although not significant, the trends regarding the near misses allow to predict that the trained participants would be more prone to report near misses, both caused by themselves and by other road users. If confirmed, this result would represent another evidence in favor of the effectiveness of the HRT training. Indeed, identifying (and thus, reporting) a higher number of occurred near misses may be a consequence of the improved hazard perception mechanism, that allows to identify more easily the risky situations. This would also be testified by the high percentage of participants (85%) that have reported the occurrence of at least one HRT hazardous scene on the real road, confirming the consistency of the virtual reconstructions proposed by the simulator with the real world too.

Considering that the control group need to be completed before drawing firm conclusions, the main limits of the present data seem to rely on the general sample size, that, although adequate, can be increased to assure more statistical reliability. Moreover, although necessary for the present study, the exclusion of experienced adolescent road users prevents us from drawing conclusions regarding the generalizability of all the effects to a more experienced sample. Finally, both follow up and control group data are measured through self-report interviews, having the limits of these types of measures.

General discussion

The present Ph.D. project had two main research aims. The first concerns the development of a training protocol for on-road hazard perception in novice drivers by means of a driving simulator, the HRT. Moreover, since some sort of evaluation of the generalizability of the effects to the real world is necessary when driving simulators are employed, a follow up study and the preliminary results regarding the comparison between a trained and a control group are provided and discussed. The second aim consists in testing the possibility to employ the HRT as tool for driving assessment, developing a method to identify different driving profiles with specific degrees of risk proneness.

Coherently, the research work followed two branches. One focused on the hazard perception training, studying at the same time its underpinning mechanisms from a psychophysiological point of view. The other branch dealt with the development and testing of a driving behavior assessment method on novice and inexperienced drivers by means of the identification of driving profiles, also exploring their relationship with decision-making ability, a wide range of personality traits, and self-reported aberrant driving behaviors. As seen, the results of both the branches led to the need to test the generalizability of the effects to the real road.

Importantly, we focused our efforts on a peculiar type of road users, that is, young and novice road users. As pointed out by recent Italian and international reports (ERSO, 2018; ISTAT, 2018), this is the most at-risk category for both crashes and their negative consequences, due to a combination of developmental, experiential and personality factors (McKenna, 2012). Moreover, the identification of different driving profiles among young novice drivers may be crucial to handle the so-called "young driver problem" as additional to the "problem young driver" (Scott-Parker *et al.*, 2013).

As regards the first research branch, we started from the evidence that experienced drivers have a better hazard perception, in that they show an accurate identification of on-road hazards (Crundall, 2016), paralleled by a peculiar gaze pattern while exploring hazardous scenes (Crundall *et al*., 2002; 2003) and by more frequent SCRs (Kinnear *et al*., 2013). Moreover, Tagliabue and Sarlo (2015) found that, when comparing a passive hazard perception training (watching video-clips of hazardous scenes) and an active training on the HRT (five courses presenting a variety of hazardous scenes that the participants had to face when driving the simulator), the active training results in a higher psychophysiological activation in terms of percentage of SCRs.

Thus, we reasoned that the experienced drivers' peculiar psychophysiological pattern, reported by various research, may be the consequence of a more efficient somatic marker mechanism as compared to novice drivers. As seen in Chapter 2, in the framework of the SMH by Damasio (1994) and Bechara *et al*. (1994; 2000a; 2000b; 2005), when we need to take decisions in risky and ambiguous contexts (such as traffic), the somatic marker is re-activated on the basis of our experience with the outcomes of our previous decisions. The activation is evident in the so-called anticipatory SCRs, *i.e.*, SCRs that occur before acting in a way that we already experienced as leading to adverse outcomes (Bechara *et al*., 2000b). These SCRs act as a signal of potential hazard, allowing to prevent the occurring of the risk itself. On the basis of this body of evidence, we hypothesized that an active training on the HRT should improve novice drivers' hazard perception by acting on the somatic marker. Indeed, since experience is necessary to develop the somatic marker, having the chance to experience hazardous situations in a safe and virtual environment should improve hazard perception without being exposed to the risks of the real road.

Thus, we designed a two-step research. In the first step, the aim was to assess the behavioral and psychophysiological changes occurring in a two-session active training on the HRT, employing the same courses in the two sessions (Tagliabue *et al*., 2017). The main results of this first study indicate that the improvement in the driving performance (in terms of decrease in the crash rate in the virtual environment) was paralleled by a decrease in the percentages of SCRs along the courses in both sessions and, more importantly, by an anticipation of the SCR onset of about 3 m in the second session, *i.e.*, when facing the same courses anew. The decrease in the SCRs percentage may indicate that, because participants could act in the virtual environment in such a way to prevent the development of risky situations, less serious hazards occurred (as testified also by the reduction in the crashes) and, consequently, less frequent SCRs were elicited. Moreover, as predicted, when participants faced anew the hazardous scenes of the first session, the SCRs onset was anticipated, thus testifying that the somatic marker was activated, and that its activation probably underpinned the improvement in the safety of driving performance.

The second step of the research (Tagliabue *et al*., 2019) was aimed at investigating whether the behavioral improvement and the SCRs onset anticipation are generalizable to new situations. Moreover, we wanted to test the presence of differences, in psychophysiological terms, between active and passive trainings. To do so, we compared the electrodermal activity of two groups of participants during a three-session active (*i.e*., driving the HRT) or passive (*i.e*., watching hazardous video-clips) training. Importantly, the first two sessions employed the same courses on the HRT, whereas the last included new courses, so as to test the generalizability of the effects. The results indicated that the active group was not only able to reduce the frequency of crashes and near misses on the simulator, but also to overall improve the safety of their driving performance, as shown by the increase in the number of scenes scored as “totally safe” and “almost safe”. Concerning the psychophysiological results, three main aspects need to be considered. First, the anticipation in the SCRs onset was present only in the second session that included the same courses the participants already experienced. This indicates that, independently of the type of training (active or passive), experience (meant as having the chance to be exposed at least two times to the same hazard) represents a key factor to activate the somatic marker and to allow the development of on-road hazard perception. Second, an increase in the SCRs percentages was present in the passive group only when an accident occurred, unlike in the active group. This may indicate a failure in the passive group to discriminate among different levels of hazard, suggesting that a passive training protocol is less efficient than an active one in training hazard perception. Finally, the accurate discrimination of the active group between different degrees of risk, proved by higher psychophysiological reactivity in “almost safe”, “near misses” and “crash” scenes, emerged in the second session and was maintained through the third one, providing partial evidence of generalization to new simulated hazardous situations.

Taken together, the results of these two studies clearly suggest that hazard perception can be trained by means of a simulated environment. Specifically, an active training on the HRT allows to improve hazard perception, leading to the development of a safe driving performance through an activation of the somatic marker. This activation is evident in the more frequent (SCR percentages) psychophysiological activation to potential hazards. However, as seen, experience is necessary to train the mechanism that allows the anticipation of hazards and to generalize the anticipation to other situations, as predicted by the SMH (Damasio, 1994; Bechara & Damasio, 2005). Thus, ideally, a multi-session (*i.e.*, at least two sessions) active training in a safe environment (such as a simulator), presenting for at least two times the widest range possible of hazardous scenes, would allow a proper improvement in on-road hazard perception.

The last conclusion is also supported by the results of a third study (Gianfranchi *et al.*, submitted) that aimed at assessing the changes in some attentional ERP components during a training session on the HRT. We presented the participants with an auditory multi-feature passive oddball task while they drove along six courses on the HRT. The deviant oddball stimuli were divided into three categories: Deviant pure tones, human sounds, and traffic-related sounds. On the basis of previous evidence (Corbetta & Shulmann, 2002; Lavie *et al.*, 2014; Wester *et al.*, 2008; 2010) we expected two main results, namely an improvement in the driving performance (decrease in crash rate), paralleled by a an increase in the amplitude of the attentional ERP components (N1, MMN, P3a) elicited by the three categories of deviant stimuli in the last courses as a consequence of the increase in the amount of available attentional resources. Moreover, we expected a reduction in the amplitude of the ERP components when the participants were driving the simulator compared to a control condition (*i.e*., being presented with the oddball stimuli while watching a clip with hazardous scenes).

Overall, our hypotheses were confirmed. However, the increase in the amplitude of the ERP components in the last phase of the training was not homogenous: It was evident only for the MMN elicited by the deviant pure tones and for the P3a elicited by the traffic-related sounds. Because previous evidence (Tagliabue *et al*., 2017; 2019) pointed out that a multi-session training is necessary to achieve a safe driving performance and a good degree of hazard perception, it is possible to hypothesize that a single session on the HRT is not adequate to "free" enough attentional resources to monitor the environment. The consequence would be a strategical employment of the available resources, that would be used to monitor salient stimuli, such as traffic-related sounds, and low-complexity stimuli such as deviant pure tones. These results, beside shedding light on the attentional effects of the driving training, nicely fit with the previous results (Tagliabue *et al*., 2017; 2019), confirming that an effective training on the HRT needs to include at least two sessions. However, still no firm conclusions can be drawn regarding the generalizability of the effects to the real road context.

The second research branch of the present Ph.D. project focused on the development of an assessment protocols for driving behaviors in novice and inexperienced drivers. The idea stemmed from two types of evidence: The limits and the inconsistency of the results reported by studies that identified personality profiles differently related to various degree of self-reported risky driving behaviors (e.g., Lucidi *et al*., 2010; 2019; Ulleberg & Rundmo, 2003) or to driving performance on a simulator (Deery & Fildes, 1999; Marengo *et al*., 2012), and the need to find a robust and replicable method of assessment of driving abilities, possibly leading to the identification of "weak" driving profiles that may need *ad hoc* training, in the framework of the assessment of the "young driver problem" (Scott-Parker *et al*., 2013).

Thus, we reversed the common methodology that was based on the identification of *personality* profiles, employing the HRT to identify *driving* profiles on the basis of a wide number of objective driving parameters. We employed this assessment protocol in a first study (Gianfranchi *et al*., 2017a) aimed at identifying driving profiles among novice road users, hypothesizing that the driving profiles would be differently related to self-reported aberrant and risky driving behaviors, as measured by the DBQ (Reason *et al*., 1990) and the 3DI (Dula & Ballard, 2003) questionnaires. Moreover, the same assessment protocol was adopted in a different study (Gianfranchi *et al*., 2017b) with the aim to explore the relation between identified driving profiles and two psychological variables usually linked to driving behaviors, *i.e.* SS and decision-making ability, measured by the SSS V (Galeazzi *et al*., 1993; Zuckerman, 1994) and by the IGT (Bechara *et al*., 1994) respectively. A cluster analysis was performed in both the studies on 18 driving indices extracted by the HRT during one experimental session with five courses. The indices can be considered as a "picture" of the overall the performance, since they reflect a wide range of aspects, such as speed, brakes, overspeeding behaviors, crashes, and safety of the driving behavior.

In the first study, this methodology allowed us to identify two driving profiles with opposite driving patterns, labelled as "Prudent" and "Imprudent" respectively. Among the 18 driving parameters considered, only some of them (such as mean speed and frequency of overspeeding behaviors) resulted significant in discriminating the two profiles. However, preliminary correlations among all the parameters showed that they were all highly correlated. These correlations, although unsurprising, have two important implications. On the one hand, they confirm the presence of the so-called vehicle response validity of the HRT simulator. The vehicle response validity corresponds to the internal validity in the field of VR and driving simulators (Shechtman *et al*., 2009), that is one of the aspects that need to be considered when a simulator is employed for research or training purposes. Moreover, as seen, this collinearity ensures flexibility in the choice of variables to consider in assessing driving behaviors. The second part of the first study results indicate that, overall, Imprudent drivers showed peculiar trends in terms of self-reported aberrant driving behaviors, although always modulated by gender, on-road exposure, and order of the task. To illustrate, higher levels of on-road exposure, male gender and answering the questionnaires before facing the HRT courses contribute to increase the scores of aberrant and dangerous driving behaviors among Imprudent drivers. Consequently, since a wide amount of aspects seem to deserve consideration when assessing driving profiles, an extended and integrated assessment protocol may represent the key to reach a solution for the "young driver problem" (Scott-Parker *et al*., 2013).

In the second study, three driving profiles emerged, namely "Prudent", "Imprudent" and "Insecure". These profiles, among others, resemble those already reported by Marengo *et al*. (2012), that were identified on the basis of self-reported personality traits and correlated to the driving performance on the HRT at a later stage. Moreover, differences between the three profiles, modulated by gender, emerged in terms of SS (specifically, in its sub-dimensions TAS and DIS), as TAS level was lower in Insecure males, when compared to Insecure females and Prudent males, whereas DIS level was higher in Prudent males compared to Prudent and Insecure females. These results suggest that higher levels of SS are not directly linked to a riskier driving performance, differently from what is reported in most part of the scientific literature (*e.g*., Jonah, 1997). Indeed, they seem to be modulated by gender and, more interestingly, also by decision-making ability. The final analysis, in fact, showed that risky behaviors on the HRT (high crash rate and bad Evaluation scores) were associated to high levels of TAS only when the decision-making ability was low. Because decision-making ability was assessed through the IGT, as a behavioral measure of the somatic marker mechanism, a bad performance on the IGT may indicate a failure in the somatic marker that would lead to worse hazard perception, and, if concomitant to a high thrill-seeking level, leading to a less safe driving performance.

Overall, the results of these two studies of the second research branch (Gianfranchi *et al*., 2017a; 2017b) point out that (1) one session on the HRT is enough to identify different driving profiles among novice drivers with different degrees of on-road exposure; (2) these driving profiles are comparable to those reported in previous studies and identified on the basis of personality traits; (3) a wide number of driving parameters allows to deeply assess participants' driving performance, although the collinearity of the parameters gives also the chance to select only the most interesting ones; (4) the parameters collinearity confirm that the HRT has a good degree of vehicle response validity, *i.e*. internal validity; (5) as expected, differences among the driving profiles emerge in terms of aberrant and risky driving behaviors, as well as of SS and decision-making skills, confirming both the validity of these methodology of assessment and the correspondence with previous evidence; (6) the complex relation between personality traits, aberrant and risky driving behaviors, decision-making ability and driving profiles represent a further confirmation of the need for an integrated assessment method in the field of drivers' assessment. However, no conclusions regarding the generalizability of these findings to different road users and to the real world can be drawn from these two studies.

Thus, the next necessary step was to attempt to extend this methodology also to different types of road users, controlling for the on-road exposure variable. We decided to apply the same assessment protocol to a sample of totally inexperienced adolescents (Gianfranchi *et al*., 2018), extracting the HRT driving parameters after a first session on the simulator, so as to exclude potential confounding factors connected to driving inexperience. Again, three driving profiles emerged: Imprudent, Prudent, and Insecure, with driving patterns comparable to those found in Gianfranchi *et al.* (2017b). Moreover, a further analysis indicated that the groups did not differed in terms of driving safety in the first session on the HRT, whereas during the second the riskier attitude of the Imprudent drivers clearly emerged. This first result points out that among adolescents without any driving experience it is possible the identification of specific "embryonic" driving profiles that resemble those identified on more experienced participants and that, maybe, can be the target of *ad hoc* training protocols. The next step was the identification of different patterns of personality traits able to predict the inclusion in the driving profiles. As expected, various personality traits resulted significant. For instance, the inclusion among Prudent drivers was predicted by low levels of thrill-seeking, an internal LC, low levels of aggressiveness, and the tendency to evaluate peers' driving behavior as risky. The opposite pattern predicts the inclusion in Imprudent drivers, as compared to the Prudent; Moreover, inclusion in the Imprudent drivers was predicted, among others, also by higher levels of thrill-seeking and higher external LC than Insecure drivers. These findings confirm the importance of the personality assessment in the evaluation of driving behaviors, being consistent with both previous (e.g., Lucidi *et al*., 2010; Marengo *et al*., 2012) and more recent evidence (Lucidi *et al*., 2019) regarding the role of personality traits in influencing adolescents' driving behaviors.

However, the last and necessary step of the present Ph.D. project was still missing, that is the evaluation of the degree of generalization of the findings to the real world, both in terms of training and driving assessment. Thus, we designed a follow up involving the adolescents of the study by Gianfranchi *et al*. (2018). Indeed, the two sessions on the HRT in which they were involved were substantially a hazard-perception training. Consequently, a 12-months follow up consisting in a brief telephonic interview was developed. One-hundred-one participants answered questions regarding their crash and near misses rate (either with bike, two-wheeled vehicles and car) in the 12 months after the completion of the training. Moreover, they also evaluated the usefulness of the training on the HRT.

Concerning the circumstantial effects of the training, *i.e.*, the improvement in the safety of the driving performance *during* the training, differences between the driving profiles already identified in the previous study emerged in the learning trends, although all the participants improved their driving performance. To illustrate, during the second session, Imprudent participants, although showing an overall safer driving performance compared to the first session, reported the highest percentage of crashes and the lowest percentage of safe and almost safe scenes. The results regarding the follow up pointed out that the most part of the participants rated the training as “quite” or “very” useful. Moreover, the majority of them (85.1%) reported to have seen happening on the real road one or more of the scenes administered by the simulator. As regards the comparison in their crash rate between the 12 months before and after the HRT training, a significant difference emerged in the bike crash rate, that was lower at the follow up independently of the driving profile. Finally, a preliminary comparison between a sub-sample of the trained participants with a control group matched for age, gender and on-road experience, showed that the trained group reported significant less bike crashes in the last 12 months.

Overall, these results are promising since they point out the following aspects: (1) the HRT training is perceived as useful from the majority of the participants, who also rated as enough consistent with reality the scenes faced on the simulator; (2) the training produces effects on the simulated driving performance (by means of an improvement in hazard perception) also on totally inexperienced adolescents, confirming the findings by Marengo *et al*. (2012), Tagliabue *et al*. (2013), and Vidotto *et al*. (2011); (3) a two-session training allows to generalize its effects (crash rate) up to some degree to the real world and to a different vehicle (i.e., bike); (4) the preliminary results derived from a comparison between a sub-sample of the trained participants and a control group indicate that the effects found at the follow up are unrelated to changes in age or on-road experience, but more probably to the training itself; (5) the absence of differences at the follow up modulated by the driving profile of the participants indicates that the training is effective independently of the individual characteristics.

To summarize, the studies included in the present Ph.D. project indicate that a multi-session active training on the HRT allows the improvement of on-road hazard perception in a safe and simulated environment, with many possible positive effects for road safety. The improvement in hazard perception is based on the somatic marker, trained through the repetition of the same virtual courses at least two times and resulting in an ability to anticipate the upcoming of a hazard so as to prevent the consequent risks. At the behavioral level, the generalization to other situations can develop without the need of re-exposure: However, from a psychophysiological perspective, repeated exposure to the same hazards is necessary to train the somatic marker. Clearly, this exposure is more feasible in a safe and simulated environment than in the real world. The HRT training has also an impact on the attentional resources and on how they are distributed, in that already at the end of one session participants' attentional resources do not need to be totally focused on the driving task and are available to monitor the environment looking for salient stimuli.

On the other hand, one session on the HRT is enough to discriminate between different driving profiles. The assessment method here developed allows to identify different driving profiles among novice and even totally inexperienced road users by means of a wide range of objective driving parameters. The identified driving profiles have proved to be differently related to self-report aberrant and risky driving behaviors, as well as to SS and decision-making ability, and to be predicted by specific patterns of personality traits. Thus, it is possible to conclude that the proposed assessment method can become a tool to identify at-risk driving profiles so as to develop *ad hoc* training interventions, even *before* starting to drive in the real world (*i.e.*, adolescents).

Finally, the first evidence of a certain degree of generalization of the effects to real on-road context were reported, with encouraging results. Specifically, a decrease in the bike crash rate is present after 12 months from the end of the training in the trained group and the first data that compare the trained group to an untrained control group indicate a riskier driving performance for the control group. However, more data are needed to draw firm conclusions regarding this last result.

Concerning the main limitations of the present work, besides the fact that the control group of the last study is still incomplete and need to be increased in its size, the first is linked to the impossibility of changing the features of the scenes presented by the HRT. Indeed, if it is true that the exposition to the widest possible range of hazardous situations would be necessary to achieve a perfect degree of training, unfortunately the HRT software cannot be employed for this aim. Another limitation deals with the sample size of the studies included in this work. Indeed, although they were all generally adequate for the purposes of the studies, it would be necessary the involvement of a higher number of participants, so as to have the chance to test all the aspects related to both the training and the assessment protocol.

Not only sample size, but also other characteristics of the samples employed here need to be considered. Indeed, the focus of the work on a specific type of road users (*i.e.*, novice and inexperienced drivers) prevented us from driving predictions regarding other types of road users that would be important to assess (*e.g.*, high experienced drivers, elderly, clinical populations). Moreover, in each phase of the study self-report measures were employed to collect data regarding on-road experience, personality traits and, in the follow up, also regarding crash rate and near misses (up to now, only related to bike data). Clearly, these measures have some limitations that need to be considered (*e.g.*, social expectancy bias). Future studies may also consider attempting the development of a structural equation model able to connect personality predictors, identified driving profiles and follow up data, so as to disentangle the effects of each variable and to systematize the present findings. Finally, the fact that the positive effects of the training seem to generalize to the real world independently of the participants' driving profiles may be interpreted as an index of unreliability or uselessness of the assessment method here proposed. Actually, the reliability of the assessment method is independent of the generalization effects here discussed, in that there may be different degrees of generalization (for instance, depending on the considered vehicle) that cannot be tested with the present data. Moreover, the usefulness of the methodology is not limited to the degree of generalizability of the driving profiles to the real world but includes also the chance to use the classification in different profiles to design customized training protocols.

Conclusions

The present Ph.D. work represents a contribution to the solution of the novice drivers' on-road fatalities issue. The idea underlying the whole project was to provide and test a VR training for on-road hazard perception and to develop an assessment protocol for the identification of driving profiles with different degrees of risk proneness. We decided to employ VR and, specifically, the HRT simulator because it guarantees to expose the participants to a wide range of hazardous scenes in a safe environment and, at the same time, to precisely monitor their performance through a high number of driving parameters.

The two branches of the project nicely converged to the evidence that hazard perception can be improved through a multi-session training on the HRT, with a good degree of generalization to both different virtual scenes and real road situations, and that the HRT can be used as a tool to assess driving profiles differently related to psychological variables. The main application of the present findings may be represented by the future development of an integrated protocol of training and assessment of driving abilities, for novice and inexperienced road users. Moreover, the same protocol would be suitable also for other types of road users, depending on the needs and the characteristics of other target populations.

REFERENCES

Allen, J. P., & Brown, B. B. (2008). Adolescents, peers, and motor vehicles: the perfect storm? *American Journal of Preventive Medicine*, *35(3)*, S289-S293.

Arnett, J. J., Offer, D., & Fine, M. A. (1997). Reckless driving in adolescence: 'State'and 'trait' factors. *Accident Analysis & Prevention*, *29(1)*, 57-63.

Baxter, J. S., Manstead, A. S., Stradling, S. G., Campbell, K. A., Reason, J. T., & Parker, D. (1990). Social facilitation and driver behaviour. *British Journal of Psychology*, *81(3)*, 351-360.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52(2)*, 336-372.

Bechara, A., Damasio, H., & Damasio, A. R. (2000a). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10(3)*, 295-307.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50(1-3)*, 7-15.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Sciences*, *9(4)*, 159-162.

Bechara, A., & Martin, E. M. (2004). Impaired decision making related to working memory deficits in individuals with substance addiction. *Neuropsychology, 18*, 152–162.

Bechara, A., Tranel, D., & Damasio, H. (2000b). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123(11)*, 2189-2202.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57(1)*, 289-300.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., *et al.* (2012). Publication recommendations for electrodermal measurements. *Psychophysiology 49*, 1017–1034.

Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, *54(1)*, 106-148.

Buelow, M. T., & Suhr, J. A. (2013). Personality characteristics and state mood influence individual deck selections on the Iowa Gambling Task. *Personality and Individual Differences*, *54(5)*, 593-597.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63(3)*, 452-459.

Carroll, P. S. (1971). *The meaning of driving exposure*. HIT-Laboratory Report. Highway Safety Research Institute: University of Michigan.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3(3)*, 2011-2016.

Crundall, D. (2016). Hazard prediction discriminates between novice and experienced drivers. *Accident Analysis & Prevention*, *86*, 47-58.

Crundall, D., Chapman, P., Phelps, N., & Underwood, G. (2003). Eye movements and hazard perception in police pursuit and emergency response driving. *Journal of Experimental Psychology: Applied*, *9(3)*, 163-174.

Crundall, D., Underwood, G., & Chapman, P. (2002). Attending to the peripheral world while driving. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *16(4)*, 459-475.

Dahlen, E. R., Edwards, B. D., Tubré, T., Zyphur, M. J., & Warren, C. R. (2012). Taking a look behind the wheel: An investigation into the personality predictors of aggressive driving. *Accident Analysis & Prevention*, *45*, 1-9.

Dahlen, E. R., Martin, R. C., Ragan, K., & Kuhlman, M. M. (2005). Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving. *Accident Analysis & Prevention*, *37(2)*, 341-348.

Damasio, A. R. (1994). *Descartes' error*. New York, USA: Random House.

Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research*, *30(4)*, 225-236.

Deery, H. A., & Fildes, B. N. (1999). Young novice driver subtypes: Relationship to high-risk behavior, traffic accident record, and simulator driving performance. *Human Factors*, *41(4)*, 628-643.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134(1)*, 9-21.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage*, *34(4)*, 1443-1449.

De Winter, J. C. F., & Dodou, D. (2010). The Driver Behaviour Questionnaire as a predictor of accidents: A meta-analysis. *Journal of Safety Research*, *41(6)*, 463-470.

De Winter, J., Van Leeuwen, P. M., & Happee, R. (2012). Advantages and disadvantages of driving simulators: A discussion. In Spink, A., Grieco, F., Krips, O., Loijens, L., Noldus, L., & Zimmerman, P. (Eds), *Proceedings of Measuring Behavior 2012. 8th International Conference on Methods and Techniques in Behavioral Research* (pp. 47-50). Utrecht, The Netherlands: Noldus Information Technology

Di Stasi, L. L., Álvarez-Valbuena, V., Cañas, J. J., Maldonado, A., Catena, A., Antolí, A., & Candido, A. (2009). Risk behaviour and mental workload: Multimodal assessment techniques applied to motorbike riding simulation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12(5)*, 361-370.

Di Stasi, L. L., Contreras, D., Cándido, A., Cañas, J. J., & Catena, A. (2011). Behavioral and eye-movement measures to track improvements in driving skills of vulnerable road users: First-time motorcycle riders. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14(1)*, 26-35.

Dula, C. S., Adams, C. L., Miesner, M. T., & Leonard, R. L. (2010). Examining relationships between anxiety and dangerous driving. *Accident Analysis & Prevention*, *42(6)*, 2050-2056.

Dula, C. S., & Ballard, M. E. (2003). Development and evaluation of a measure of dangerous, aggressive, negative emotional, and risky driving. *Journal of Applied Social Psychology*, *33(2)*, 263-282.

European Road Safety Observatory – ERSO (2018). *Annual accident report 2018*. Available at: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2018.pdf.

Farah, H., Yechiam, E., Bekhor, S., Toledo, T., & Polus, A. (2008). Association of risk proneness in overtaking maneuvers with impaired decision making. *Transportation Research Part F: Traffic Psychology and Behavior*, *11*, 313–323.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18(4)*, 1030-1044.

French, D. J., West, R. J., Elander, J., & Wilding, J. M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accident. *Ergonomics, 36*, 627–644.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, *25(4)*, 355-373.

Galeazzi, A., Cavallini, G., and Levrero, A. (1993). *La ricerca di forti sensazioni in militari, sportivi e gruppi di controllo: la Sensation Seeking Scale di Zuckerman*, in Atti del III Congresso Internazionale "Latini Dies", 6-8 Maggio 1993, Toulouse.

Gandhi, R. D., & Patel, D. S. (2018). Virtual reality: Opportunities and challenges. *International Research Journal of Engineering and Technology*, *5(1)*, 482-490.

Gianfranchi, E., Mento, G., Chierchia, C., Duma, G.M., Sarlo, M., & Tagliabue, M. (submitted). Electrophysiological correlates of attentional monitoring during a complex driving simulation task.

Gianfranchi, E., Spoto, A., & Tagliabue, M. (2017a). Risk profiles in novice road users: relation between moped riding simulator performance, on-road aberrant behaviors and dangerous driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *49*, 132-144.

Gianfranchi, E., Tagliabue, M., Spoto, A., & Vidotto, G. (2017b). Sensation seeking, non-contextual decision making, and driving abilities as measured through a moped simulator. *Frontiers in Psychology*, *8*, 2126.

Gianfranchi, E., Tagliabue, M., & Vidotto, G. (2018). Personality traits and beliefs about peers' on-road behaviors as predictors of adolescents' moped-riding profiles. *Frontiers in Psychology*, *9*, 2483.

Gidron, Y., Davidson, K., & Ilia, R. (2001). Development and cross-cultural and clinical validation of a brief comprehensive scale for assessing hostility in medical settings. *Journal of Behavioral Medicine*, *24(1)*, 1-15.

Ginsburg, K. R., Winston, F. K., Senserrick, T. M., Garcia-España, F., Kinsman, S., Quistberg, D. A., Gross, J. G., & Elliott, M. R. (2008). National young-driver survey: teen perspective and experience with factors that affect driving safety. *Pediatrics*, *121(5)*, e1391-e1403.

Hink, R. F., Hillyard, S. A., & Benson, P. J. (1978). Event-related brain potentials and selective attention to acoustic and phonetic cues. *Biological Psychology*, *6(1)*, 1-16.

Horswill, M. S., & McKenna, F. P. (2004). *Drivers' hazard perception ability: Situation awareness on the road.* In S. Banbury & S. Tremblay (Eds.). A Cognitive Approach to Situation Awareness (pp.155-175). Aldershot, UK: Ashgate.

Horváth, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction?. *Biological Psychology*, *79(2)*, 139-147.

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and individual Differences*, *32(3)*, 401-414.

Istituto Nazionale di Statistica –ISTAT (2018). *Incidenti stradali. Anno 2017*. Available at: https://www.istat.it/it/files/2018/07/Incidenti-stradali_2017.pdf

Jasper, H.H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371-375.

Jonah, B. A. (1997). Sensation seeking and risky driving: a review and synthesis of the literature. *Accident Analysis & Prevention*, 29(5), 651-665.

Kaptein, N. A., Theeuwes, J., & Van Der Horst, R. (1996). Driving simulator validity: Some considerations. *Transportation Research Record*, *1550(1)*, 30-36.

Kinnear, N., Kelly, S. W., Stradling, S., & Thomson, J. (2013). Understanding how drivers learn to anticipate risk on the road: A laboratory experiment of affective anticipation of road hazards. *Accident Analysis & Prevention*, *50*, 1025-1033.

Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: A review and synthesis. *Biological Psychology*, *45(1-3)*, 19-56.

Lage-Castellanos, A., Martínez-Montes, E., Hernández-Cabrera, J. A., & Galán, L. (2010). False discovery rate and permutation test: an evaluation in ERP data analysis. *Statistics in Medicine*, *29(1)*, 63-74.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133(3)*, 339-354.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127(2)*, 267-286.

Lucidi, F., Mallia, L., Giannini, A. M., Sgalla, R., Lazuras, L., Chirico, A., Alivernini, F., Girelli, L., & Violani, C. (2019). Riding the adolescence: characteristics of personality sub-types of moped riders. *Frontiers in Psychology*, *10*, 300.

Lucidi, F., Giannini, A. M., Sgalla, R., Mallia, L., Devoto, A., & Reichmann, S. (2010). Young novice driver subtypes: relationship to driving violations, errors and lapses. *Accident Analysis & Prevention*, *42(6)*, 1689-1696.

MAIDS Motorcycle Accidents In Depth Study (2004). In-Depth Investigation of Motorcycle Accidents. Available at: http://www.maids-study.eu/

Marengo, D., Settanni, M., & Vidotto, G. (2012). Drivers' subtypes in a sample of Italian adolescents: relationship between personality measures and driving behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour*, *15(5)*, 480-490.

McKenna, F. P. (2012). Young driver challenges. *Psychologist*, *25(9)*, 672-675.

McKenna, F. P., & Crick, J. L. (1991). *Hazard perception in drivers: a methodology for testing and training. Final Report.* Behavioural Studies Unit, Transport and Road Research Laboratory, Crowthorne, UK.

Martinussen, L. M., Lajunen, T., Møller, M., & Özkan, T. (2013). Short and user-friendly: The development and validation of the Mini-DBQ. *Accident Analysis & Prevention*, *50*, 1259-1265.

Mayhew, D. R., Simpson, H. M., & Pak, A. (2003). Changes in collision rates among novice drivers during the first months of driving. *Accident Analysis & Prevention*, *35(5)*, 683-691.

Mazuryk, T., & Gervautz, M. (1996). *Virtual reality. History, applications, technology and future*. Technical Report. Institute of Computer Graphics. Technical University of Vienna.

Megías, A., Cándido, A., Maldonado, A., & Catena, A. (2018). Neural correlates of risk perception as a function of risk level: an approach to the study of risk through a daily life task. *Neuropsychologia*, *119*, 464-473.

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. *Safety Science*, *49(10)*, 1388-1393.

Miceli, R., Settanni, M., & Vidotto, G. (2008). Measuring change in training programs: an empirical illustration. *Psychology Science*, 50(3), 433-447.

Michon, J.A. (1985) *A critical view of driver behavior models. What do we know, what should we do?* In: Evans, L., Schwing, R.C. (Eds.), Human Behavior and Traffic Safety. Plenum Press, New York.

Montag, I., & Comrey, A. L. (1987). Internality and externality as correlates of involvement in fatal driving accidents. *Journal of Applied Psychology*, *72(3)*, 339-343.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42(4)*, 313-329.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24(4)*, 375-425.

Oltedal, S., & Rundmo, T. (2006). The effects of personality and gender on risky driving behaviour and accident involvement. *Safety Science*, *44(7)*, 621-628.

Organization for Economic Co-operation and Development [OECD] - International Transport Forum (2015). *Improving Safety for Motorcycle, Scooter and Moped Riders*. Paris: OECD Publishing.

Özkan, T., & Lajunen, T. (2005). Multidimensional Traffic Locus of Control Scale (T-LOC): factor structure and relationship to risky driving. *Personality and Individual Differences*, *38(3)*, 533-545.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology*, *118(1)*, 177-185.

Parmentier, F. B., Elsley, J. V., Andrés, P., & Barceló, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition*, *119(3)*, 374-380.

Parker, D., Reason, J. T., Manstead, A. S., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics*, *38(5)*, 1036-1048.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32(1)*, 3-25.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109(2)*, 160-174.

Preusser, D. F., Ferguson, S. A., & Williams, A. F. (1998). The effect of teenage passengers on the fatal crash risk of teenage drivers. *Accident Analysis & Prevention*, *30(2)*, 217-222.

Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the roads: a real distinction? *Ergonomics*, *33(10-11)*, 1315-1332.

Richer, I., & Bergeron, J. (2012). Differentiating risky and aggressive driving: Further support of the internal validity of the Dula Dangerous Driving Index. *Accident Analysis & Prevention, 45*, 620-627.

Sanders, R. L. (2015). Perceived traffic risk for cyclists: The impact of near miss and collision experiences. *Accident Analysis & Prevention*, *75*, 26-34.

Sagberg, F., Selpi, G., Bianchi Piccinini, G. F., & Engström, J. (2015). A review of research on driving styles and road safety. *Human Factors*, *57(7)*, 1248-1275.

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. *Traffic Injury Prevention*, *10(4)*, 379-385.

Scott-Parker, B., Watson, B., King, M. J., & Hyde, M. K. (2013). Revisiting the concept of the 'problem young driver'within the context of the 'young driver problem': Who are they? *Accident Analysis & Prevention*, *59*, 144-152.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis: An International Journal*, *24(2)*, 311-322.

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15(6)*, 322-325.

Slovic, P., Peters, E., Finucane, M. L., & MacGregor, D. G. (2005). Affect, risk, and decision making. *Health Psychology*, *24(4S)*, S35-S40.

Smorti, M., & Guarnieri, S. (2016). Exploring the factor structure and psychometric properties of the manchester driver behavior questionnaire (DBQ) in an Italian sample. *Testing, Psychometrics, Methodology in Applied Psychology*, *23(2)*, 185–202.

Soares, A. P., Pinheiro, A. P., Costa, A., Frade, C. S., Comesaña, M., & Pureza, R. (2013). Affective auditory stimuli: Adaptation of the international affective digitized sounds (IADS-2) for European Portuguese. *Behavior Research Methods*, *45(4)*, 1168-1181.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38(4)*, 387-401.

Steg, L., & van Brussel, A. (2009). Accidents, aberrant behaviours, and speeding of young moped riders. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12(6)*, 503-511.

Sutherland, I. E. (1965). *The ultimate display*. Paper presented at Information Processing Congress – IFIP.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, *8*, 1-13.

Tagliabue, M., Da Pos, O., Spoto, A., & Vidotto, G. (2013). The contribution of attention in virtual moped riding training of teenagers. *Accident Analysis & Prevention*, *57*, 10-16.

Tagliabue, M., Gianfranchi, E., & Sarlo, M. (2017). A first step toward the understanding of implicit learning of hazard anticipation in inexperienced road users through a moped-riding simulator. *Frontiers in Psychology*, *8*, 768.

Tagliabue, M., & Sarlo, M. (2015). Affective components in training to ride safely using a moped simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*, 132-138.

Tagliabue, M., Sarlo, M., & Gianfranchi, E. (2019). How can on-road hazard perception and anticipation be improved? Evidence from the body. *Frontiers in Psychology*, *10*, 167.

Taylor, D. H. (1964). Drivers' galvanic skin response and the risk of accident. *Ergonomics*, *7(4)*, 439-451.

Ulleberg, P. (2001). Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. *Transportation Research Part F: Traffic Psychology and Behaviour*, *4(4)*, 279-297.

Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. *Safety Science*, *41(5)*, 427-443.

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14(6)*, 435-446.

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.

Vidotto, G., Bastianelli, A., Spoto, A., & Sergeys, F. (2011). Enhancing hazard avoidance in teen-novice riders. *Accident Analysis & Prevention*, *43(1)*, 247-252.

Vidotto, G., Sica, C., and Baldo, S. (1995). *Vie di sviluppo e nuovi strumenti psicologici per l'individuazione di soggetti a rischio* in Qualità Della Vita E Sviluppo Delle Risorse Umane Nelle Forze Armate, eds L. Manfredi and L. Salvatico (Torino: UPSEL), 53–70.

Vidotto, G., Tagliabue, M., & Tira, M. D. (2015). Long-lasting virtual motorcycle-riding trainer effectiveness. *Frontiers in Psychology*, *6*, 1653.

Warner, H. W., Özkan, T., & Lajunen, T. (2010). Can the traffic locus of control (T-LOC) scale be successfully used to predict Swedish drivers' speeding behaviour? *Accident Analysis & Prevention*, *42(4)*, 1113-1117.

Wester, A. E., Böcker, K. B. E., Volkerts, E. R., Verster, J. C., & Kenemans, J. L. (2008). Event-related potentials and secondary task performance during simulated driving. *Accident Analysis & Prevention*, *40(1)*, 1-7.

Wester, A. E., Verster, J. C., Volkerts, E. R., Böcker, K. B., & Kenemans, J. L. (2010). Effects of alcohol on attention orienting and dual-task performance during simulated driving: an event-related potential study. *Journal of Psychopharmacology*, *24(9)*, 1333-1348.

Wolfe, A. C. (1982). The concept of exposure to the risk of a road traffic accident and an overview of exposure data collection methods. *Accident Analysis & Prevention*, *14(5)*, 337-340.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16(1)*, 121-134.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. Cambridge: Cambridge University Press.

Zuckerman, M., Bone, R. N., Neary, R., Mangelsdorff, D., & Brustman, B. (1972). What is the sensation seeker? Personality trait and experience correlates of the Sensation-Seeking Scales. *Journal of Consulting and Clinical Psychology*, *39(2)*, 308-321.

RINGRAZIAMENTI

Ringrazio di cuore la Prof.ssa Michela Sarlo, per la presenza, l'aiuto costante e incondizionato, per l'umanità e per la saggezza. Ringrazio la Prof.ssa Marialena Tagliabue per avermi resa più forte, sotto ogni punto di vista.

Muchas gracias también al Profesor Antonio Maldonado Lopez, que no solo fue un supervisor, sino también una ayuda, guía y referencia durante los seis meses en España.

Grazie al Prof. Giulio Vidotto, per la fiducia e gli insegnamenti; ad Andrea Spoto, per la presenza discreta, i consulti statistici, la pazienza e la schiettezza; a Giovanni Mento per avermi dato la possibilità di mettermi in gioco in un nuovo campo e per l'aiuto lungo il percorso.

Grazie a tutti i tirocinanti (pre e post-lauream) con cui ho condiviso gioie e dolori della raccolta dati (e qualche avventura spagnola) e agli adolescenti che hanno creduto, insieme alle loro famiglie, nel mio progetto.

Grazie alla mia famiglia, mamma Anna e papà Remo, a mia sorella Monica e al mio compagno Federico. Senza voi, i momenti più cupi sarebbero stati insormontabili: grazie per il sostegno e la fiducia nei miei confronti, che avete rinnovato giorno per giorno, ricordandomi che non sono e non sarò mai sola. Grazie anche a Oriano e Antonietta, per tutto il supporto lungo il percorso.

Un grazie speciale ad alcuni compagni di avventura:

A Umberto Granziol, per le lunghe giornate in B09, per avermi fatto tenere il sangue freddo, per le risate, il sostegno, il buon cibo e l'amicizia;

A Francesca Serra, per i consigli da sorella maggiore, la schiettezza, i troppi spritz, le cene di laboratorio, i consulti (personali e statistici), il supporto e le sigarette contro l'ansia;

A Luca Menghini per la fiducia e l'instancabile ottimismo. A Cosimo Leserri, per le riunioni al volo, le raccolte dati disperate, le focacce, il rispetto. A Gian Marco Duma, per la presenza rassicurante e gli incoraggiamenti. A Enrico Sella, el mejor compañero de todo el mundo, per i (sempre troppo pochi) caffè, i confronti fruttuosi e la reciproca fiducia.

Gracias también a los nuevos amigos españoles: a Laura y Maria, mi guapisimas compañeras de piso, y a José Miguel, Sabina, Irene y Clara que han sido mucho mas que compañeros de trabajo.

Un grazie agli amici di una vita, che, nonostante la distanza e i problemi quotidiani, non sono scappati durante questo lungo percorso che mi ha vista passare da studentessa a Ph.D.: Marta, Fulvio, Phil, Greta, Nicla, David, Marco, Vins e Giulia, Arianna, Chiara M. e Chiara G.

Infine, grazie a Berny, che dal 2017 mi veglia da lassù. Questa tesi è anche tua, perché se lungo il percorso non mi fossi ricordata ogni giorno della forza con cui hai affrontato le tue sfide, soprattutto la sfida contro la morte, avrei ceduto nei momenti più difficili.