BACHELORTHESIS, JULY 2012

ELECTRODERMAL ACTIVITY OF NOVICE DRIVERS DURING DRIVING SIMULATOR TRAINING – AN EXPLORATIVE STUDY

Raphaela Schnittker (s1016350)

FACULTY OF BEHAVIORAL SCIENCES

DEPARTMENT COGNITIVE PSYCHOLOGY AND ERGONOMICS

EXAMINATION COMMITTEE

First supervisor: Dr. Matthijs Noordzij
Second supervisor: Teun Lucassen, Msc

UNIVERSITY OF TWENTE.

Abstract

The goal of this study is to gain insight into physiological patterns of novice drivers (zero years of experience, never driven before) during their first driving simulation training. This training has the aim to prepare the learners for driving on-road. Physiological levels were obtained by means of electrodermal activity (EDA) measurement. EDA was traced by *Affectiva Q-sensors,* which are unobtrusive wristbands with a wireless biosensor. Five participants aged between 17 and 18 were followed during their entire simulation training program. EDA patterns manifested in orienting reflexes (skin conductance responses, indicators for cognitive involvement) and habituation effects were investigated. We expected that there are fewer orienting reflexes on the highway, due to less interaction with road users and other traffic control devices. Furthermore, there should be a negative correlation between driven speed and EDA. The theoretical framework for this was the risk homeostasis theory which claims that individuals adjust their behavior to maintain a constant level of perceived risk. The explorative analysis indicated a high variation in EDA levels across sessions and participants. Moreover, opposite habituation effects were found, showing habituation (lower EDA levels) from session to session, but an increase in EDA levels within sessions. Furthermore there were high variations in EDA levels present within sessions, which suggests variations in mental workload within sessions. Significant differences in EDA levels and habituation effects were also found between participants. Also, significant differences were found between sessions, with less orienting responses on the highway, thereby accepting the first hypothesis. A risk homeostasis could not be found in this study. Generally this study indicates that EDA measurement by means of *Affectiva Q*-sensors is a valid, unobtrusive approach to assess individual differences in EDA in novice drivers during their driving simulator training. Moreover, the study reveals that the used driving simulator is capable of eliciting predicted orienting responses in specific driving situations (i.e. differences in responses in situations on the highway) which underlines their fidelity. As this explorative study gained insight in EDA patterns of novice drivers in a simulated driving environment, it can give direction for more specific follow up research. Also it is considered to be relevant for the simulation industry, driving schools, traffic psychology and psychophysiology research in the field of human factors and ergonomics.

Samenvatting

Deze studie heeft als doel om fysiologische patronen van onervaren bestuurders (nul jaar ervaring, nog nooit eerder gereden) tijdens een rijsimulator training programma in kaart te brengen. Dit training heeft als doel om de leerlingen op echte rijlessen voor te bereiden. Fysiologische levels werden door middel van elektrische huidgeleiding (Engels: *electrodermal activity, EDA*) achterhaald. Deze werden door *Affectiva Q*-sensors gemeten, decente armbanden met een draadloze biosensor. Vijf participanten tussen de 17 en 18 jaar werden gedurende hun eerste rijsimulator training programma gevolgd. Daarbij werd een focus gelegd op oriënterende reflexen (indicatoren voor cognitief involvement) en habituatie effecten. We verwachtten dat er minder schommelingen in elektrische huidgeleiding op de snelweg te vinden zijn omdat er sprake is van minder interactie met verkeersfaciliteiten en andere bestuurders. De tweede hypothese stelde dat er een negatieve samenhang tussen snelheid en huidgeleiding bestaat. Het theoretisch kader hiervoor is de risk homeostasis theory, welke veronderstelt dat individuen hun gedrag aanpassen om een constant level van subjectief risk te houden. De exploratieve analyse vertoont een hoge variatie in EDA levels in participanten en het gehele trainingsprogramma. Verder werden tegenovergestelde habituatie effecten gevonden. Aan de ene kant heeft habituatie (afname in EDA levels) plaatsgevonden binnen het gehele programma, van sessie tot sessie. Aan de andere kant werd een toename in EDA levels binnen sessies gevonden. Ook is er een hoge variatie binnen sessies, wat suggereert dat er verschillen in de cognitive werkbelasting binnen sessies bestaan. Bovendien werden significante verschillen tussen sessies geconstateerd, met minder oriënterende reflexen op de snelweg en dus een bevestiging van de eerste hypothese. Een risk homeostasis werd niet terug gevonden. Verschillende aspecten kunnen hiervoor in aanmerking komen. Algemeen laat deze studie zien dat *Affectiva Q*-sensors een valide methode bieden om verschillen in huidgeleidingen van onervaren bestuurders vast te stellen. Daarnaast toont de studie aan dat de gebruikte rijsimulator in staat is om in specifieke rijsituaties voorspelde reacties uit te lokken (verschillen in oriënterende reflexen op de snelweg). Dit onderstreept de echtheidswaarde van de simulator. Deze gevonden resultaten van deze exploratieve studie kan richtig geven voor meer specifieke vervolgstudies. Verder worden de gevonden resultaten als relevant beschouwd voor de rijsimulator industrie, rijscholen, verkeers-psychologie and psychofysiolgisch onderzoek in de richting van human factors en ergonomie.

Table of contents

1. Introduction

The group of young, novice drivers without experience is the one with the highest number of accidents (e.g. De Craen, Twisk, Hagenzieker, Elffers, & Bruikhuis, 2009).The number of deathly crashes is highest in the group of people between 18 and 25 (WHO, 2004; CBS, 2010; SWOV 2012). According to the World Health organization, teenagers have the greatest risk of being involved in an accident, especially in the first 12 months after getting their driver's license. Research shows that the risk of accidents drops after the first year of independent driving and decreases with age (Twisk & Stacey, 2007). Although the overall accident risk of young drivers seemed to drop in the last years (CBS, 2010) the risk for male drivers is increasing (SWOV, 2012). In regard to the issue of traffic safety and driving skills of youngsters, driving schools play a central role. Here, the involvement of driving simulators as a virtual learning environment rose in popularity. Today, driving simulations earn a great focus of research in the field of human factors and ergonomics (e.g., De Groot, De Winter, Mulder & Wieringa, 2011; Brookhuis & de Waard, 2010; Lewis-Evams, de Waard, Jolij & Brookhuis, 2012).

In this study, physiological arousal of novice drivers during their first driving simulator training is explored. This simulator training aims to prepare novice drivers for the following real driving lessons on-road. Levels of arousal are obtained by means of electrodermal activity measurement. In this introduction, more background information regarding the context of this study will be outlined. First, the use of driving simulations as an education- and research instrument will be addressed. Then, the combined use of driving simulations with physiological measurement will be approached, with a later focus on EDA measurement. This will eventually result in a review of conducted research about physiological measurements in driving simulators and situations on-road. Finally, the goals and hypotheses are exemplified. Then, the method and analyses of this study are addressed. After that, the results are specified, followed by an elaborated discussion. Finally, recommendations for follow up research are given and a summarized conclusion for the findings of this study is presented.

1.2 The use of driving simulators as an instrument for education and research purposes

Already years ago one was aware of the beneficial effects of virtual reality as a learning environment. The feeling of immersion, being present in the immediate situation, evoked through personal control and movement was a new pathway for education and learning (Psotka, 1995). Research in the 1990's shows the power of virtual reality for acquiring various skills (e.g. Regian, Shebilske & Monk, 1992; Johnson, 1994 & Magee, 1993; Seymour, Gallagher, Sanziana, O'Brien, Bansal, Andersen & Satava, 2003). Today, virtual reality is widely used for the study and attainment of driving skills in experienced as well as novice drivers. Kappe and Emerik (as cited in De Winter, Wieringa, Kuipers, Mulder, & Mulder, 2007) suggested that about 100 driving simulators are used in Dutch driving schools. Today, even higher prevalence of driving simulators can be expected taking into account the constant and rapid development of technology (Boyle & Lee, 2010). The use of driving simulations offers the opportunity to train learner drivers in a controlled environment to prepare them for driving on-road. Demanding tasks in which sensorimotor skills play an important role (e.g., passing roundabouts, interactions with other road users) can be trained in a safe, virtual environment (De Winter et al, 2007). Research suggests that there is a 7% higher chance to pass the driving test when simulator training was followed. Moreover, the duration of driving training on-road is reduced when it was previously trained in a simulator (De Winter et al, 2007).

Besides their educational use, driving simulators are also used as a research instrument for the study of various driving behaviors, because they offer the possibility to study the nature of human driving behavior in a controlled environment (Fisher, Caird, Rizzo & Lee, 2011). Furthermore, users can be confronted with goal-oriented virtual worlds and their driving performance can be evaluated objectively (De Winter, Groot, Mulder, Wieringa, Dankelman & Mulder, 2009). According to Brookhuis and de Waard (2011), driving simulators are especially useful in the context of traffic safety. Due to the increased likelihood of accidents during suboptimal mental conditions (e.g. stress, fatigue) there is a compelling need for doing more research about those mental conditions. According to the authors, driving simulations thus offer a useful tool for studying this in a standardized and controlled environment.

Furthermore, also physiological measures are useful for studying driving behavior as they can access the mental conditions of an individual. The underlying basic assumption is that all facets of human behavior express themselves in physiological body reactions (Bouscein & Backs, 2008). The advantage is that physiological reactions are involuntary and uncontrollable (Bouscein & Backs, 2008). Thus, they offer a valid representation of the human's mental state and are therefore useful for studying underlying cognitions for various (driving) behaviors. As such, Brookhuis and de Waard (2011) advocate the combination of driving simulators with physiological measures. The concept of arousal is fundamental in the measurement of physiology; therefore it will be elaborated upon the next paragraph.

1.3 The physiological state of arousal

Generally speaking, arousal is a physiological state of an individual. It is commonly considered as one of the two dimensions of emotion. The state of arousal goes along a continuum, ranging from low to high. For example, it can go from deactivated to activated or from calm to excited (Picard, 2009). States of low arousal are for instance during sleep, while states of high arousal are, for example, when a person has strong emotions or engages in a demanding task. The second dimension of emotion is valence, which goes from positive to negative; or pleasant to unpleasant (Picard, 2009). Sympathetic arousal is constituted by different physiological reactions; for example an elevated heart rate, respiratory rate (electrodermal activity) or pupillary dilation. Those physiological reactions can be assessed by means of different measurement (see below for details on the measurement relevant for this study). Important to note is that those measurements can only measure arousal, not valence. Hence, obtained arousal levels can be interpreted in various ways, as they can be positive or negative when related to emotion. Arousal levels also vary for non-emotional reasons. For instance, variations in arousal levels can be due to changes in cognitive workload.

1.4 Electrodermal activity as physiological measurement

Measurement of electrodermal activity (EDA) presents one possibility to assess physiological arousal levels. Electrodermal activity is produced by the activity of the sweat glands which has its origins in the sympathetic subdivision of the autonomous nervous system. The sympathetic part handles external challenges and initiates the well-known fight-or-flight response (Picard, 2009). This means that the sympathetic branch prepares the body for action; amongst others through elevation of heart rate, blood pressure and sweating. This activity of the sympathetic subdivision is called sympathetic arousal. Sweat excretion gives rise to changes in skin conductance at the surface, labeled as electrodermal activity (Picard, 2009).

EDA is commonly split into *tonic* and *phasic* phenomena (Boucsein, 2012). Whereas the tonic part reflects the general **level** of EDA, the phasic part mirrors spontaneous electrodermal **responses** due to an external event (Boucsein, 2012).

In this study, an exosomatic recording is applied, which is the current standard. Here, through the used measurement device (elaborated in the method section), a direct current is exerted on the skin of the participants. This results in different dependent variables. The current study focusses on non-specific skin conductance responses (**NS.SCR’s**) which are derived from the phasic part of EDA. NS.SCR’s are electrodermal responses elicited by an external or internal stimulus. These are called non-specific because they cannot be related to specific events (Boucsein, 2012), which is the case in our study, where participants continuously encounter stimuli. Two parameters of NS.SCR's are analyzed: The frequency (quantity of NS.SCR's in a certain time window) and amplitude (height of those NS.SCR's in the given time window). For those parameters, both **total values (**per session) and **values per minute** (per session) are used in this study for analysis.

In laboratory settings, the measurement of EDA fulfills different purposes. It is amongst others used for assessing orienting responses, habituation effects and arousal- and stress levels (Bouscein, 2012). Furthermore, EDA measurement has proven to be useful for detecting changes in mental activity and underlying resources for task performances (Kramer, 2006). Therefore, it is also used for the assessment of cognitive workload (Kramer, 2006; recently confirmed by Mehler, Reimer and Coughlin (2012). All in all, EDA is regarded as an authentic indicator of emotion, arousal, stress-strain and cognitive workload (Brookhuis & de Waard, 2011). As cognitive workload thresholds can have crucial impacts on driving performance, the concept of cognitive workload is especially important in traffic psychology and hence in this study. The next section will go into this in more detail.

1.5 Cognitive workload and its role in the context of traffic psychology

According to Brookhuis & de Waard (2010) suboptimal mental conditions of the operator play a central role in the occurrence of accidents (e.g. through distraction, inattention or fatigue) . However, it is important to know that current views on human error deviate from this opinion (see for example Dekker, 2006, Reason, 1990). Still, the cognitive workload of the human involved is a pervasive concept in the field of human factors and ergonomics; and hence in this study. There are various classifications and definitions of workload. Generally, cognitive workload refers to the proportion of available cognitive resources an individual has for meeting task demands (Young & Stanton, 2006). Problems arise when task demands exceed the individual's cognitive capacities. Here, the multiple resource model is commonly used (Wickens, 1992). According to this model there are distinct resources for different types of cognitive processing. If a task exceeds an individual's cognitive resources, mental overload arises and level of performance decreases (Tsang, 2006). This relation is already dictated by the *Yerkes-Dodson law* (1908) which describes the relation between arousal and performance as an inverted U-shaped curve.

In the context of traffic psychology, the division of mental over- and underload is especially useful, because both states have important implications for the context of traffic safety. Mental overload expresses itself, amongst others, in stress, insufficient information processing and distraction. On the contrary, diminished attention and alertness are consequences of mental underload. More research needs to be carried out about levels of mental workload and its implications for reduced performance; and thus the occurrence of accidents (de Waard & Brookhuis, 2011). This can be achieved by combining EDA measures with traffic research.

1.6 EDA measurement in traffic research

Combining EDA measures with the task of driving is already done for several years. In 1962, Michaels (as cited in Bouscein, 2012) used EDA successfully for assessing driver's stress levels. In 1969, Preston (as cited in Bouscein, 2012) found an increased EDA rate in participants driving on country roads then in the city. He concluded that in the town, participants are more limited by traffic control devices (e.g., traffic lights, road signs), whereas country roads allow risky behavior, which explains the higher EDA rates. Several other studies proved the feasibility of this combination (e.g. Mallis & Dinges as cited in Brookhuis & de Waard (2010), Lee & Kim (2009), Helander (as cited in Bouscein, 2012). Tanaka, Ishida, Kawagoe & Kondo (as cited in Bouscein, 2012) proved that EDA is useful for tracing the level of driver's arousal.

Driving behavior studies with EDA measurements also discovered the occurrence of a risk homeostasis. According to this theory, people adjust their driving behavior to compensate for the risk they perceive. Wilde (1994) states that people continuously try to find a balance between their subjective perceived risk and the amount of risk they accept to perceive. Hence, if drivers are experiencing lower levels of risks than acceptable for them, they are prone to engage in activities which increase their perceived risk levels. Contrary, if their experienced risk exceeds the acceptable level, drivers try to reduce this by taking more caution, e.g. driving more slowly. Thus, drivers always try to keep a constant level of experienced risk (Wickens et al, 2004). As the perception of risk is accompanied by anxiety levels, EDA offers a possibility for assessing risk perceptions. Taylor (1969) was the first one who found a negative correlation between average speed and EDA. He observed that in risky situations, when high levels of EDA were present, participants slowed down. In contrast, when low levels of EDA were present, participants drove at higher speeds. Also, he found that the level of skin conductance maintained relatively constant over time. His results were thus consistent with the presence of a risk homeostasis. The concept of risk homeostasis is also investigated in this study (see end of the next paragraph).

1.7 Purpose of this study

This study aims to gain insight into the driving behavior of young learners during their first steps of driving skills acquisition. All participants of this study have zero years of driving experience on-road and start to learn driving for the first time (individuals with these characteristics are further labeled as ‘novice’). First step of this learning process is the driving simulator training in order to prepare the learners’ for driving on-road. The current study obtains EDA levels of five participants during their complete driving simulator training. It is thus an observational study, which is a unique characteristic compared to other investigated driving simulator studies with physiological measures. Those studies were conducted with participants who were specifically recruited for research purposes. Hence, they happened out of the real life context of the participants (e.g., Ting, Hwang, Doong, & Jeng, 2008, De Waard, Jessurun, Steyvers, Raggatt & Brookhuis, 1995). Moreover, those studies were carried out at one specific point of time. Thus, they could not provide a deeper insight into individual physiological reactions over a longer period of time.

This study, on the contrary, captures physiological reactions of novice drivers in a real life setting which guarantees real motivations and ambitions of the participants. Furthermore it adopts an ideographic approach by observing a small number of individuals over an extended period of time, that is, over the whole period of time the study wants to give insight about. According to Picard (2009) such an ideographic design offers the best approach to cope with the complexity of EDA; as it can change on a daily basis within the same individual in similar situations. Thus, this study offers new, unique conditions for exploring levels of arousal in a simulated driving environment.

Exploring levels of arousal in novice drivers is important in the regard of two different aspects. Arousal levels give information about the cognitive workload of an individual. As cognitive workload is a fundamental construct in the context of driving and traffic safety, it is crucial to explore its role during the acquisition of driving skills. It can be assumed that novice drivers, without any experience on road and during the process of acquiring driving skills, have completely different cognitions and emotions than experienced drivers. Gaining insight into the arousal of novice drivers hence promises a new and valuable field of research. As mentioned earlier, youngsters are the group with the highest accident rate; and most accidents are happening shortly after getting the driver’s license. Attaining more knowledge about the mental state of novice drivers during the acquiring of driving skills is thus beneficial.

There is also a second, more methodological aspect. This study is considered relevant for the simulation industry because it can give important information about the fidelity of simulations; as arousal levels are an indicator of emotional and cognitive involvement. This is especially important for novice drivers as they are preparing for driving in the real-world. Also, this study can help finding an optimal level of task complexity to assure optimal learning effects in driving simulations in the future. Finally, this study can provide additional evidence for the working of *Affectiva Q*-sensors (the used EDA measurement device, more information given in the methods) in the context of simulations.

Specific research aims and hypotheses

As mentioned before, a study like this with its ideographic and observational characteristics is not done before. Therefore, a first goal of the analysis is to be explorative. The aim is to investigate how electrodermal activity of novice drivers during their first driving simulation training is shaped in terms of orienting responses and habituation effects. Both are widely used electrodermal indices (Boucsein, 2012). An **orienting response** presents a reflex which is triggered by an external, environmental change (Boucsein, 2012). From an evolutionary perspective this is crucial for survival; as an orienting response provides the direction of attention to a source. The electrodermal response (EDR) presents one of the most important indicators for orienting responses; visible in peaks and oscillations (NS.SCR's). **Habituation effects** are identified by decreases in electrodermal response (EDR) frequency and intensity resulting from repeated stimulation (Boucsein, 2012). They are regarded as the most fundamental indicator of the learning process. Traditionally, the acquisition of skills is associated with an increase of orienting responses during the learning process, and decreases in orienting responses after the skill was acquired (e.g. Kintsch, 1964, Bower, 1961). As such, both orienting responses and habituation effects are interesting indices to investigate in the context of this study.

It is important to note that there are general individual differences in the reactivity of skin conductance. There is distinguished between individuals who generally show a high level of EDA (labeled as "electrodermal labiles") and individuals who show a general low level of EDA ("electrodermal stabiles"; Boucsein, 2012). Research indicated that electrodermal stabiles also have a faster habituation rate than electrodermal labiles (Dawson, Schell & Filion, 2007).

Another aim is to find out how electrodermal activity of novice driver's differs between different driving situations; especially on the highway. It can be assumed that on the highway, there are less orienting responses than in other driving situations, due to less environmental triggers. In urban situations there are a lot more interactions with other road users and traffic control devices (i.e. road signs, traffic lights, crosswalks). Hence, it is assumed that there is a higher frequency of NS.SCR's than on the highway, where much less control devices and interactions (thus environmental triggers) are present.

Moreover the correlation between speed and electrodermal activity on the highway will be explored. The theoretical framework for the investigation of this correlation is the above mentioned principle of risk homeostasis. In this study it is focused on driving faster or slower than the road speed permits. This is considered interesting in view of two aspects.

First, it is interesting to explore how the correlation of electrodermal activity and road speed is shaped in a simulated driving environment. Some of the above mentioned studies investigated this on-road with more experienced drivers. However, electrodermal activity of inexperienced drivers on a virtual highway is not explored before. Second, from a methodological point of view it is worth to study whether, in a study like this, a statistical calculation can shed usable outcomes. This is important for the prospective research of novice drivers in driving simulations; and hence interesting for the simulation industry as well as the empirical sciences.

Based on the risk homeostasis theory it is hypothesized the following: There is a negative correlation between driven speed and arousal. Thus, lower EDA levels are presented when driving faster than road speed. Oppositely, higher EDA levels are presented when driving slower than road speed; both in order to keep level of risk on a constant level. Summarized, besides the explorative approach the hypotheses (and sub-hypotheses) of this study are:

H1 - There are less NS.SCR's during sessions on the highway than during other sessions
H2 - There is a negative correlation between driven speed and electrodermal activity when driving on the highway.

S1 - How more often it was driven lower than permitted road speed, how higher is the level of EDA

S2- How more often it was driven faster than permitted road speed, how lower is the level of EDA.

2. Method

2.1 Participants

A total of five novice drivers participated. Of those, three were male and two female. Four participants were aged 17 and one 18 years. All of the participants were novice drivers without any experience on-road. Furthermore, all of them intended to start with driving lessons on-road after completing the simulation training. Every participant (or, if under age of consent, their parents) signed an informed consent.

2.2 Materials

2.2.1 Driving simulator

A driving simulator designed by the company Green Dino BV., located in Wageningen (the Netherlands) was used. It is one of the first driving simulators brought to the market, carrying the label “classic”. The frame of the driving simulator is metal and it is equipped with all necessary instruments for driving and controlling the virtual car (see Green Dino BV, www.greendino.nl). The driver’s position is located on the left hand side similar to real automobiles in the Netherlands. Exposure to the virtual driving environment is provided through large beamers, which surround the user on both his left and right side and the front.

Figure 1. Example of the driving simulator “classic”

2.2.2 Physiological measures

For measuring electrodermal activity, an exosomatic recording device was used, i.e. *Q-sensors*, developed by the company *Affectiva* (www.affectiva.com).
Those *Q-sensors* can be worn as wristbands without any need of connecting it to a different device. Thus, they offer a wearable measurement device without being obtrusive; and thereby facilitating measurements that do not interfere with how people would normally act in a driving simulator. *Q-sensors* are equipped with a wireless biosensor which detects skin-conductance as well as temperature and motion. They measure EDA in microsiemens (μS) with a sample frequency of 32 Hertz.

Levels of EDA are obtained by means of two silver electrodes on the back of the *Q-sensors*. Those two electrodes exert a constant voltage on the individual’s skin. As a result, changes in EDA can be obtained by variations of the partial voltages (Boucsein, 2012). The top of the *Q-sensor* has a small button. This button can be pressed in order to mark specific events, which can be found back later in the data output. In the visualization of the obtained EDA data this marker can be recognized as a blue water drop at the respective time where it was set.

Figure 2. Example of the *Affectiva Q*-sensor used in this study

2.2.3 Log files simulator

The simulator output log files for 116 different variables with a sample frequency of 8 Hertz. For every block of the training sessions, the simulator produced a distinct a set of log files. Two of those 116 variables are relevant for testing the second hypothesis this study. These are the variables “driving faster than road speed” and “driving slower than road speed”. Output of these variables is given in dichotomous terms. Value “0” means that the participant did not drive slower or faster than road speed. Value “1” means that the participant was actually driving slower or faster than the given speed limit. The tolerated band width for speed was 3 km/h. Thus, when it was driven slower than 47 km/h, the value of the variable “lower than road speed” changed from 0 to 1. Also, when it was driven faster than 53 km/h, the value of the variable “driving faster than road speed” changed from 0 to 1.

2.2.4 Videos

Every session was also recorded on video by means of a recording program on a laptop which stood next to the driving simulator. The linked webcam was located in the upper right corner in the back of the simulator. Video recordings contain visual as well as audio data.

2.2.5 Time registration document

For each unit, the time was recorded on a paper-based document. The time noted was based on the time of the computer from where the videos where recorded. Recorded start time was the moment where the participant actually sat in the virtual car, ready to start the engine. Recorded end time was the moment where the virtual drive was finished; i.e. when the car was stopped by the simulator software by the end of each unit. In the progress of the data analysis it was found out that those time registration could not be used as a reliable research device. This was due to the fact that the clocks of the simulator, the administration computer and the video recording device were not synchronous. In Section 2.4.1 it is explained how synchronization of the various data streams was realized.

2.2.6 Training program

Every participant had 20 training units. Each session consists of 2 units, thus every participant had 10 training sessions in total. Every session had a different content (see Appendix 1 for a complete overview of the training program). Every unit contained of 2-3 blocks. In those blocks, different aspects of the driving task were covered. For each unit, a corresponding report was made by the simulator software. This report contained an overview of the performance of the participant for the different blocks. After finishing the first unit, the participant shortly evaluated this report with the driving instructor. Then the participant started with the second unit of the session, also followed by a short evaluation.

2.3 Procedure

The study took place in the driving school *Auto- en motorrijschool Lucassen* located in Ruinerwold in the Netherlands (www.rijschoollucassen.nl). This driving school has one Green Dino simulator *classic*. The simulator is placed in a separate room in order to assure a quiet environment for the learners. The simulator was started via the administrator computer in the next room. The procedure of the study was the same for every participant. Before starting with the first session, each of the five participants received information about the aim and course of the study. At that time they also signed the informed consent. After that, the *Q-sensor* was put on the participants left arm. All participants used the gear shift in the simulator with their right hand, thereby producing rapid arm movements. Thus, in order to reduce motion artifacts, the left arm was used. After putting on the *Q-sensor*, the participant started to read the respective theoretical background of the session. This took usually about 5 to 10 minutes. During this time, the *Q-sensor* had enough time to find a steady skin conduction signal. Then, the participant stepped in the driving simulator. At that time, the video recording was started by the researchers. Before starting with the actual lesson, the participants set a marker by pressing the little button on top of the *Q-sensor*. That moment was always visible on the video recordings in order to facilitate later synchronization of the different data sources. Then, the simulator gave the user instructions about the unit that followed. After the instructions, the actual lesson began and the view was “zoomed” in the virtual car. That moment was noted as “start time” by the researchers on the time registration document. When the participant finished the unit, the “end time” was noted.

Each unit consisted of 2-3 blocks which were followed by a short evaluation. Then, the participant could skip the indicator switch to start with the next block. After finishing one unit, the participant stepped out of the simulator and shortly evaluated the report of the unit with the driving instructor on the administration computer. During this time, the video still recorded. Thus, all videos contain a short break of approximately 3 minutes, without a person sitting in the simulator. After the evaluation, the participant started with the second unit of the session. At the end of the second unit, the session was finished and the “end time” was administered again. After the simulator training was finished, the participant – still sitting in the simulator – set a second marker on the *Q*-sensor. The pressing of the button was, as with the first marker, clearly visible on video. Then, the participant stepped out the simulator and the recording was stopped. It followed a second evaluation with the driving instructor. After that, the *Q-sensor* was removed from the participant’s arm. Finally, the participant filled in a mental workload questionnaire, the *NASA-TLX* (for more information see Sharek, 2011). Due to the fact that results of this survey are not necessary for achieving the goals of this study, they are not included in the analyses and hence not further mentioned.

2.4 Data Analysis

2.4.1 Synchronization of data sources

The first step of the analysis was synchronization of the different data sources; i.e. synchronization of EDA data, log files and videos. The sample frequency for the EDA data was 32 Hertz and for the log files of the simulator 8 Hertz. A challenge was the format of the log files. Instead of coming in a coherent stream, they were broken up in fragments. For each block, a distinct fragment was produced. As a consequence, there were moments where the simulator did not produce log files; the moments between those fragments (see Figure 1).

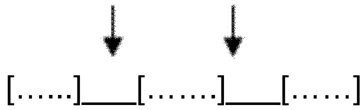


Figure 1. Example of the form of log files produced by the simulator software. The marked spaces are the production breaks between the blocks.

To overcome this, the exact moments of those breaks have to be identified in order to remove them. This was accomplished through the video recordings. For each single session, different time moments had to be found out:

1) The moment where the first marker was set by the participant

2) The moment where the production of the log files started (for all blocks the respective unit consisted of).

These steps were carried out by two bachelor students. The exact determination of the time moments required good visual abilities which could differ between individuals. This is why data inter-rater reliability was calculated for ten percent of the data (Cohen's Kappa = .787). Then, the moment where the first marker was set had to be subtracted from the moment where the production of log files started. The results of all 50 video recordings were documented in Excel. Besides, also the EDA data had to be transformed in a version which is usable for further analysis. This was accomplished through using the *Q*-software where the visualized EDA data was transformed into a comma-separated-values (CSV) format. Now it was possible to synchronize both with *MATLAB*. During this process, four missing values were detected. This concerned session two and three of participant 2 and session four and seven of participant 4. Then, variables were transmitted to SPSS where further steps of analysis were executed.

2.4.2. Analysis of synchronized data

Amount and total amplitude of NS.SCR's for each session and participant were calculated in *MATLAB*. NS.SCR's were determined by filtering the data with a Butterworth filter (cut-off frequency = 1/32 Hertz). Then, the first derivative from this signal was used to search for data points with a minimum speed difference of .0004 µSiemens/ms. There had to be a time gap of at least 700 ms between two detected NS.SCR's, otherwise they were combined to one. Also, the "through to peak" value had to be higher than 01 µS. The rest of the analysis took place in *SPSS 18.* Data plotting was carried out with *Microsoft Excel*. First of all, all data was tested for normality. Then, an explorative data analysis was conducted.

Therefore descriptive statistics calculations were carried out. For exploring differences between subjects and sessions, Mann-Whitney-U tests, Kruskall-Wallis tests and comparisons of mean ranks were conducted. For exploring habituation effects, linear regression analyses were carried out with EDA parameters as dependent variables.

Finally, for investigating the correlation between speed and electrodermal activity, the output of the log files had to be transformed into different variables for making the sub-hypotheses testable. The information of the log files was converted into two ratio values which gave values for each minute. The first ratio value presents the proportion that was driven too fast in that minute in relation to “normal” driving behavior, thus driving at the permitted road speed. The second ratio presents the proportion that was driven too slow in that minute in relation to driving at the permitted road speed. Both ratios have a maximum value of 1. Here, how greater the value of the ratio (thus, the closer to 1), how higher the proportion of driven too fast or too slow in that respective minute. For those ratio values, Kendall’s tau correlations were calculated for both EDA parameters.

All tests were undertaken with a maximum α of 0.05.

3. Results

Before carrying out any analyses, the data was tested for normality with the *Kolmogrov-Smirnov* test. The null hypothesis assuming normality was rejected with p = 0.00 (df= 1943, α= 0.05). This concerned both EDA parameters (total amplitude of NS.SCR’s and total quantity of NS.SCR’s). As a consequence, all used statistical tests were non-parametric.

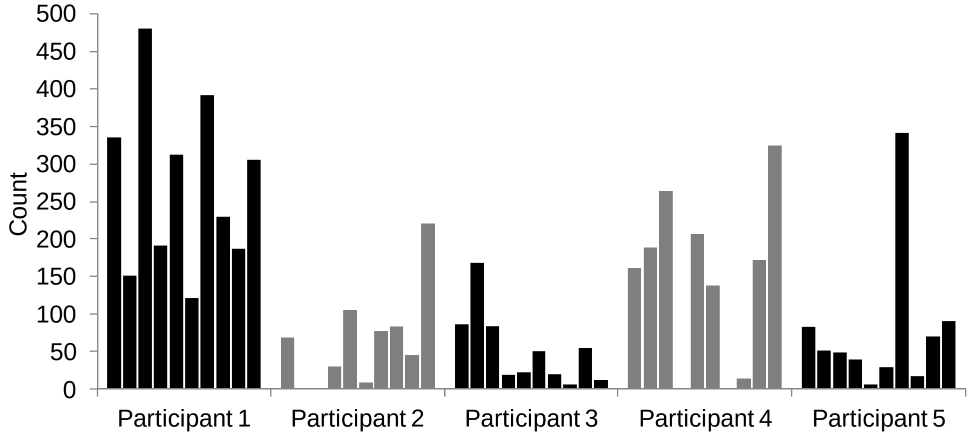
3.1 Explorative data analysis

Regarding the broadness of a data exploration, the analysis is divided into several sections. The first section explores the data by means of total EDA values for each session: *Total quantity of NS.SCR’s* per session and *total amplitude of those NS.SCR’s* per session. This is done for giving a general overview of the distribution of EDA levels. In the next section, more detailed parameters are used for making the exploration more precise: *NS.SCR’s quantity per minute* and *amplitude of those NS.SCR’s per minute*. In this section the occurrence of habituation effects was investigated as well.

3.1.1 Total NS.SCR’s quantity per session and total amplitude of those NS.SCR’s

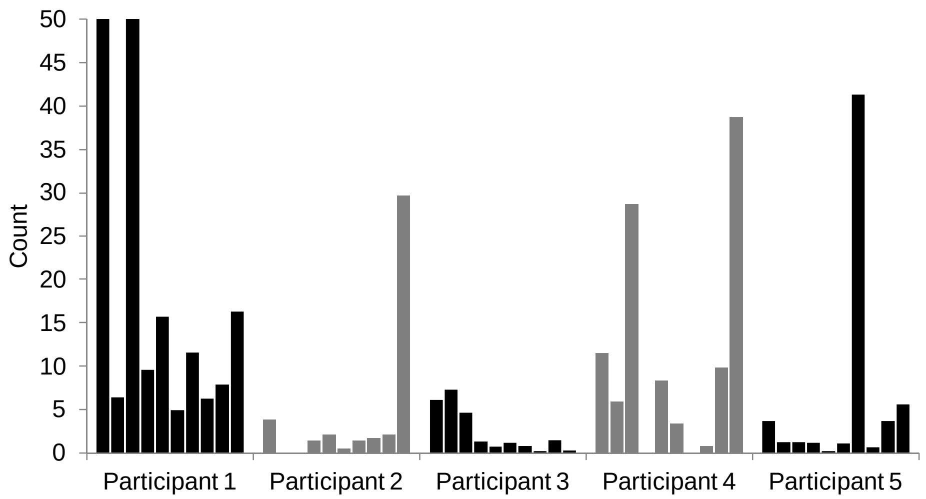
Figure 3 illustrates the quantity of NS.SCR’s for all participants for each respective session. The same overview for the total amplitude of these NS.SCR’s is given in Figure 4.

Figure 3. Compared overview of total quantity of NS.SCR'S per session



Note. Missing values for participant two (session 2 and 3) and participant 4 (session 4 and 7).

Figure 4. Compared overview of the total amplitude of NS.SCR's per session



Total NS.SCR values **per session** range from 6 (minimum) to 480 (maximum). Participant 1 has with 2707 the highest quantity of NS.SCR's in the **complete training program**; followed by participant 4 with 1466, participant 5 with 792 and participant 2 with 638 total NS.SCR's. Participant 3 has with 522 the lowest total of NS.SCR's. Values for the amplitude range from .16 to 55.29. Again, participant 1 has the highest total NS.SCR's amplitude (184.79); followed by participant 4 (106.9), participant 5 (59.57) and participant 2

(42.59). Also here, participant 3 shows the lowest values (23.75). In Table 1 you can find a summary of descriptive statistics for both parameters.

Table 1

Descriptive statistics for NS.SCR quantity and total amplitude for the complete driving simulation training

Variable	*N*	*M*	*Mdn*	*SE*	*Min*	*Max*
NS.SCR's quantity	46	132.80	85	119.33	6	480
Total amplitude of NS.SCR's	46	9.09	3.73	13.53	0.16	55.29

3.1.2 NS.SCR's quantity per minute and amplitude of those NS.SCR's per minute

In the next two tables descriptive statistics were calculated for each participant and session separately. The first table illustrates values for the NS.SCR's quantity per minute. The total average for each participant across all sessions can be found in the last row. Total average values for each participant range from .87 to 5.25. Participant 1 has the highest average of NS.SCR's per minute, participant 3 the lowest. See table below for all values. Appendix 2 shows NS.SCR quantity per minute for the first session of each participant compared with the last session. This was done to give an example for differences in distributions of NS.SCR's quantity.

Table 2.

Descriptive statistics for NS.SCR's quantity per minute for each participant and session separately

	Participant									
	1		2		3		4		5	
	M	SD	M	SD	M	SD	M	SD	M	SD
1	9.37	4.96	2.12	1.86	2.88	2.96	4.38	3.40	2.69	2.87
2	3.08	2.55	-	-	3.35	2.62	4.00	2.67	.86	1.05
3	9.48	3.25	-	-	1.40	2.00	4.95	3.28	.88	1.27
4	3.57	3.23	.41	.69	.19	.66	-	-	.64	.84
5	5.85	3.22	2.13	1.80	.29	.69	3.66	2.53	.09	.37
6	2.06	2.54	.17	.38	.91	.82	2.49	1.67	.32	.59
7	7.65	3.80	1.30	1.19	.11	.37	-	-	6.42	3.22
8	4.10	1.97	1.36	1.90	.02	.14	.19	.57	.17	.52
9	3.40	3.21	.81	1.05	.91	1.29	2.98	2.00	1.02	1.50
10	5.51	4.18	1.48	3.32	.16	.45	5.89	3.42	1.77	2.46
Total	5.25	4.07	1.48	2.06	.87	1.70	3.47	3.04	1.44	2.50

Table 3 shows mean and standard deviation for the amplitude of the found NS.SCR's per minute. Also here, values are separated by participant and session. Average amplitudes per minute across all sessions range from .04 to .34. Highest average amplitude per minute across all sessions is obtained from participant 1, lowest from participant 3.

Table 3.

Descriptive statistics for the amplitude of found NS.SCR's per minute for each participant and session separately

	Participant									
	1		2		3		4		5	
	M	SD	M	SD	M	SD	M	SD	M	SD
1	1.56	1.55	.11	.12	.21	.34	.32	.39	.11	.14
2	.13	.13	-	-	.15	.14	.11	.09	.02	.04
3	.98	.61	-	-	.08	.15	.49	.58	.02	.03
4	.17	.27	.01	.02	.01	.03	-	-	.02	.03
5	.28	.24	.04	.04	.01	.02	.13	.13	.00	.01
6	.06	.19	.00	.03	.02	.02	.06	.07	.01	.02
7	.21	.14	.02	.02	.01	.05	-	-	.80	.67
8	.10	.06	.03	.05	.00	.01	.00	.02	.00	.01
9	.14	.21	.03	.05	.02	.04	.13	.17	.03	.06
10	.26	.36	.45	.52	.00	.01	.59	.64	.10	.17
Total	.34	.62	.08	.23	.04	.12	.22	.39	.11	.33

Kruskall-Wallis-Test and confidence intervals

A Kruskall Wallis test was carried out with session and participant as grouping factors. Results show a significant effect of "session" with $\chi 2(df=9, N=1943) = 221.3$, $p < 0.001$ on the quantity of NS.SCR's per minute. This effect was also significant for the amplitude of those NS.SCR's per minute with $\chi 2$ (df = 9, N=1943) =257, $p < 0.001$. Furthermore, a significant effect of "participant" was found on both parameters with $\chi 2$ (df = 4, N=1943) = 452.59, $p < 0.001$ for NS.SCR's quantity and χ2 (df = 4, N=1943) = 528, 37, $p < 0.001$. The differences between the participants were further analyzed by comparing mean ranks for both parameters (see Table 4).

Table 4.

Mean ranks for differences between participants

	Mean Rank	
Subject	*NS.SCR's quantity per minute*	*Total amplitude of NS.SCR's per minute*
1	1390.74	1348.74
2	841.69	843.64
3	666.71	683.51
4	1203.42	1206.83
5	770.30	791.74

Participant 1 shows the highest NS.SCR's quantity per minute, participant 3 the lowest. The same is true for the total amplitude of those NS.SCR's per minute.

3.1.3 Habituation effects

To find out if habituation took place, different regression analyses were carried out for both EDA parameters as dependent variables.

NS.SCR's quantity per minute. The first regression analysis was conducted in order to search for habituation within and between sessions. For exploring habituation effects within sessions, "elapsed minutes in the simulator" was used as predictor variable. For exploring habituation effects between sessions, "session" was used as predictor variable. Note that session 10 was excluded from the analysis, as this lesson took place in a different motor vehicle (bus or truck) and could therefore bias the results for habituation effects. See Table 5 for results.

Table 5.

Results of the linear regression analysis for NS.SCR's quantity per minute

	B	SE	Beta	t	Sig.
Session	-.271	.031	-.221	-8.789	.00
Time (in minutes) in the simulator	.017	.006	.067	2.798	.00

With R=.21, R^2 = .04 and SEE= 3.127

Amplitude of those NS.SCR's per minute. The same predictor variables as for the first parameter were used. See table below for results.

Table 6.
Results of the linear regression analysis for amplitude of NS.SCR's per minute

	B	SE	Beta	t	Sig.
Session	-.035	.004	-.226	-9.477	.00
Time (in minutes) in the simulator	.002	.001	.053	2.206	.03

With R= .223, R^2= .05 and *SEE* = .38.

The Kruskall-Wallis test showed a significant effect of the number of on the levels of both EDA parameters. Therefore, a regression analysis is carried out for each participant separately in order to search for individual habituation effects.

Participant 1

Quantity of NS.SCR's per minute. Session was found as a **significant predictor** with *Beta* = -.282 (t = -5.71, p = 0.000). Time (in minutes) in the simulator was also found to be a **significant predictor** with *Beta* = .225 (t= 6.185, p = 0.000). Model summary: R= .328, R^2= .11, *SEE*= 3.85.

Total amplitude of found NS.SCR's per minute. Session is a **significant predictor** with *Beta*= -.432 (t=-9.127, p < 0.000). Time (in minutes) in the simulator was **no significant predictor**. Model summary: R = .425, R^2 = .18, SSE = .58

Participant 2

Quantity of NS.SCR's per minute. Session was found as **a significant predictor** with *Beta* = -.125 (t = -2.148, p =0.03). Time (in minutes) in the simulator was **no significant predictor**. Model summary: R= .126, R^2= .016, *SEE*= 1.50.

Total amplitude of found NS.SCR's per minute. Session is a **significant predictor** with *Beta*= -.292 (t=-5.177, p = 0.00). Time (in minutes) in the simulator is **not significant**. Model summary: R= .289, R^2= .08, *SEE*= .05.

Participant 3

Quantity of NS.SCR's per minute. Session was found as a **significant predictor** with Beta = -.408 (t = -8.494, p =0.00). Time (in minutes) in the simulator is **no significant predictor**. Model summary: R= .411, R^2= .17, *SEE*= 1.63.

Total amplitude of found NS.SCR's per minute. Session is a **significant predictor** with Beta= -.382 (*t*= - 7.832, *p* = 0.00). Time (in minutes) in the simulator is **not significant**. Model summary: *R*= .38, R^2= .14*, SEE*= .12.

Participant 4

Quantity of NS.SCR's per minute. Session is a **significant predictor** with *Beta* = -.415 (*t* = -7.615, *p* =0.00). Time (in minutes) in the simulator is also **significant** with *Beta* = .118 (t= 2.173, *p* = 0.00). Model summary: *R*= .4, R^2= .17, *SEE*= 2.56.

Total amplitude of found NS.SCR's per minute. Session is a **significant predictor** with *Beta* = -.364 (*t* = -6.642, *p* = 0.00). Time (in minutes) in the simulator is also a **significant predictor** with *Beta* = .246 (with *t*=4.48, *p* = 0.00). Model summary: *R*= .4, R^2 = .16, *SEE*= .28.

Participant 5

Quantity of NS.SCR's per minute. Session is **no significant predictor**. Time (in minutes) in the simulator is **a significant predictor** with *Beta* = -.122 (*t*= -2.343, *p* = 0.02). Model summary: *R* = .14, R^2 = .02

Total amplitude of found NS.SCR's per minute. Session is a **significant predictor** with *Beta*= .165 (*t*= 3.172, *p* = 0.00). Time (in minutes) in the simulator is **not significant**. Model summary: *R*= .166, R^2 = .03, *SEE*= .34.

Summarized, individual regression analyses showed different (opposite) habituation effects between participants. Also, the variance of EDA levels which is explained by those habituation effects differs between participants. In participant 1, the NS.SCR's rate is increasing between and within sessions. The amplitude of those NS.SCR's is decreasing between sessions, no significant changes within sessions are found. For participant 2 was found that the NS.SCR rate is decreasing between sessions. Within sessions no significant changes were found. The amplitude of those NS.SCR's is decreasing between sessions. Also here, no significant changes within sessions were found. For participant 3 the NS.SCR rate is decreasing between sessions. Within sessions, no significant changes were found. The total amplitude of those NS.SCR's is decreasing between sessions as well. Within sessions no significant differences were found. For participant 4, the NS.SCR rate is decreasing between sessions. Within sessions, the NS.SCR rate is increasing. The amplitude is decreasing between, but increasing within sessions. Finally, for participant 5 no significant changes were found between sessions for NS.SCR's quantity. Within sessions, NS.SCR rates are decreasing. The amplitude of those NS.SCR's is increasing between sessions; within sessions no significant changes were found.

3.2 Situations on the highway

In this section, a general overview of the obtained EDA data in highway situations will be given. Thus, this section will concentrates on EDA data obtained in session 7 and 8.The other part of the analysis will focus on the relation of EDA and driven speed.

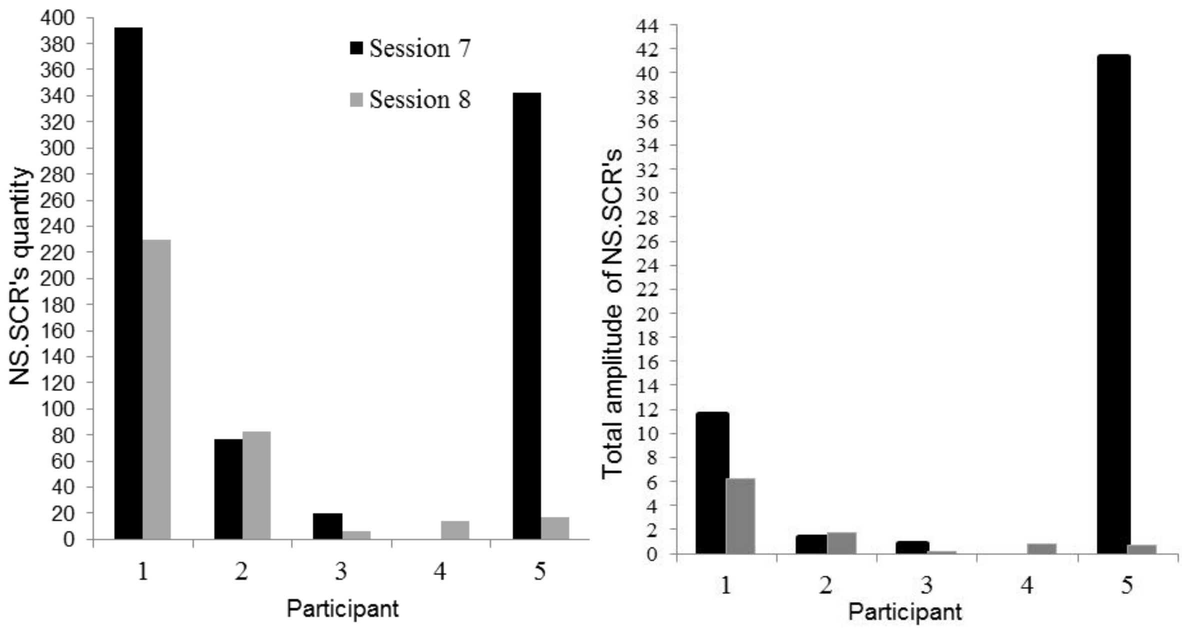
3.2.1 Total NS.SCR's quantity and total amplitude of those NS.SCR's

Descriptive statistics were calculated for all moments on the highway (see Table 7). Additionally, data plots were made to visualize the distribution of both total NS.SCR's quantity and amplitude for each participant separately (see Figure 5).

Table 7.
Descriptive statistics of both parameters for sessions on the highway (session 7 and 8)

Variable	*N*	*Mean*	*Mdn*	*SE*	*Min*	*Max*
NS.SCR's quantity	9	131.22	77	150.95	6	392
Total amplitude of NS.SCR's	9	7.18	1.39	13.34	0.17	41.32

Figure 5.
Distribution of NS.SCR's quantity and amplitude during highway sessions



3.2.2 NS.SCR's quantity per minute and total amplitude of NS.SCR's per minute in different driving situations

Data plots were made to visualize differences in EDA in highway and urban traffic situations. In Appendices 4 and 5, plots for the comparison of one example session in urban traffic with a session on the highway are illustrated. Content of the used urban traffic session were passing of small round about traffics, speed bumps and right of way situations. Goal of the highway session was to experience different speeds. The data plotting suggests higher levels of both parameters during urban traffic sessions than during highway sessions. However, for participant five the reverse is visible. Thus, no stable pattern can be recognized for these two example sessions. Further analysis is conducted in order to explore possible patterns.

H1 - Less NS.SCR's during sessions on the highway than during other sessions

For testing this hypothesis, a Mann-Whitney-U test was carried out. A nominal variable was added which separated all sessions of the whole training program by the criteria "on the highway or not". (With 0 = not on the highway; 1 = on the highway), which was then used as the grouping factor. Results show that there is a significant difference between the highway- and non-highway session's median NS.SCR's quantity per minute ($Z = -3.226$, $p = 0.001$). Furthermore, mean ranks demonstrate that lessons on the highway elicited a significant lower NS.SCR quantity than other sessions (Mean rank for on-highway sessions = 896.92; Mean rank for not-on highway sessions = 992.89).

H2 - There is a negative correlation between driven speed and electrodermal activity when driving on the highway.

Kendall's tau correlations were calculated for the two ratio values (i) driving faster than permitted road speed / permitted road speed and (ii) driving slower than permitted road speed / permitted road speed. See Table 8 for results.

Table 8.

Kendall tau rank correlations for EDA parameters and speed ratios on the highway.

EDA parameters	*Ratio too fast/ at permitted speed*	*Ratio too slow/ at permitted speed*	*N*
NS.SCR's quantity per minute	-.035 (n.s.)	-.045 (n.s.)	421
Amplitude of those NS.SCR's per minute	-.033 (n.s.)	-.050 (n.s)	421

Note. n.s. = not significant.

Both sub-hypothesis were rejected with $p>0.05$ for both NS.SCR's quantity and amplitude per minute. Thus, no significant correlation was found between levels of EDA and driving faster or slower than road speed permitted. This shows that a risk homeostasis could not be reflected in this study.

4. Discussion

Aim of this study was to explore EDA patterns of novice drivers during their first driving simulation training program. Goal of this training is to prepare the novice drivers for driving on-road. Arousal patterns of participants were observed by means of an individualized EDA measurement device. This was done during their complete training program. Hence, this study adopted an ideographic approach which made it possible to measure individual EDA patterns over an extended period of time in a real-life setting. As mentioned in the introduction, levels of EDA are highly varied within and between individuals (Picard, 2009). With its ideographic design the study gives consideration to this complexity. Regarding the novelty of this study, the first goal was an explorative analysis of the present situation; thus how EDA levels are shaped during the training program. This was guided through the investigation of two EDA parameters: Frequency and amplitude of non-specific skin conductance responses (NS.SCR's). (See Appendix 3 for EDA recordings obtained by the *Q*-sensors). For clarity of exposition, the discussion is divided into the same parts as the result section.

Total NS.SCR's quantity per session and total amplitude of those NS.SCR's per session

Summarized minimum and maximum values for all participants and sessions indicate that EDA levels varied throughout the training program and participants. The standard deviation underlines this with its even higher value for the total amplitude than the mean. As stated in the introduction, variation in EDA levels is, amongst others, an indication for variations in cognitive workload (Kramer, 2006). This suggests that cognitive workload levels were highly alternating across the training program. Total values for individual participants show a great variation between sessions as well (see Figure 3 and 4 for a review). Mehler, Reimer & Coughlin (2012) showed that skin conductance levels (SCL's) increased with task demands and, as a consequence they concluded that skin conductance levels are useful for distinguishing between different levels of task difficulty. Taking these findings into consideration, the variations in the total levels of EDA between sessions can be seen as an indicator that sessions were perceived as different cognitive demanding. The fact that general levels of EDA are different between individuals (Boucsein, 2012; Picard, 2009) is reflected in Figure 3 and 4 as well.

On the first view, no specific pattern in the distribution of EDA levels can be recognized within participants. However, the total values of NS.SCR's quantity and amplitude show differences in EDA levels between participants. Both differences (within and between participants) will be discussed in more detail upon the next section.

NS.SCR's quantity per minute and amplitude of those NS.SCR's per minute

As these two variables concern EDA levels per minute, a more detailed analysis was possible. First we are concentrating on results obtained *within* participants.
Average values per session within each participant show that EDA levels per minute vary highly *in* and *between* sessions. The fact that some standard deviations even exceed the average values per minute point outs that within sessions, EDA levels fluctuate on a high range. This is also visible in Appendix 2, which shows that the rate of NS.SCR's vary at different moments during sessions. Hence, the arousal within sessions is not stable, but changes continuously. Reason for that could be the fact that the participants are continuously busy with accomplishing the task requirements given by the simulator software, thereby showing increased cognitive workload (Kramer, 2006). Moreover, Verwey (2000) suggests that inexperienced drivers show more cognitive workload as they continuously have to think about what to do next, whereas experienced drivers possess about a repertoire of standard solutions. The same could apply to the participants in this study. As complete

novice drivers they do not have an available set of standard solutions yet. Thus, they continuously have to consider how to deal with upcoming situations, which is reflected in continuously changes in levels of EDA. This would also be an indicator for the perceived genuineness of the driving simulator; as it elicits similar arousal patterns that are suggested to occur in the real world of inexperienced drivers.

Besides, total averages per minute demonstrate a wide scope of fluctuation of EDA levels *between* the five participants. Indeed, individual differences in EDA levels are significant in this study. Thus, there is evidence of individual differences in cognitive workload levels amongst the participants. However, as already noted, are there general individual differences regarding EDA reactivity (Dawson, Schell & Filion, 2007). Thus, the found EDA differences could also (partly) be due to already existing baseline differences in EDA. As this study has no baseline comparisons, no absolutely valid conclusions can be drawn about EDA differences induced by the simulation environment itself (see recommendations section for an elaboration). Mean ranks show that the first participant experienced significant higher levels of arousal than the other four participants. In Figure 3 and 4 these differences are easy to detect. Investigating specific reasons for individual differences lies beyond the scope of this study, but would be interesting to explore in follow up research. Boucsein (2012) underlines that EDA levels are influenced by various others factors, which can account for individual differences, i.e. personality, demographic characteristics or climatic conditions.

From another point of view, the findings of individual differences in EDA levels provide evidence for the usefulness of *Affectiva Q*-sensors. They are able to detect individual differences of cognitive workload in novice drivers; thereby underlining their validity for obtaining physiological reactions in the context of simulated driving environments.

Habituation effects

Although the data plots in Figure 3 and 4 did not suggest a certain pattern of total EDA levels, an opposite habituation effect for within- and between-sessions was found. From session to session, the NS.SCR's quantity is dropping. The same is found for the amplitude. On the contrary, within sessions NS.SCR's frequency and amplitude is rising, also called sensitization (Boucsein, 2012).

The found habituation effect between sessions provides evidence that, across all participants, general habituation took place during the simulator program. As the initial situation stayed the same (exposure to a virtual driving environment), this is in line with the literature which states, that habituation effects occur after repeated, external stimulation (Boucsein, 2012, Dawson, Schell & Filion 2007). This suggests that – disregarding the content of the different sessions – there is a general habituation for driving in a virtual environment. This seems logical, as the participants get used to guide the virtual vehicle and improve their sensorimotor skills. This habituation effect is interesting, as it could account for several other things. The already mentioned *Yerkes-Dodson law* states that low levels of arousal are detrimental for performance; especially when the task is simple. Thus, habituation effects by means of EDA measurement can shed light into reasons for performance levels of novice drivers in simulated environments. It could further prevent for misinterpretations of performance reports. For example, if low levels of performance are observed it is generally assumed that more exercise is necessary. However, if low performance is due to a low level of arousal resulted from habituation (Dawson, Schell & Filion, 2007), the key for a better performance would be to enhance the task difficulty instead of exercising the same task over and over again. With the found habituation effects, this study gives some direction for further research in finding optimal levels of task complexity. This, in turn, could be significant for driving schools and the simulation industry.

On the contrary, within sessions NS.SCR's quantity per minute is rising with the number of minutes the participants were in the simulator. Reason for that could be the fact that the difficulty of tasks to accomplish increases within sessions. For example, in the beginning of one session the participants have to negotiate a crossing. Here, they have to give way to other road users. Then, when exercised a couple of times, they have to negotiate crossings with more traffic, thereby enhancing task difficulty. Thus, increased EDA levels towards the end of the session could be due to enhanced task difficulty, what also would be in line with the findings of Mehler, Reimer & Coughlin (2012). The rising of EDA levels within sessions could give an explanation for the role of the lessons content. It can be argued that the content of lessons is no underlying factor for the found habituation effects *between* sessions, because EDA levels still increase *within* sessions.

However, the R^2 for both habituation effects suggests that only a very low proportion of the variation in EDA levels is explained by both models. Thus, the greatest part of EDA variance cannot be explained by habituation effects. In fact, various other factors are responsible for the observed variance. The above mentioned differences in task difficulty within sessions could play a role, as well as gender and personality differences (Boucsein, 2012).

Aside from that, habituation effects for each participant were investigated. Results indicate three interesting findings. First, the occurrence of (opposite) habituation effects differed between participants. Second, also the portion of explained variance differed between participants. Thus, in this study, some participants are more responsive for getting habituated or sensitized to the virtual driving environment than others. Third, individual (opposite) habituation effects deviated from the general found habituation effects discussed before.

Participant 1 did not show habituation effects in the rate of NS.SCR's. Both within and between sessions the NS.SCR's quantity is increasing in this participant. Only the total amplitude is decreasing between sessions. Participant 3, on the contrary, shows no significant increases in arousal levels within sessions, but habituation effects between sessions. Also participant 2 only showed habituation effects and no increases in EDA levels within sessions. Interestingly, participant 1 was also the one with the highest and participant 2 and 3 the ones with the lowest total levels of EDA (see section 3.1.1 for a review). This suggests a relation between responsiveness to habituation and general arousal levels, which is in line with the literature. As stated earlier, there are general differences in EDA reactivity. There are individuals with a general high level of EDA ("labiles") and others with a general low level of EDA ("stabiles") (Dawson, Schell & Filion, 2007, Boucsein, 2012). Research shows that this general reactivity is also related to habituation rates. It was found that individuals with a general high NS.SCR's frequency show slow habituation rates and individuals with a general low NS.SCR's frequency are habituating faster (Boucsein, 2012). Significant correlations between habituation indices and the general individuals' NS.SCR's frequency were found (Martin & Rust, 1976; Siddle & Heron, 1976; Vossel & Roßmann, 1982 as cited in Boucsein, 2012). This was confirmed by factor-analytic studies of Martin and Rust (1976, as cited in Boucsein, 2012) which indicated that general reactivity is an underlying factor for habituation indices.

Results of this study are in line with these findings. Participant 1 did not show significant habituation effects in NS.SCR's quantity during the training program; and he is also the one with the highest EDA levels. On the contrary, he is showing increases within and between sessions, and these increases explain a moderate portion of the variance in his EDA levels. Both confirm the above mentioned research findings. In contrast, participant 3 has the lowest levels of EDA and shows habituation between sessions. This habituation also explains a relatively high portion of the variance in his EDA levels. Furthermore he does not show significant increases of EDA levels within sessions. This suggests a general lower EDA reactivity. Similar findings were obtained for participant 2. For participant 4 and 5 no direct relations between general EDA reactivity and habituation effects can be observed. However, for participant 5, the variance in EDA levels explained by habituation effects is very low anyway, indicating that other factors play more important roles.

More research needs to be conducted about this, as this study is not representative regarding its ideographic design and sample characteristics. Still, differences in habituation effects could have important implications on traffic safety. They could shed more light in the occurrence of accidents. As levels of EDA drop with habituation, this could lead to mental underload. As discussed earlier, mental underload can lead to reduced alertness and attention (Brookhuis & de Waard, 1997); both can have implications for driving safety.

However, also here, the low R^2 for some of the found habituation effects have to be taken into consideration when drawing conclusions. In some cases only a small portion of the variance can be explained by habituation effects; whereas (the combination of) other factors offer better explanations for the distribution of EDA levels (see above). Interestingly, the portion of variance explained by the regression models differs between genders. For the female participants (2 and 5) the explained portion is very low, whereas for the male participants (1, 3 and 4) the portion is much higher (review section 3.1.2 for specific values). Investigating reasons lies beyond the scope of this thesis, but would be interesting to explore in follow-up research.

Highway

Another aim of this study was to investigate differences in EDA in different driving situations. Here, it was focused on differences between EDA levels in situations on the highway and all remaining driving situations (e.g. urban traffic). A general data exploration illustrates that, again, both EDA parameters display a great variance with very low minimum values and high maximum values as well as standard deviations. Thus, arousal levels vary also on highway sessions. For further investigating differences in driving situations, data plots for comparisons between EDA levels in one highway and one urban traffic session were made (see appendices 4 and 5). The data plots reveal high variations in NS.SCR's frequency and amplitude between participants on the highway and in urban traffic. A specific pattern cannot be recognized in all participants. In the first four participants the amount of NS.SCR's seems to be lesser on the highway. However, this pattern is reserved for participant five. Thus, no overall conclusions can be drawn for the first instance. Nevertheless, the data plots only show comparisons for respectively **one** session. More statistical analyses have to be carried out which include the comparisons all highway and non-highway sessions (see next paragraph).

Hypothesis 1

The first hypothesis stated that there are less NS.SCR's during situations on the highway than during other driving situations. Argumentation for this was that there should be fewer orienting responses (i.e. NS.SCR's) on the highway than during other driving situations. As NS.SCR's arise through orienting responses which are elicited by environmental stimulation (Boucsein, 2012), there should be less NS.SCR's on the highway; due to much less interactions with other road users and traffic devices. Results indeed indicate that the amount of NS.SCR's is significant lesser on the highway, which is in line with what is logically deducted from the literature (Boucsein, 2012). This can be seen as another proof for the fidelity of the used driving simulator; as it can reproduce physiological responses which are derived from what is expected in the real world. Nevertheless, another point has to be considered. Mehler, Reimer & Coughlin (2012) state that situations on the highway give participants more control over their task demands than other driving situations, as they can select for speed or the lane they are using. In other driving situations is this not possible, as conditions are more determined by, for instance, traffic control devices or interactions with other road users. The impact of this self-regulation of cognitive workload levels could provide some additional explanation for the lower EDA levels on the highway.

Hypothesis 2

The second hypothesis aimed to test if there is a negative correlation between driven speed and EDA levels on the highway. Theoretical framework for this hypothesis was the principle of risk homeostasis, as past research already claimed its existence during the task of driving (e.g. Wilde, 1994, Taylor, 1964, Preston, 1969, Heino, van der Molen & Wilde (1990). The first sub-hypothesis stated that, the more the participants drove slower than the permitted road speed, the higher the level of EDA would be. This hypothesis was rejected for both NS.SCR's quantity and amplitude. The second sub-hypothesis was that the more participants' drover faster than the permitted road speed, the lower the level of EDA would be. Also this hypothesis was rejected for both EDA parameters. Thus, a risk homeostasis could not be found in this study.

For an interpretation of these findings, methodological as well as theoretical aspects have to be taken into consideration. First, the conducted statistical calculation does not offer a perfect solution for investigating a direct correlation. For testing the concept of risk homeostasis as Wilde (1994) suggested, absolute speed rates must have been used for the calculation of a correlation with levels of EDA. However, absolute speed rates were not recorded by the simulator program. Thus, ratios instead of absolute values were used, which makes the calculation more indirect.

Besides, the presence of a risk homeostasis in driving situations is challenged by some critics. Slovic and Fischhoff (1982, as cited in Trimpop, 1996) reveal that a major drawback is the negative connotation of risk taking. Risk taking is also a motivational behavior to gain positive feelings like proud, mastery of one's own skills or success. Risk homeostasis theory does not take these positive aspects of taking risks into consideration, although especially those could have serious implications for traffic safety (Trimpop, 1996). Also, Orr (1982, as cited in Trimpop, 1996) refers to the difficulty an individual has to assess the likelihood of risks at the direct moment of decision making (e.g. selection of road speed). Other critics challenge Wildes' claim that there is a *target* level of risk that drivers' strike to achieve. This presumes that driver's constantly compare their perceived and target risk, a claim for which no sufficient evidence is found yet (Rothengatter, 2002). Related to this it is suggested that individuals are not necessarily able to adapt their driving behavior to changes in situations where only a low risk is present (Fuller, McHugh & Pender, 2008), similar to the claims of Orr (1982) mentioned above. Thus, the nature and underlying assumptions of risk homeostasis are still questioned, which could provide another explanation that it could not be found in this study.

Furthermore, the interpretation of EDA levels has to be taken into consideration. As mentioned before, EDA measurement provides indications for levels of arousal, but not for the valence (Picard, 2009). Consequently, there remains ambiguity in interpreting patterns of arousal, as they can stand for different emotions and cognitive states. EDA measures are widely used for the assessment of emotions, information processing, attention, cognitive workload, anxiety, abnormal behavior and more (Boucsein, 2012, Picard, 2009, Dawson, Schell & Filion, 2007). Clearly, regarding the context, changes in arousal levels were here mainly interpreted as changes in cognitive workload. However, the theory of risk homeostasis is about risk perception and thus linked to the feeling of anxiety (Taylor, 1964). As stated in the introduction, for this hypothesis, changes in EDA levels are interpreted as changes in risk perception. Noticeably, there is thus some ambiguity in the interpretation of EDA levels. Perhaps, changes in EDA levels do not indicate changes in risk perception in this study.

A related explanation could be that, in this study, the driving simulator is not able to elicit risk perceptions in relation to the driven speed on the highway. This would be contrary to the found results in a real-world setting found by Taylor (1964). . Although some other studies found a risk homeostasis in a driving simulator (e.g. Jackson, 1994), those studies did not measure EDA but limited conclusions to observable declines in speed. Also, they had an experimental design and therefore did not have the challenge of possible habituation effects. Here, the general found habituation effect between sessions could have an influence on the risk perception of the participants. As highway sessions take place in the end of the training curriculum, the participants could have gotten used to the virtual environment and thereby to the virtual risk they are perhaps exposed to. As a consequence, they do not automatically adapt their driving behavior to differences in EDA patterns, as these patterns do not stand for a perception of risk. Thus, no significant correlation between EDA levels and driven speed could be found, as it is the case in this study.

Finally, the general risk perception of youngsters has to be addressed. Research persistently states that the risk perception of young drivers is a general problem which is greatly involved in the high accident rates among young drivers (review introduction for statistics). For example, Fergusion (2003) reveals that inexperienced drivers do not have sufficient experiences to appraise the degree of a given hazard. This, on the other hand, is one condition for the occurrence of a risk homeostasis (Wilde, 1994). Thus, the participant in this study could simply not have enough abilities to estimate upcoming risks, and hence did not see any reason to adapt their speed

4.1 General limitations

There are some general limitations in this study which need to be considered. First, the results cannot be generalized to the whole population of novice drivers. This study used an ideographic approach with a very small sample size. Thus, as mentioned before, there was a clear focus on individualized, long-term measurement instead of drawing nomothetic conclusions for the general population. Besides, the rather quantitative analysis method which was used in this study offers only a limited explanation for oscillations in EDA. Regarding its complexity, statistical concepts like mean, median and standard deviations provide a roughly, but of course no absolute representation of the truly obtained data. Although this study tried to combine different analyses and statistical descriptive concepts, it cannot capture all facets of EDA due to the scope of a bachelor thesis. For instance, more state-of-the art analysis techniques could have been put to test on this data set (see for example Bach, Friston & Dolan, 2010, Benedek & Kaernbach, 2010).

Also, the parameters quantity and amplitude were analyzed separated all the time. This enhanced the difficulty in interpreting EDA as a whole, coherent physiological response. Besides, it was not controlled for possible motion artifacts, although they were already greatly reduced by putting the *Q-sensor* on the left wrists.

Furthermore, for a more detailed understanding of specific electrodermal responses, the content of the lessons have to be taken into account. Here, it was only distinguished between highway and non-highway situations. It would be interesting to investigate differences in EDR between more specific contents. This study did not distinguish between lessons with a perhaps monotone, low arousing content and a perhaps exciting, high arousing content. Still, a problem would be to distinguish between monotone and exciting sessions, as differences in perception are natural.

Finally, this study has no baseline measures. Thus, EDA levels obtained in the simulator cannot be compared to EDA levels under resting conditions. Consequently, no conclusions about increases or decreases in EDA levels compared to situations outside the simulator can be drawn. Also, as noted earlier, individuals differ in their general electrodermal reactivity, which is also related to habituation rates (Boucsein, 2012). Thus, for drawing more valid conclusions about differences between individuals, EDA levels also have to be obtained under resting conditions (Boucsein, 2012). This is one recommendation for follow-up research (elaborated upon the next paragraph).

4.2 Recommendations for further research

Results of this study open the field for several new research questions. Follow up research should investigate reasons for individual differences of novice driver's during their driving simulation training (e.g. role of gender, personality traits, mood or time of training). Besides it is recommended to further explore differences in habituation effects between participants; and relate those to the actual performance levels obtained by the simulator logfiles and reports. For both, a larger sample size is implied to achieve more generalizable conclusions.

Moreover, as mentioned before, for further research it is advocated to add control moments. Making comparisons to baseline levels lies beyond the scope of this study, but would be interesting to investigate in the future. Also, as mentioned before, EDA can be regarded as a stable trait of an individual (Dawson, Schell & Filion, 2007); i.e. individuals can have the disposition to show high rates (labiles) or low rates (stabiles) of NS.SCR's. Baseline levels would control for this when drawing conclusions about individual differences during the simulator training.

Finally, it would be interesting to connect the virtual with the true reality. Exploring the relation between obtained EDA levels in the simulator and real driving lessons would reveal great insights into the fidelity of simulators. However, at least till today, the related ethical concerns restrict studies with novice drivers to the use of simulators.

4.3 Conclusion

This study investigated the physiology of novice drivers during their driving simulation training by means of electrodermal activity measurement. It was generally found that EDA levels manifest high variability. This was observed in different contexts. First, individual participants presented a high variability in his or hers EDA levels. This concerned EDA levels within sessions, suggesting that contents of sessions demanding alternating cognitive involvement. Furthermore, total EDA levels also showed a strong variability between sessions. This indicates that cognitive workload was alternating for different sessions. There is also evidence for the occurrence of opposite habituation effects. On the one hand, within each session, EDA levels increase towards the end. On the other hand, EDA levels tend to decrease between sessions, thus becoming less for each additional undertaken session. Both findings are roughly in line with what past research suggests about the relation between acquisition of skills and EDA (Klintsch, 1964, Bouwer, 1966). However, both models only explain a little variance in EDA levels, which underlines the importance of other, further to investigate factors such as individual differences.

Despite the overall variability, significant differences were found between participants.

This differences concerned general levels of EDA, as well as differences in habituation effects. Some participants were found to be more responsive for habituation than others. Habituation effects showed relation to total EDA levels of the participants, indicating that general electrodermal reactivity played a role in the habituation rate, which is in line with the literature (Boucsein, 2012) Moreover, significant differences between sessions were found.

As hypothesized, sessions on the highway produced less non-specific skin conductance responses than sessions with other driving circumstances. This suggests that driving simulations have the ability to produce research-based predicted behavior in novice drivers, which is a sign for their fidelity and usefulness for statistical analysis in this area, although all results have to be interpreted with caution. A risk homeostasis has not been found, which could be due to both theoretical and methodological aspects. Besides general critics about the nature of risk homeostasis, this finding could also be an indication that the simulator is not able to elicit risk perceptions in this study; or that the participants are not able to adapt their driving behavior to risks they perceive. But, there remains ambiguity in interpreting levels of EDA, because it can only provide insight in arousal patterns, but cannot give information about valence. Overall, the study showed that *Q-sensors* are useful for assessing EDA levels in simulated virtual environments. Also, the place on the sensor on the left wrist of the participants turns out to be a good, unobtrusive location. Generally, skin conductance measures are proven to be useful for discriminating between different levels of cognitive workload in this study, thereby confirming findings of Mehler, Reimer & Coughlin (2012).

Acknowledgements

First, I would like to thank Matthijs Noordzij and Teun Lucassen for being reliable supervisors of this research project. Then, thanks to *Auto- en motorrijschool Lucassen* where the collection of data was made possible. Here, especially thanks to Jaap Lucassen who integrated our research in his planning and thereby catered for an efficient course of events. In relation to this, thanks to Teun Lucassen for establishing this contact. Furthermore I would like to thank Martijn Broeks for being a reliable and pleasant research fellow. Finally, thanks to *Green Dino* for showing an open-minded attitude regarding the use of their simulators for our research purposes.

References

Bach, D. R., Friston, K. J., & Dolan, R. J. (2010). Analytic measures for quantification of arousal from spontaneous skin conductance fluctuations. *International Journal of Psychophysiology, 76*(1), 52-55.

Benedek, M., & Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology, 47*(4), 647-658.

Bouscein, W. (2012). *Electrodermal activity* (2nd ed.). New York: Springer.

Bouscein, W. (2004). Electrodermal measurement. In Stanton, N.A., Hedge, A. Brookhuis, K. Salas, Eduardo & Hendrick, H.W. *Handbook of human factors and ergonomics methods*. Boca Raton: CRC Press

Bouscein, W. & Backs, R.W. (2008). The Psychophysiology of Emotion, Arousal, and Per sonality: Methods and Models. In Duffy, V.G.(Ed.) *Handbook of Digital Human Model ing: Research for Applied Ergonomics and Human Factors Engineering*. Boca Raton: CRC Press.

Bouwer, G.H. (1961). Application of a model to paired-associate learning. *Psychometrika, 26*(3), 255-280

Brookhuis, K.A. & de Waard, D. (2010). Monitoring drivers' mental workload in driving simu lators using physiological measures. *Accident Analysis and Prevention,42*, p. 898–903

Brookhuis, K.A., de Waard, D. & Samyn, N. (2003). Effects of MDMA (ecstasy), and multiple drugs use on (simulated) driving performance and traffic safety. *Psychopharmacolo gy,173*, 440-445

Brookhuis, K.A. & de Waard, D. (2011). Measuring physiology in simulators. In Fisher, D.L, Caird, J.K., Rizzo, M., Lee, J.D.*Handbook of Driving Simulation for Engineering, Medicine, and Psychology: An Overview*. Boca Raton: CRC Press

Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (2007). *Handbook of Psychophysiology* (3rd ed.). Cambridge: Cambridge University Press

Centraal Bureau voor de statistiek (2010, May 26). *Helft jonge verkeersdoden zat achter het stuur*. Nederland: De Bruin, K. & Hoogenboezem, J.

Crain, de , S., Twisk, D.A.M., Hagenzieker, M.P., Elffers, H., Brookhuis, K.A. (2011). Do young novice drivers overestimate their driving skills more than experienced drivers? Different methods lead to different conclusions. *Accident Analysis & Prevention, 43*,1660-1665

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2011). Non-vestibular motion cueing in a fixed-base driving simulator: Effects on driver braking and cornering per formance. *Presence: Teleoperators and Virtual Environments, 20*, 117-142

Dawson, M.E., Schell, A.M., Filion, D. (2007). The electrodermal system. In Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (2007). *Handbook of Psychophysiology* (3rd ed., pp. 159-181). Cambridge: Cambridge University Press

De Waard, D., Jessurun, M., Steyvers, F.J.J.M., Raggatt, P.T.F., & Brookhuis, K.A. (1995). The effect of road layout and road environment on driving performance, driver's physiology and road appreciation. *Ergonomics, 38*, 1395-1407

De Winter, J. C. F., Wieringa, P. A., Kuipers, J., Mulder, J. A., & Mulder, M. (2007). Violations and errors during simulation-based driver training. *Ergonomics, 50*, 138-15

De Winter, J.C.F., Groot, de, S., Mulder, M., Wieringa, P.A., Dankelman, J. & Mulder, J.A (2009). Relationships between driving simulator performance and driving test results. Ergonomics, 52(2), 137-153

Dekker, S. (2006). *The Field Guide to Understanding Human Error.* Farnham: Ashgate Publishing

Fisher, D.L., Rizzo, M., Caird, J.K & Lee, J.D (2010). *Handbook of driving simulation for medicine, engineering and psychology*. Boca Raton: CRC Press

Fuller, R., McHugh, C. and Pender, S. (2008). Task difficulty and risk in the determination of driver behaviour. *Revue européenne de psychologie appliquée, 58*(1), 13-21.

Heino, A., van der Molen, H.H. and Wilde, G.J.S. (1990). Risk-homeostatic processes in car-following behavior: Electrodermal responses and verbal risk estimates as indicators of the perceived levels of risk during a car-driving task. Traffic Research Center, Groningen, NL.

Jackson, J.S.H. & Roger, B. (1994). A driving-simulator test of Wilde's risk homeostasis theory. *Journal of Applied Psychology,78*(6), 950-958

Johnson, D. (1994). *Virtual Environments in Army Aviation Training*. Proceedings of the 8th Annual Training Technology Technical Group Meeting (pp. 47-63). Mountain View, CA in Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science, 23*, 405-431

Kintsch, W (1965). Habituation of the GSR component of the orienting reflex during paired-associate learning before and after learning has taken place. *Journal of mathematical psychology,2*, 330-341

Kramer, A.F. (2006). *Cognitive Psychophysiology in Ergonomics and Human Factors*. International Encyclopedia of Ergonomics and Human Factors (2nd Ed.) Boca Raton: CRC Press

Lewis-Evams, B., de Waard, D. Jolij, J. & Brookhuis, K.A. (2012). What you may not see might slow you down anyway: Masked Images and Driving. *Traffic and environmental* *psychology group*, 7(1), 1-14

Magee, L. (1993). MARS Virtual Reality Simulator. Defence and Civil Institute of Environ mental Medicine Fact Sheet G-07. North York, Canada in Psotka, J. (1995). Immer sive training systems: Virtual reality and education and training. *Instructional Sci ence, 23*, 405-431

Mehler, B.,Reimer, B. & Coughlin, J.F (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: An on-road study across three age groups. *Human Factors and Ergonomics Society*, *54*(3), 396-412

Picard, R. W. (2009). Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3575-3584.

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *In structional Science, 23*, 405-431

Reason, J. (1990). *Human error.* Cambridge: Cambridge University Press

Regian, J. W., Shebilske, W., & Monk, J. (1992). A preliminary empirical evaluation of virtual reality as an instructional medium for visual-spatial tasks. *Journal of Communication, 42*(4): 136-149.

Rothengatter, T. (2002). Drivers' illusions – no more risk. *Transportation Research Part F:Traffic Psychology and Behaviour, 5*, 249-258

Seymour, M.E., Gallagher, A.G., Sanziana, A.R. , O'Brien, M.K., Bansal, V.K. , Andersen, D.K., Satava, R.M. (2003). Virtual reality training improves operating room perfor mance. *Annals of Surgery*, 236(4), 458-466

Sharek, D. (2011). A useable, online NASA-TLX tool. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2011, 1375-1379.

Taylor , D.H.(1964) Driver's galvanic skin response and the risk of accident. *Ergonomics* 7(4): 439–451.

Ting, P., Hwang, J., Doong, J. Jeng, M. (2008). Driver fatigue and highway driving: A simulator study. *Physiology and behavior, 94(*3), 448-453

Trimpop, R.M (1996). Risk homeostasis theory: Problems of the past and promises for the future. *Safety Science, 22*(1-3), 119-130.

Twisk, D.A.M. & Stacey, C. (2007). Trends in young driver risks and countermeasures in Eu rope. *Journal of Safety research , 38*, 245-257

Verwey, W. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics,43*(2), 187-209.

Wetenschappelijk onderzoek verkeersveiligheid SWOV (2012). *Jonge automobilisten en hun jonge passagiers*. Leidschendam, Zuid-Holland.

Wickens, C.D., Lee, J.D. Liu, Y. & Gorden Becker, S.E. (2004). *An introduction to human factors engineering* (2nd ed.). Upper Saddle River, New Jersey: Pearson Education.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance* (2nd ed.) New York: Harper Collins

Wilde, G.J.S. (1994) *Target Risk: Dealing with the danger of death, disease and damage in everyday decisions*. Toronto: PDE Publications.

World Health organization (2004). *World report on road traffic and injury prevention.* Geneva: Peden, M., Scurfield, R. Sleet, D. Mohan, D. Hyder, A.A., Jarawan, E. & Mathers, C.

Young, M.S & Stanton, N.A. (2006). Mental Workload: Theory, Measurement, and Applica tion. International Encyclopedia of Ergonomics and Human Factors (2nd Ed.) Boca Raton: CRC Press

APPENDICES

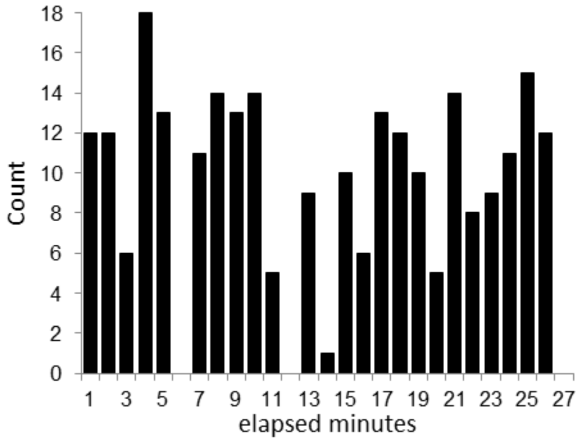
APPENDIX 1 – Complete training curriculum divided by lesson (one session contained of two lessons)

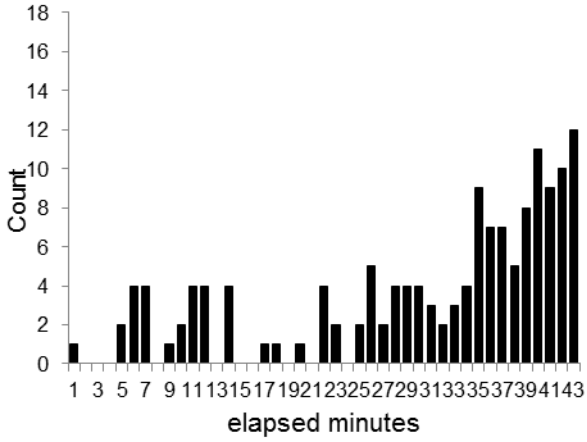
Leergang

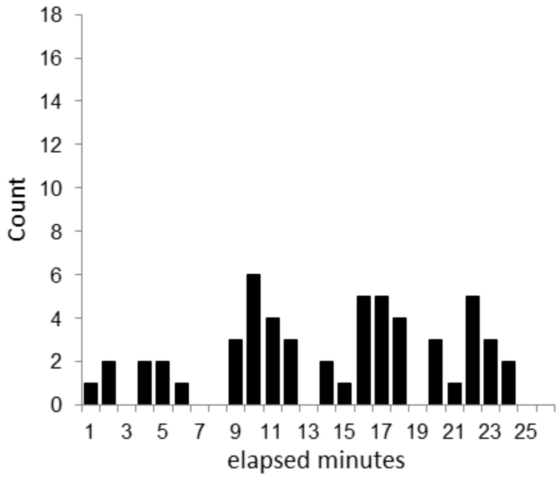
3	Bochten / Snelheid / Grip of Slip	1
4	Wegrijden / Schakelen en koppelen / Grip of Slip	1
5	Remmen / Remmen (Reactietijd) / Bochten / Gri...	2
6	Plaats op de weg (Buiten de bebouwde kom) / Pl...	2
40	Anticiperen . Afstand houden	3
42	Parkeren . toetsen vtg beheersing	3
9	Oversteken / Rechts afslaan / Links afslaan	4
10	Oversteken en afslaan / Verkeerslichten	4
11	Voorrangskruispunten	5
12	Mini-rotonden / Drempels	5
13	Gelijkwaardige kruispunten	6
14	Gevarieerde kruispunten / Toets	6
15	Snelheid / Verschillende snelheden	7
16	Inhalen / Inhalen (wisselende snelheden)	7
17	In en uitvoegen / In en uitvoegen (wisselende sn...	8
19	Weefvakken (wisselende snelheden) / Toets	8
20	Gevarieerde kruispunten met VOP's / Spoorwego...	9
21	Gevarieerde kruispunten met brug / Donker	9
24	Langzaam verkeer / Inhalen van langzamer verk...	10
25	Mist / Regen / Nacht	10

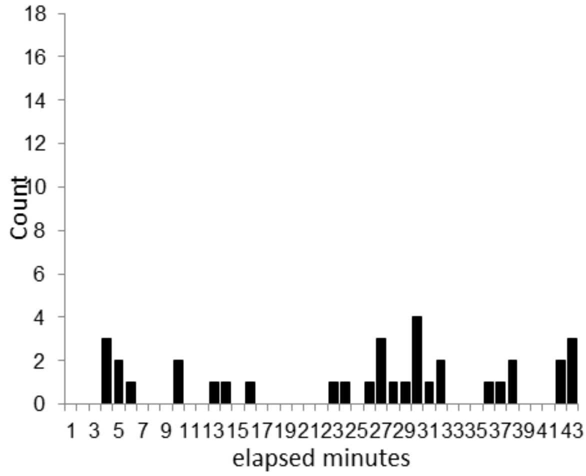
**Note.* Session 7 and 8 are highway sessions.

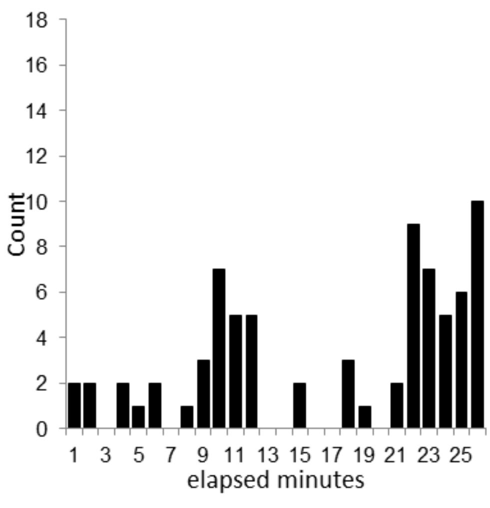
APPENDIX 2 - NS.SCR frequency per minute of the first and last session for all participants

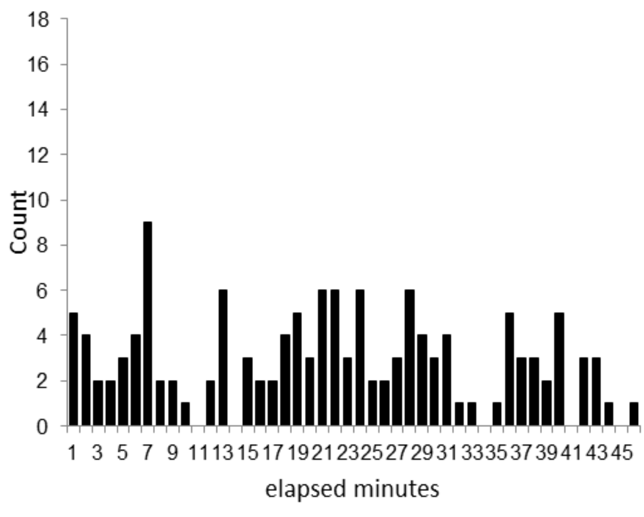


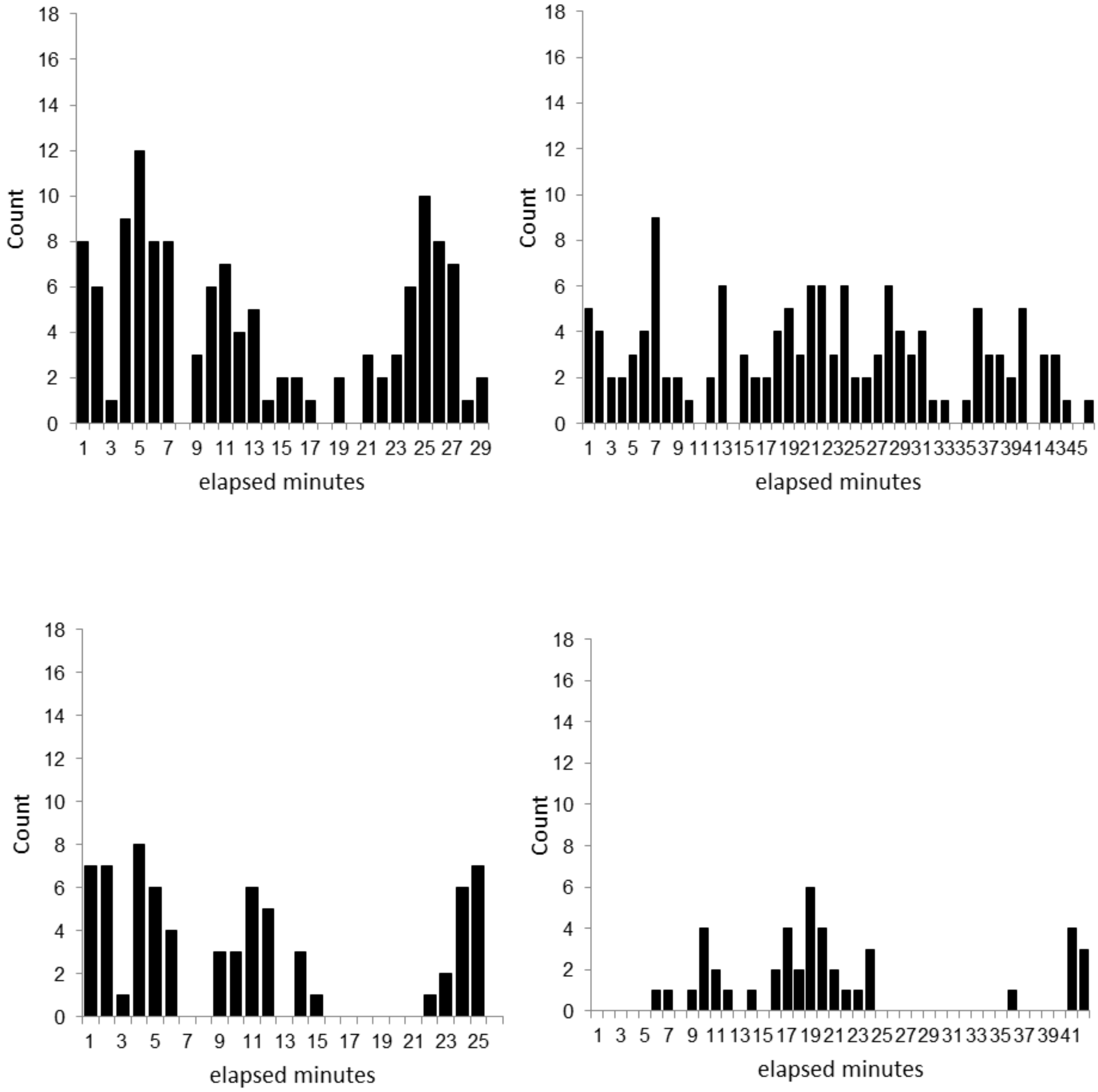








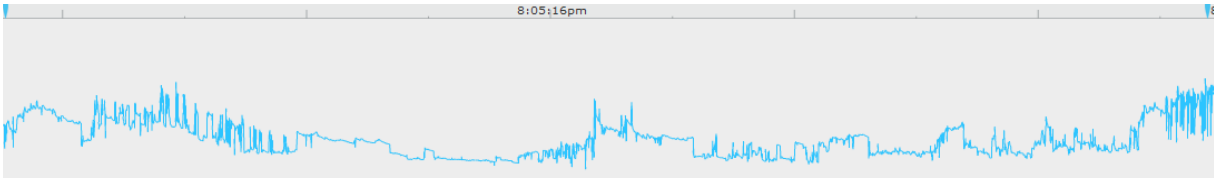



Count
elapsed minutes
Count
elapsed minutes
Count
elapsed minutes
Count
elapsed minutes

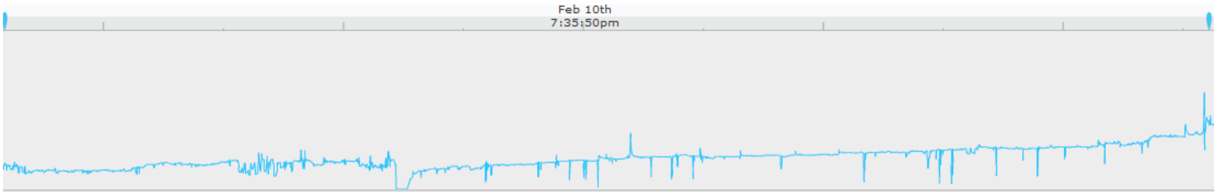
APPENDIX 3 – Electrodermal activity recordings *Q-SENSOR*

Participant 1

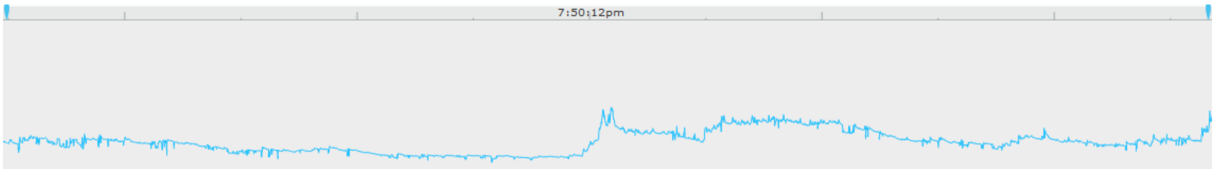
Session 1:



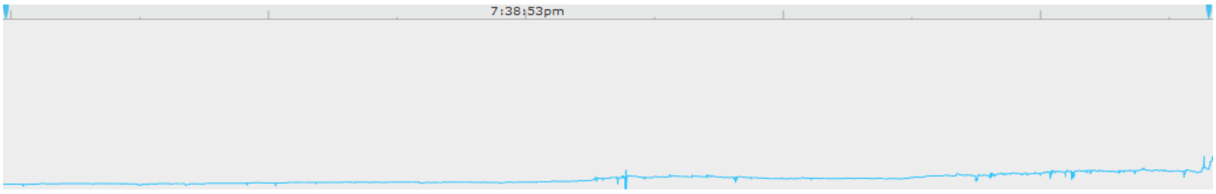
Session 2:



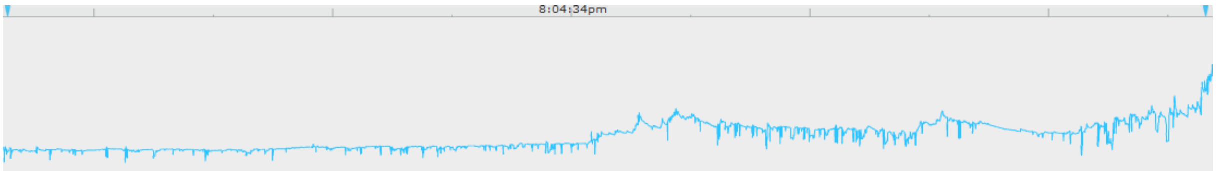
Session 3:



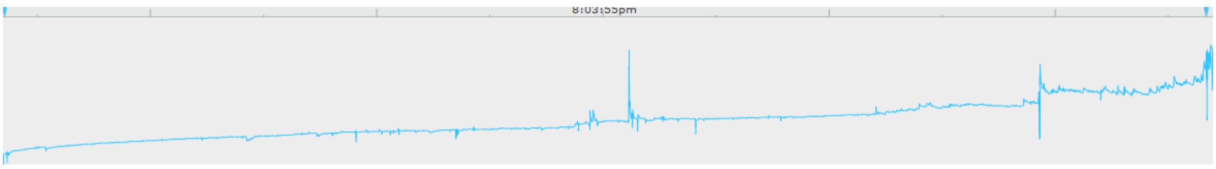
Session 4:



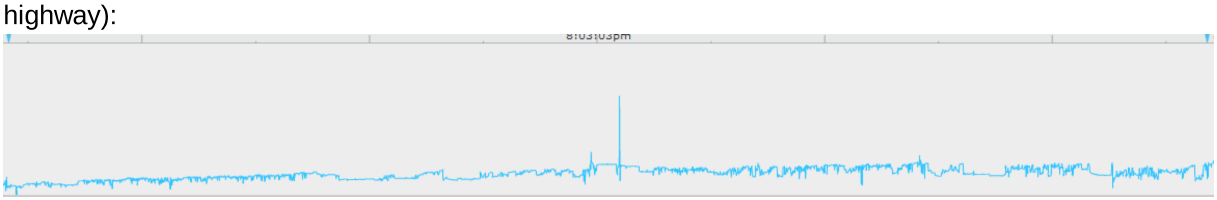
Session 5:



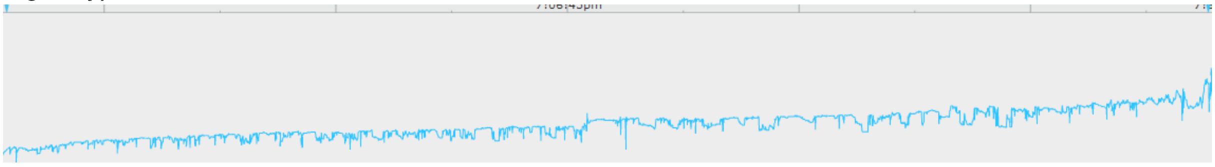
Session 6:



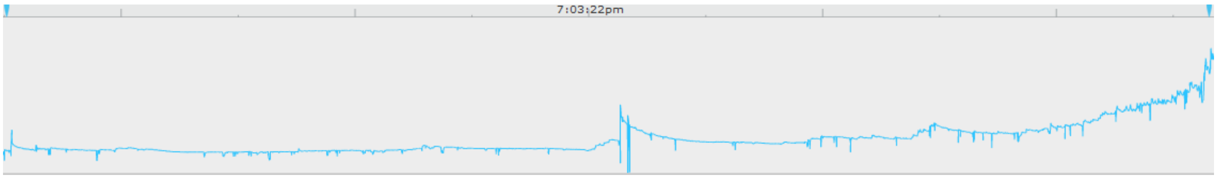
Session 7 (on the highway):



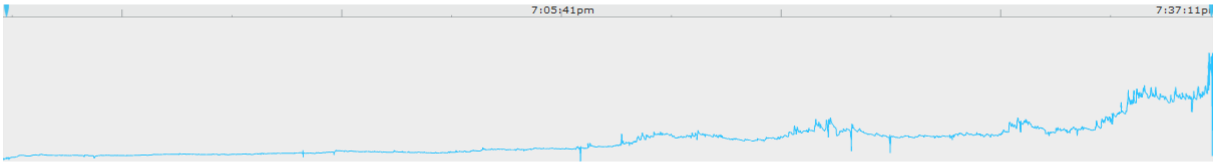
Session 8 (on the highway):



Session 9:

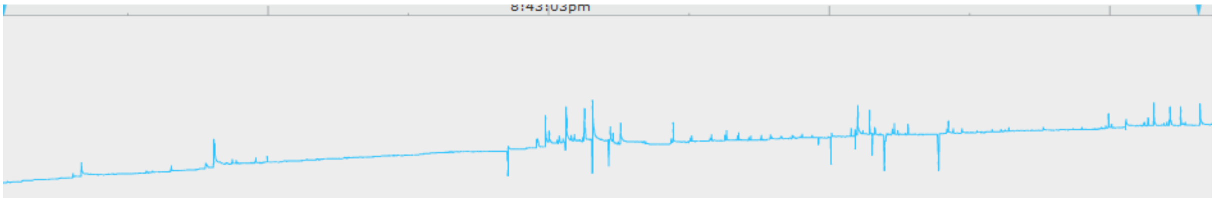


Session 10:

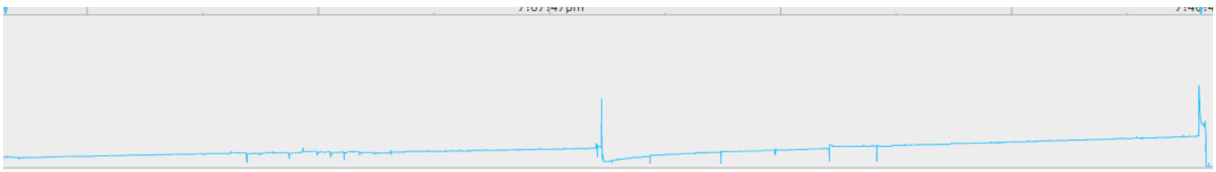


Participant 2

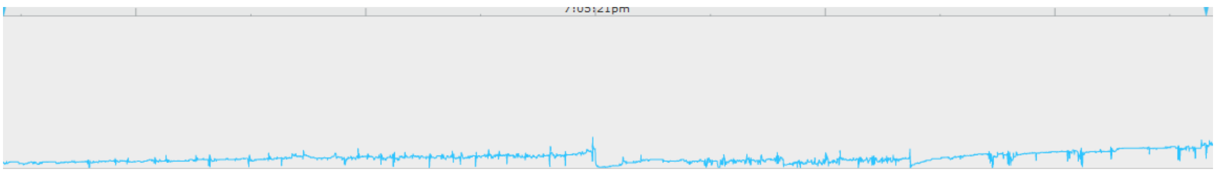
Session 1:



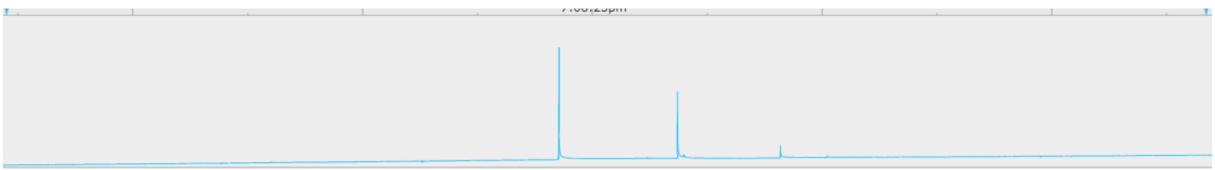
Session 4:



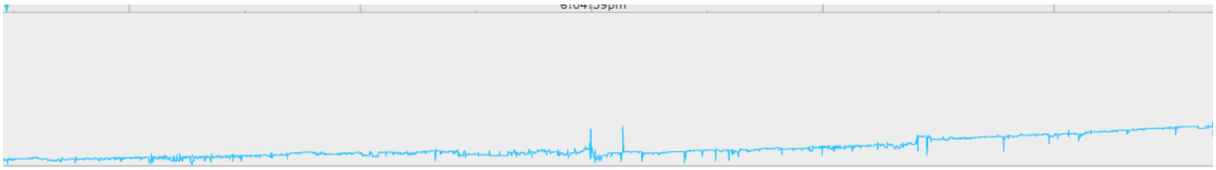
Session 5:



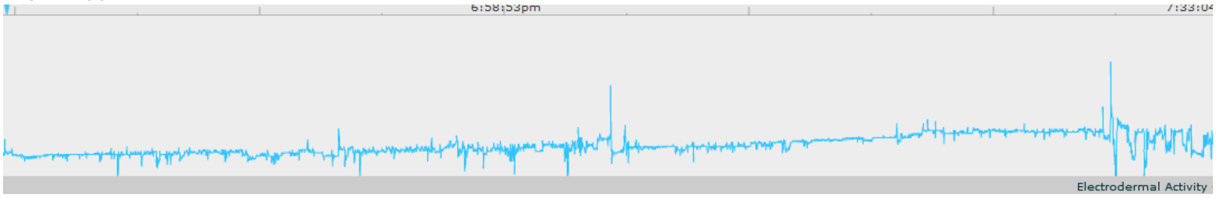
Session 6:



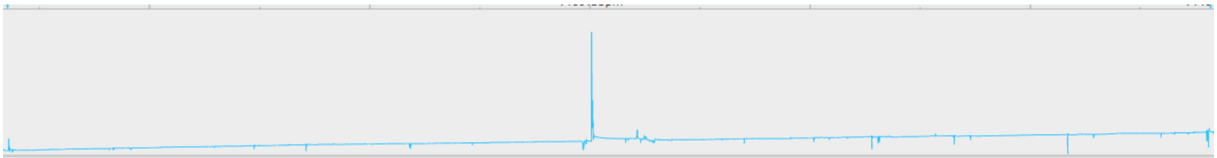
Session 7 (on the highway):



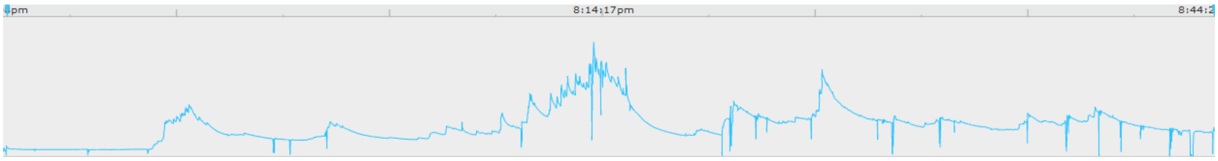
Session 8 (on the highway):



Session 9:

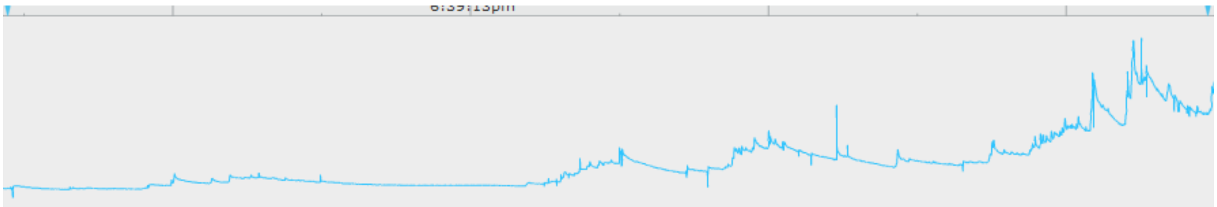


Session 10:

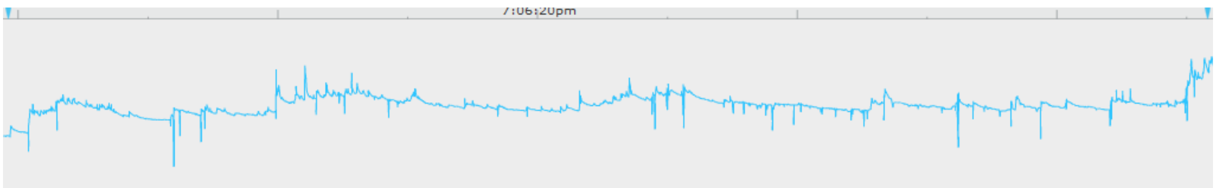


Participant 3

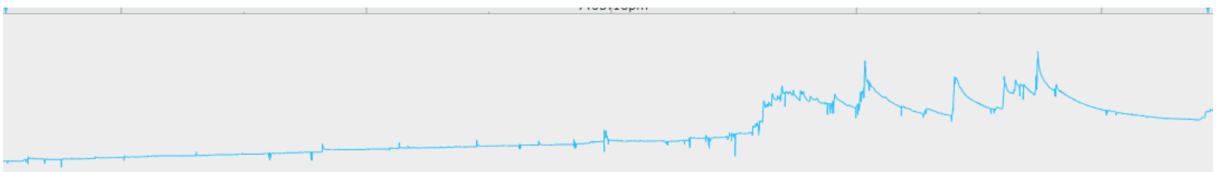
Session 1:



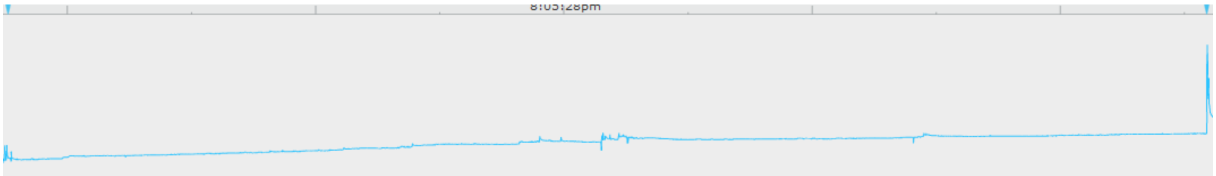
Session 2:



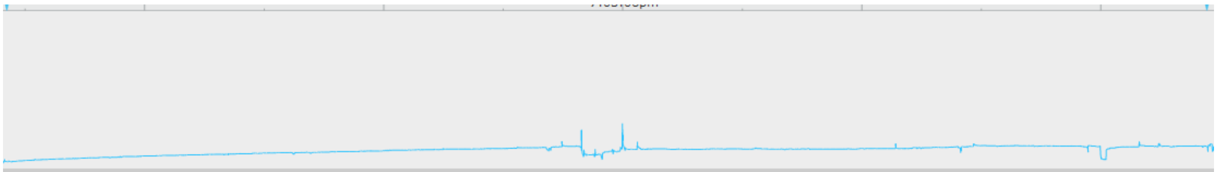
Session 3:



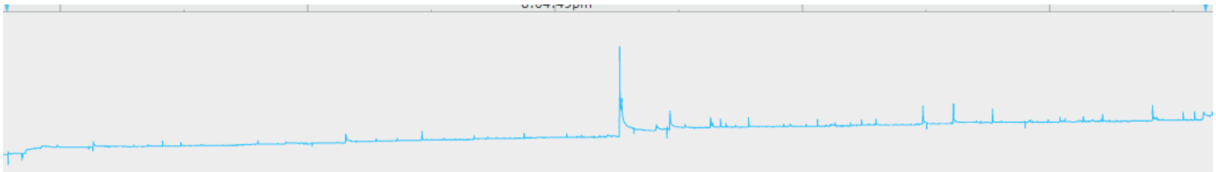
Session
4:



Session 5:

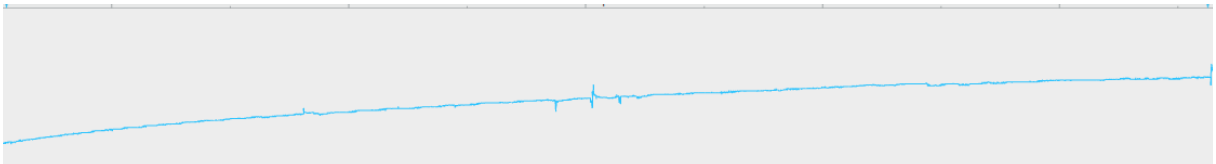


Session 6:

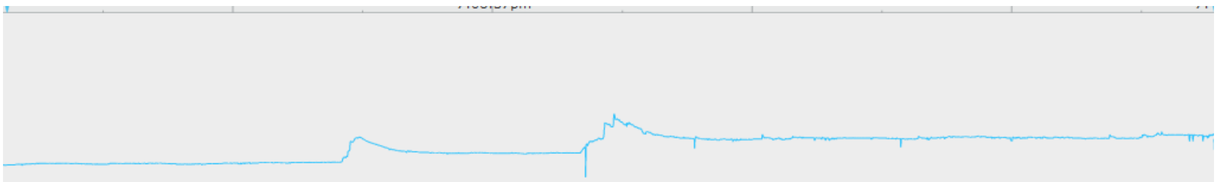


Session 7
(on the highway):

Session
8 (on the highway):



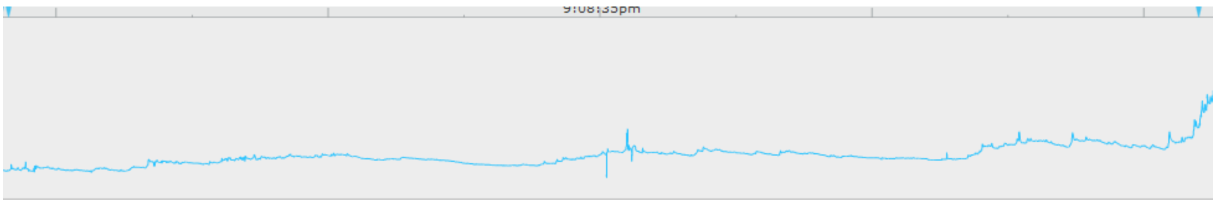
Session 9:



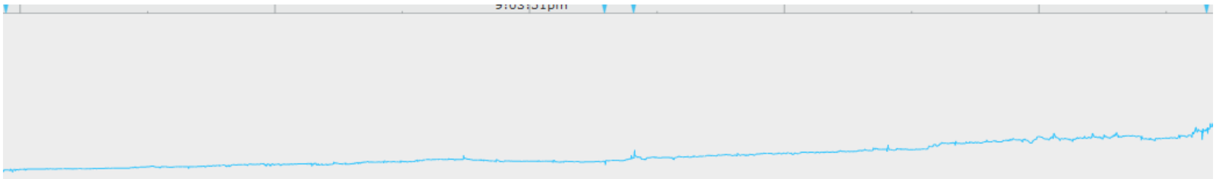
Session 10:

Participant 4

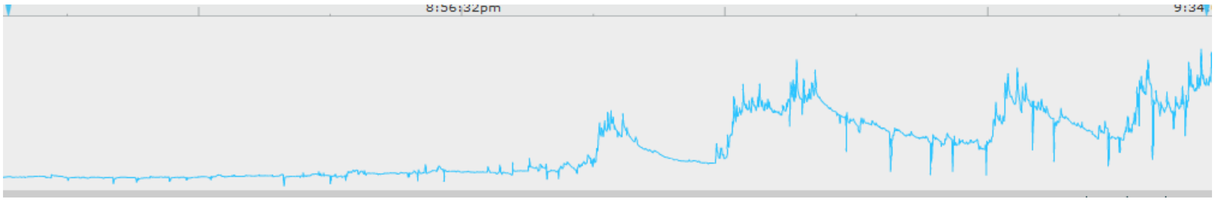
Session 1:



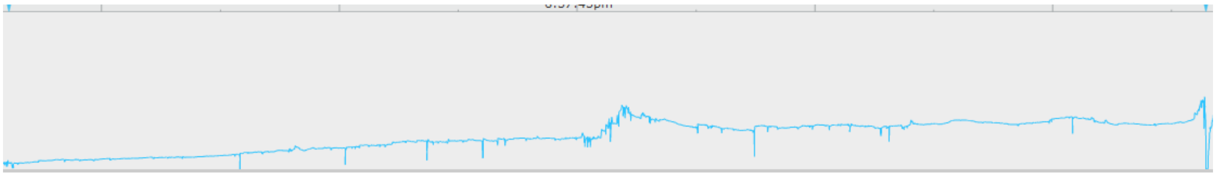
Session 2:



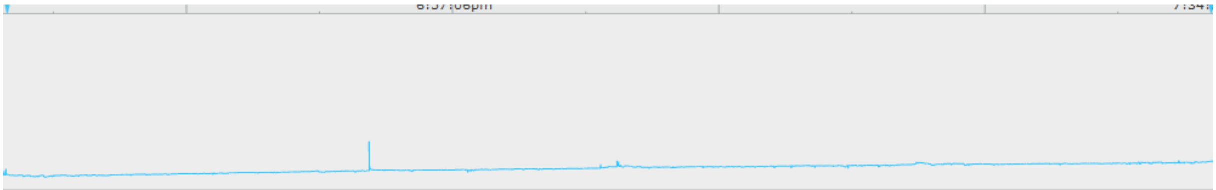
Session 3:



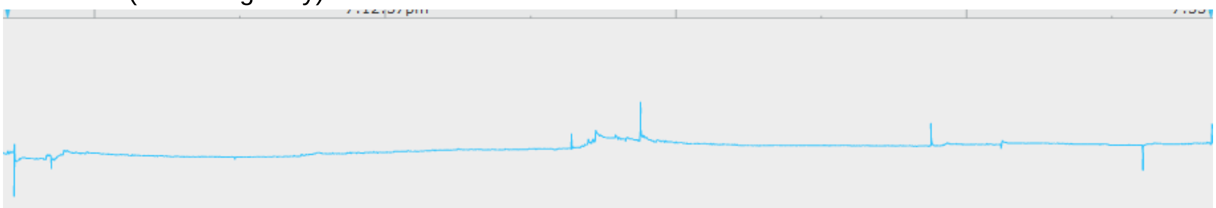
Session
5:



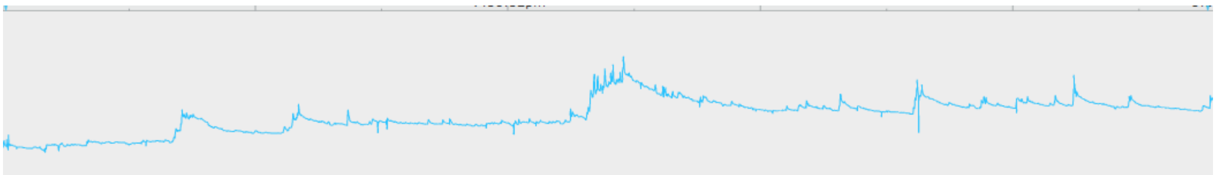
Session 6:



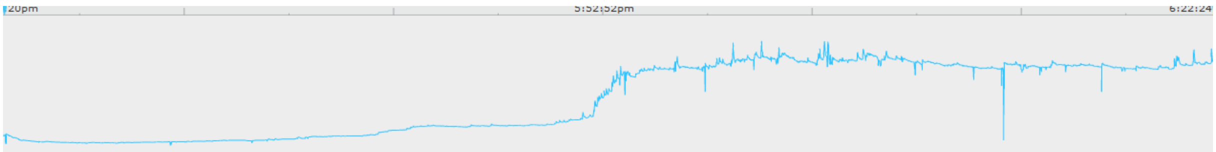
Session 8 (on the highway):



Session 9:

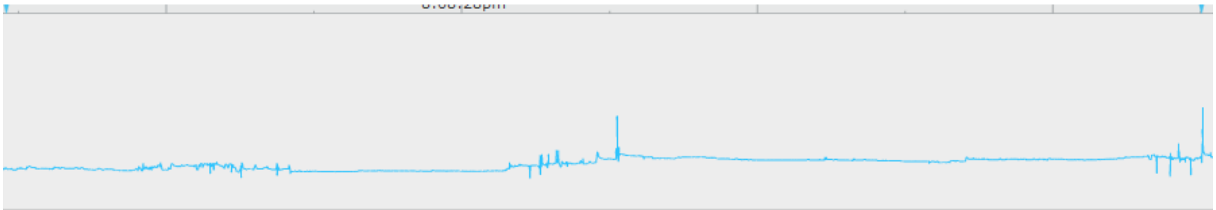


Session10:

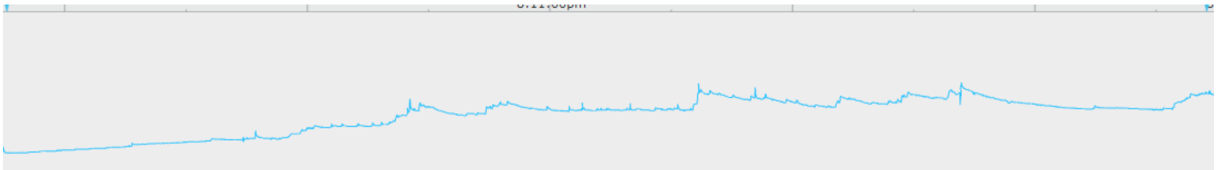


Participant 5

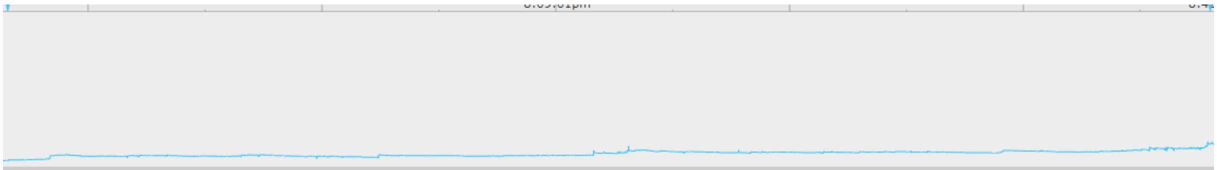
Session 1:



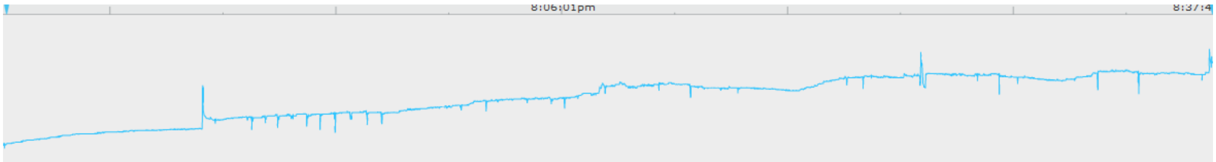
Session 2:



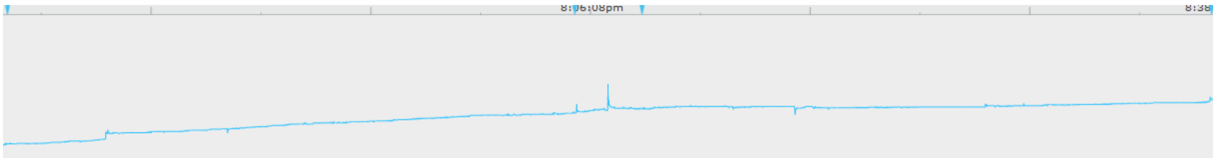
Session 3:



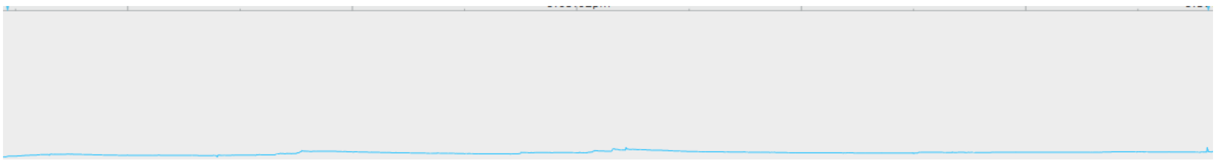
Session4:



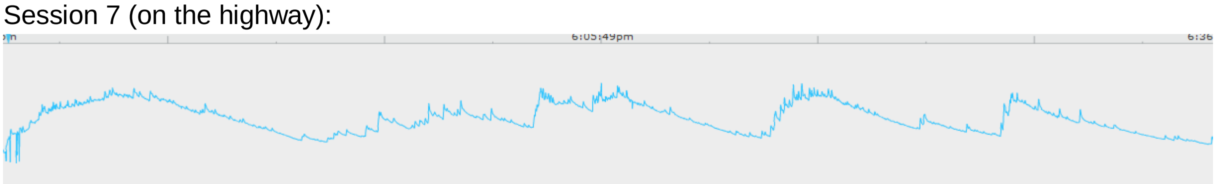
Session 5:



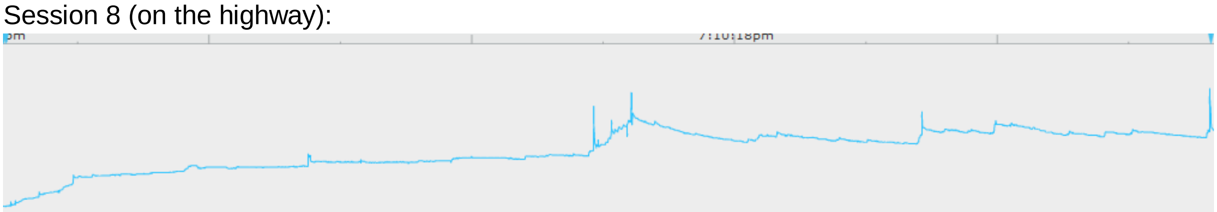
Session 6:



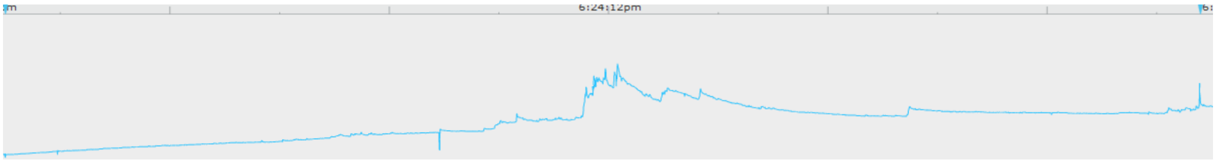
Session 7 (on the highway):



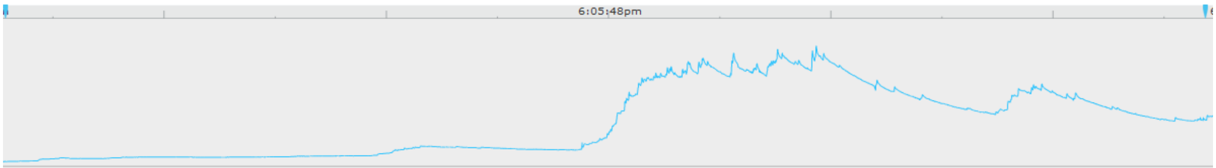
Session 8 (on the highway):



Session 9:

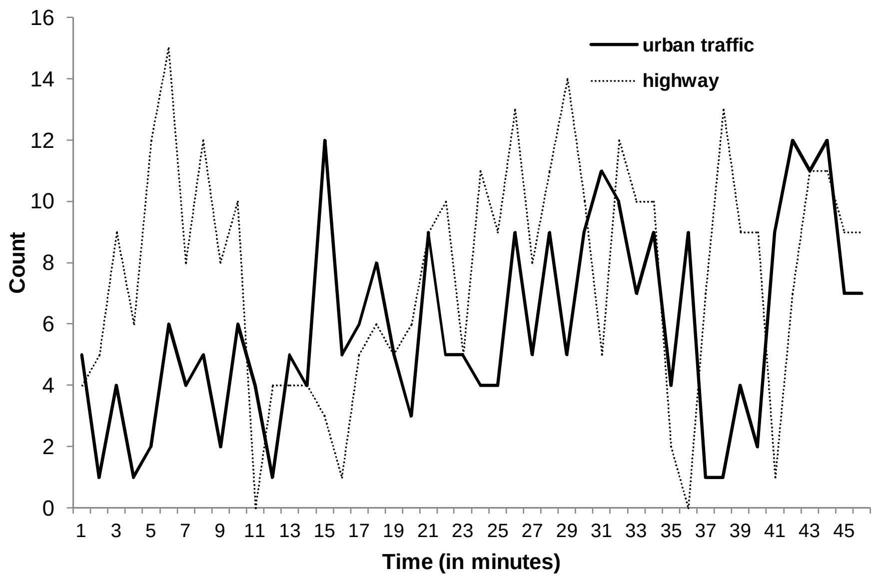


Session 10:

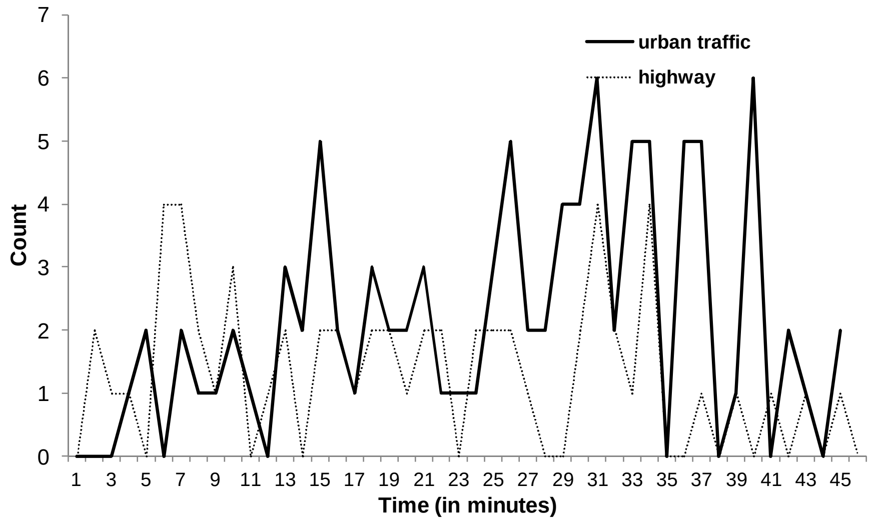


APPENDIX 4 - Compared NS.SCR's quantity of a highway and urban traffic session (Note: Different scalings on the vertical ass)

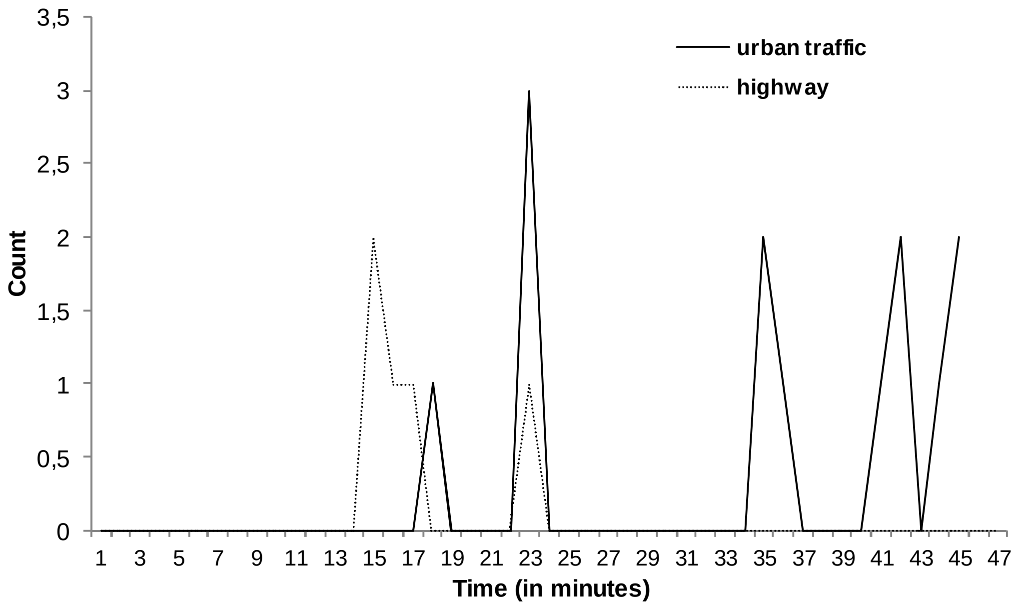
<u>Participant 1</u>



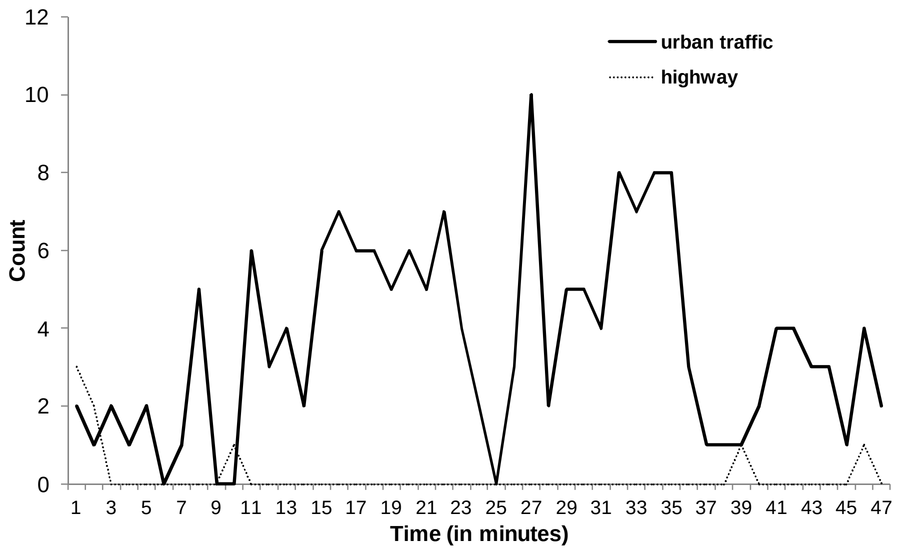
Participant 2



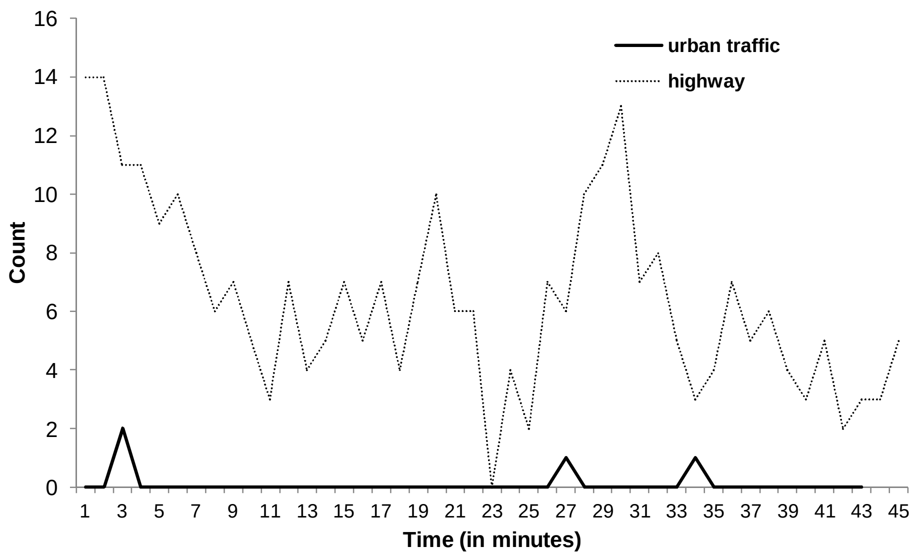
Participant 3



Participant 4

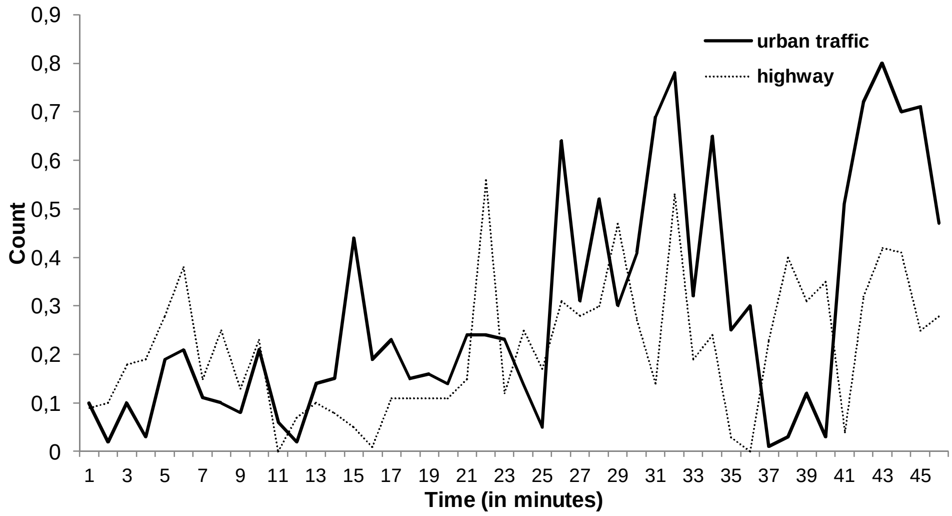


Participant 5

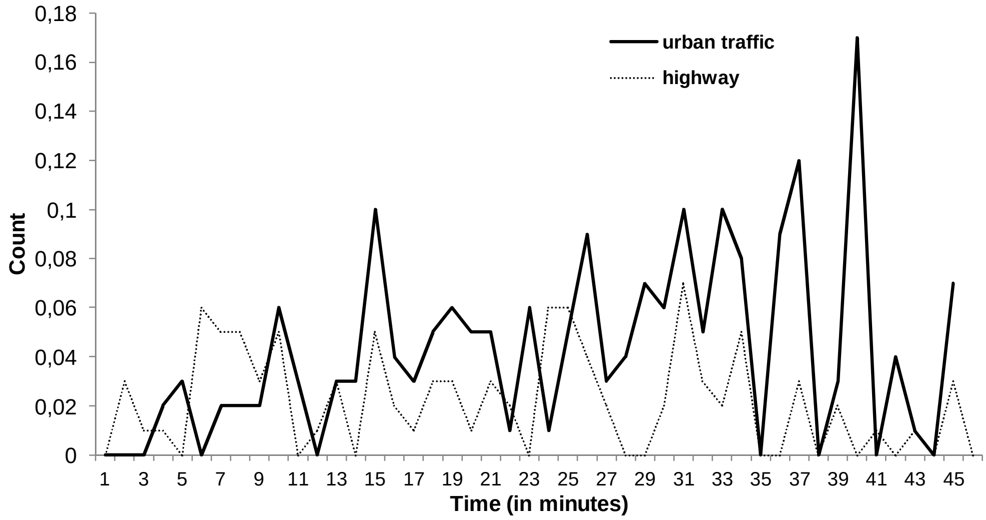


APPENDIX 5 - Compared NS.SCR's amplitude of a highway and urban traffic session.

<u>Participant 1</u>

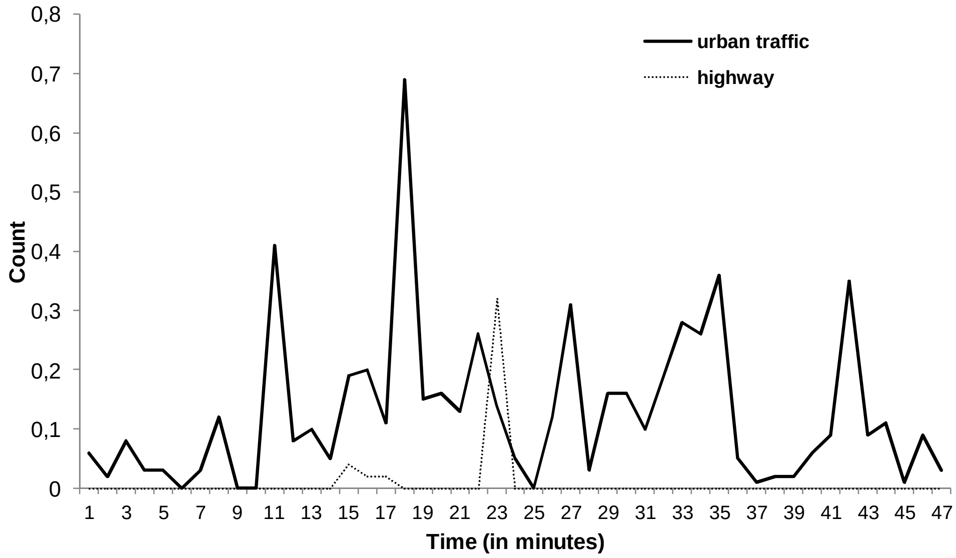


<u>Participant 2</u>

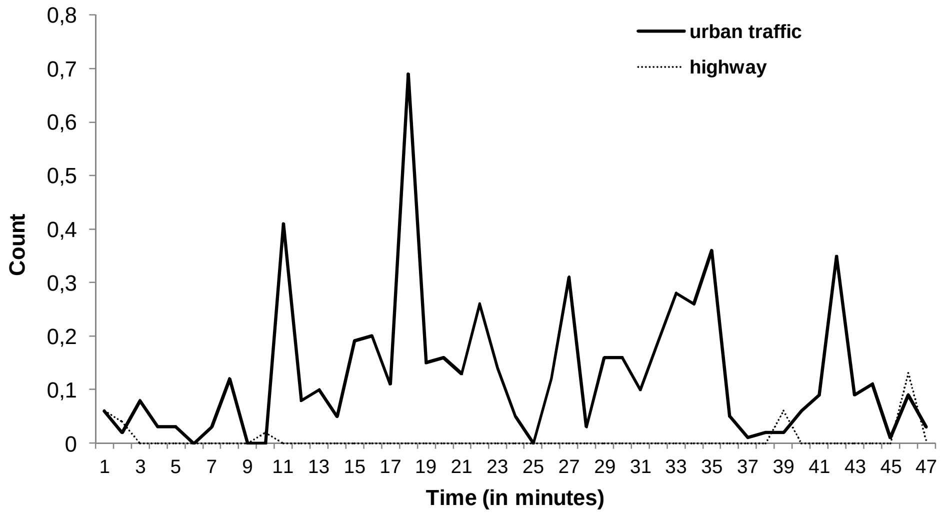


Note. Different scaling on the vertical ass

Participant 3



Participant 4



Participant 5