UNIVERSITY OF GOTHENBURG

DEPARTMENT OF APPLIED IT

INVESTIGATING THE VIABILITY OF PERFORMING DRIVING EXAMINATIONS IN A SIMULATED ENVIRONMENT

The effect of mere presence and evaluation apprehension on driving performance

Oskar Ekström, Joel Ruthberg, Erik Van Der Maarel

Thesis:	15 hp
Program:	Cognitive Science
Level:	First Cycle
Year:	2019
Supervisor:	Alexander Almér
Examiner:	Robert Lowe
Report nr:	2019:57

Abstract

This report aims to give insights into how simulators could be used to assess driving competence. For simulators to constitute viable alternatives to on-road assessments it is necessary to understand all meaningful differences. One differentiating factor between conditions could be the social presence of a driving examiner. The first part of this study therefore aims to investigate presence and evaluation effects from the perspective of social facilitation and impairment. This was done by letting 41 participants drive through a short simulated city scenario using a relatively simple single screen simulator setup at the Swedish National Road and Transport Research Institute. Using a between-groups design participants either drove alone or in the company of another non-observing individual (isolating the factor of mere social presence) and were told that their driving would or would not be assessed (isolating the factor of evaluation apprehension). To further investigate the viability of assessing certain tasks required for a Swedish class B driver's license using a simulator setup, perceived behavioral validity for tasks included in the simulated scenario were assessed by letting participants complete a two-part survey. The results indicated no significant effect of social facilitation and impairment on driving performance. Furthermore, the study indicates that performing and assessing certain elements of the Swedish class B driver test in a simulated environment are theoretically valid provided that the simulator is of sufficient fidelity. However more research is required before this can be considered practically viable.

Keywords

Driving Simulation, Driver Assessment, Swedish Driver Test, Social Facilitation, Social Impairment, Mere Presence, Evaluation Apprehension**.**

Sammanfattning

Denna rapport syftar till att ge insikter om hur simulatorer kan användas för att bedöma förarkompetens. För att simulatorer ska utgöra ett lämpligt alternativ till bedömningar på väg är det nödvändigt att förstå alla meningsfulla skillnader. En sådan skillnad kan utgöras av den sociala närvaron av en examinator. Den första delen av denna rapport syftar därför till att undersöka *mere presence* och *evaluation apprehension* utifrån perspektivet av social facilitering och social interferens. Detta gjordes genom att låta 41 deltagare köra genom ett kort simulerat stadsscenario i en relativt enkel körsimulator vid Väg- och transportforskningsinstitutet. I en mellangruppsdesign fick deltagarna antingen köra ensamma eller i närvaron av en annan icke-observerande person (för att isolera faktorn *mere presence*). Dessutom fick de instruktioner om att deras körning antingen skulle bedömas eller inte (för att isolera *evaluation apprehension*). För att vidare undersöka möjligheten att bedöma vissa aspekter av förarprovet i simulerad miljö så fick deltagarna slutföra en tvådelad enkätstudie. Resultaten indikerade ingen signifikant effekt av social facilitering och social interferens på körprestanda. Studien visar vidare att det är teoretiskt möjligt att utföra och bedöma vissa delar av det svenska förarprovet i en simulerad miljö, förutsatt att simuleringen är tillräckligt trovärdig. Ytterligare forskning krävs dock innan detta kan anses praktiskt genomförbart.

Nyckelord
Körsimulering, Förarprov, Körkortsprov, Social Facilitering, Social Interferens, Mere Presence, Evaluation Apprehension.

Foreword

This report constitutes a bachelor thesis written by cognitive science students at the University of Gothenburg and the Department of Applied Information Technology in collaboration with the Swedish National Road and Transport Research Institute in spring 2019. The authors would like to express gratitude to all those who participated in the study. A special thanks to Alexander Almér for supervising on behalf of the University of Gothenburg and Anders Lindström and Birgitta Thorslund on behalf of VTI.

Table of contents

1 Introduction

The use of simulators and virtual environments has been steadily increasing and continues to provide new alternative approaches for tackling issues in a wide range of fields and applications. The Swedish National Road and Transport Research Institute (VTI) conducts research on simulated vehicle environments and the question has been raised as to what extent simulators can be used to conduct official driving assessments. The aim of this report is to give further insight into the use of driving simulators and simulated environments as a tool for assessing driving competence.

The use of driving simulators to assess and evaluate driver competence can be both a cost- and time-efficient alternative to the currently used on-road assessments. Furthermore it may lower the step of entry for potentially competent novice drivers who are intimidated by the current process as many risks associated with on-road practice, examinations and driving can be eliminated. Another potential benefit of using simulators is the customizability; scenarios can be constructed to include a range of different hazardous events, variable conditions and alternative outcomes in a way that is not possible in on-road situations. Furthermore, identical settings can be assured between drivers and data can be objectively recorded. With these benefits in mind it is therefore of importance to investigate the potential use, limitations and applications of driving examinations in a simulated environment. This question is approached from two perspectives as described further in the following sections: Social presence effects on driving performance in a simulated environment, and an investigation into elements of the Swedish class B driving test through self-reported driver behavior prior to and after driving in the simulated environment.

1.1 Social Presence Effects and Driving Performance

For simulators to constitute viable options for assessing driver competence it is necessary to ensure that meaningful differences between on-road driving and simulated driving are well known (Campos et al., 2017). This is especially true if certain driver responses are assumed to be directly comparable between these two conditions. One aspect that could differ between on-road driving assessments and simulator-based driving assessments is the presence of another individual both in a social and evaluative context. If driving assessments are carried out in simulated environments and based on recorded measures of driving performance, the presence of an examiner could prove unnecessary since assessments could be made at a later time or at another location. It is however not yet fully understood if social presence in this particular context affects drivers performance. The aforementioned issue is approached with a quantitative study investigating the effects of social presence on driving performance (study one). Based on previous studies on the effects of social presence (for a review, see Guerin (1993)) it is hypothesized that both the mere presence of another individual and the evaluatory aspect will interact, affecting driving performance in a positive way for experienced drivers.

1.2 Driving Examinations and Assessing Driver Behavior

Driving examinations aim to distinguish safe from unsafe drivers (Filtness, Tones, Bates, Watson & Williamson, 2016) and have an important function in encouraging novice drivers to obtain a certain level of education and driving competency before attempting to acquire a driver's license (Bates, Filtness & Watson, 2018). However, practical examinations are far from perfect in assessing a novice driver’s competence. Filtness et al. (2016) cites studies that indicate a weak correlation between assessment scores and subsequent crash-risk; that newly licensed drivers appear to be the group with the highest crash-risk and that retest reliability is quite low (as many as 36% would receive a different outcome the second time they performed a practical drivers test in Great Britain). Least to say there is room for improvement, and simulators could be a possible way to increase effectiveness of evaluations. In the future it may be possible that certain procedural skills associated with unusual or hazardous events could be isolated for assessment in simulators (Filtness et al., 2016).

As it stands today a number of different procedural skills and elements are assessed in the Swedish driving test for a driver's license with eligibility B (Transportstyrelsen, 2012; Trafikverket, 2018). If simulators are to be used as a tool for assessing these skills and elements it is crucial that on-road driving behaviors are analogous with behaviors in the simulated setting (Mullen, Charlton, Devlin & Bédard, 2011; Campos et al., 2017). The simulated city scenario provided by VTI includes unprotected pedestrians, zebra crossings, signal controlled crossings, roadworks and effective braking. The second part of this report was therefore restricted to evaluate this subset of the Swedish driving test. A survey study was performed in an attempt to understand the viability of assessing these elements of the Swedish driving test in a simulated environment (study two). It has been shown that self-reported survey items can be used as a suitable alternative method of establishing validity of measures from a simulated scenario (Reimer, D'Ambrosio, Coughlin, Kafrissen & Biederman, 2006) and give indications of on-road driving behaviors (Ben-Ari, Hager & Prato, 2016). This study consists of participants' responses collected prior to and after driving in a simulated environment. The aim was to gather responses on self-reported driver behavior and the ability of the simulator to emulate the driving experience. These responses were used as a basis for further discussion.

2 Theory and Previous Research

2.1 Social Facilitation and Impairment

It is well known that social presence in and of itself has the potential to affect performance, and such effects can be grouped by the term *social facilitation and impairment* (SFI) (Belletier, Normand & Huguet, 2019). SFI effects can be defined as behavioral changes that occur due to the presence of a non-interacting conspecific (Guerin, 1993). The effects of SFI has been recognized for more than a century. In one of the first studies on the topic, Triplett (1898) found that cyclists performance improved when cycling together with a pacemaker as compared to when cycling alone. Allport (1920) later found that performance differed in group compared to solitary settings even when attempts were made to control for the factor of rivalry between subjects, something that was lacking from Tripletts study. Since the early studies on SFI it has been shown that the phenomenon is robust and widely spread, not only present in humans but also in other species such as cockroaches, birds and baboons (Zajonc, Heingartner, Herman & Mcguire, 1969; Huguet, Barbet, Belletier, Monteil, Fagot & Gauthier, 2014; Klopfer, 1958). The phenomenon has been found to have an effect on a multitude of activities ranging from eating (Herman, 2015) to performance on the Stroop task (Sharma, Brown, Booth & Huguet, 2010). Even though Triplett's early study indicated that performance was facilitated by social presence (Triplett, 1898); Allport (1920) and an abundance of later studies have shown that social presence can both impair and facilitate performance (Guerin & Innes, 1984; Zajonc 1965). These earlier studies were however unable to account for which factors that would lead to social impairment contra facilitation (Zajonc, 1965).

The contradictory findings that performance could be either impaired or facilitated was reconciled by Zajonc (1965) who proposed that drive theory (Spence, 1958) could account for the differentiating results of social presence studies. Zajonc hypothesized that the presence of others causes an increase in arousal, which in turn leads organisms to more often execute a dominant (habitual) response. Since responses to situations are thought to become habitual with learning, there is a higher probability that the dominant response is incorrect for novel or complex tasks but correct for easier or well-learned tasks. In short, Zajonc provided support for the notion that the presence of conspecifics has potential to impair performance on complex or novel tasks and facilitate performance on easier or well-learned tasks. However, even though task complexity is a common factor with which to explain SFI effects, individual differences in personality (negative or positive orientation to social presence) have been shown to potentially be a better predictor of performance. Individuals with a negative orientation tend to be negatively affected (impaired) by social presence whilst individuals with a positive orientation tend to be positively affected (facilitated) (Uziel, 2007).

2.1.1 Mere presence and evaluation apprehension

Since Zajonc's initial proposal of drive theory as the mechanism for SFI, the source of the increased arousal has been disputed (Guerin & Innes, 1984; Guerin, 1993). Is it perhaps the mere presence of conspecifics that leads to an increase in arousal or is it caused by a fear of being evaluated? The best answer seems to be that multiple factors could contribute to an increase in arousal and thereby independently lead to SFI effects (Guerin, 1986). Zajonc (as cited in Guerin, 1993) proposed that uncertainty follows from the mere presence of others, and that this uncertainty should increase an organism's level of arousal as it prepares to react. The more threatened an organism feels by the presence of another the larger the effect will be on it's behavior. Refuting this, Cottrell, Sekerak, Wack & Rittle (1968) showed that the presence of an audience actively observing facilitated the emission of dominant responses in a pseudorecognition task, while the presence of others unable to observe did not. This led Cottrell (as cited in Seitchik, Brown & Harkins, 2017) to propose the evaluation apprehension model of SFI. This model states that performance will only be affected when the conspecific is able to observe and evaluate the task being carried out. The model also states that performance will be unaffected by the mere presence of another conspecific unable to observe and evaluate performance; contrasting with Zajonc's proposal of the mere presence effect.

However, Guerin's (1993) review of 313 social presence studies disputes the previous claim that only evaluation apprehension should lead to SFI effects. In his review Guerin notes that a considerable amount of previous research in some regard has failed to successfully isolate the aforementioned effects of social presence. Therefore he defined 12 selective criteria; five for SFI, and seven to further isolate the effects of mere presence. Guerin found that eighteen of the reviewed studies met all criterion for mere presence, and that eleven of those remaining studies had produced significant results. However, effects were mainly found when the subjects were unable to observe the passive other. This led Guerin to propose that mere presence effects occur when participants are unable to monitor the other person present and that this uncertainty causes the effect. Harkins (1987) studies of social presence effects on an ideation task and a vigilance task also seemingly supports the view that both evaluation apprehension and mere presence are factors that independently are capable of affecting performance. It is however worth noting, that the presence of others could have both arousal increasing and arousal decreasing effects depending on the type of situation as well as the role of the person present (Mullen, Bryant & Driskell, 1997), it makes intuitive sense that not all social encounters should elicit fear and uncertainty.

2.1.2 Explanatory models of SFI effects

Guerin (1993) highlights that the Zajonc (1965) drive-based account leaves little room for motivation or other psychological processes and is vague in its definitions of drive and arousal. It is therefore important to note that there are more recent theories that ascribe SFI effects to behavioral demands (e.g. Bond, 1982), to changes in cognitive processing and attentional load (e.g. Muller & Butera, 2007; Belletier et al., 2019), and that the question of which explanatory model that best accounts for the underlying mechanisms of SFI effects appears to be open (Seitchik et al., 2017). The mere effort hypothesis (Harkins, 2006), which states that the anticipation of evaluation leads organisms to expend greater effort on tasks thereby facilitating performance on easy tasks while impairing performance on more complex tasks, could for instance also account for evaluation apprehension effects. Harkins (2006) has found supporting evidence for the hypothesis, showing that participants that are being evaluated expend greater effort on a remote associates task (where participants have to generate a fourth word related to a set of three previously viewed words), and that performance on this task varied with the complexity of the set. Attentional conflict between a focal task and a social stimulus could, as noted by Belletier et al., (2019), also (and without invalidating Zajonc's (1965) hypothesis) result in facilitated or impaired performance depending on task complexity. Performance on simple tasks or tasks that do not require much attention could be facilitated, as a perceived social threat would restrict attentional focus, while performance on complex or more attentionally demanding tasks could be impaired since less attentional resources are available for the focal task. The presence of conspecifics who are perceived as more threatening could then both increase arousal and simultaneously consume cognitive resources needed for top-down suppressions of habituated responses.

The underlying cognitive mechanisms are however less relevant for the present study in comparison to the clear findings of effects mediated by task complexity on performance ascribed to both evaluation apprehension (Cottrell et al., 1968; Harkins, 1987) and the mere presence of other individuals (Zajonc, 1965; Guerin & Innes, 1984; Guerin, 1986). In the present study a prerequisite of holding a valid drivers license should presumably make the simulated driving scenario a simple task for participants. The reasoning behind this presumption being that appropriate responses (e.g. braking or slowing down) should already be learned (internalised as dominant responses) for participants with driving experience. Following this line of argument performance should therefore be facilitated. These assumptions however hinges on the fact that participant behavior transfers adequately between on-road driving and the simulator setup.

2.1.3 Social presence effects and on-road driving

SFI and the effects this has on driving behavior is of importance to study since differences in performance due do the presence or non-presence of conspecifics might have serious consequences. One might think that the presence of passengers would consistently increase the risk of accidents due to attention being diverted to accommodate other persons present in the vehicle. However, studies have shown that in non-driving contexts distraction does not necessarily have to impair performance, it can in some cases actually facilitate performance on simple tasks (Baron & Glenn, 1975). If it is the case that distraction leads to an increase in arousal this seems to support an attentional or drive-based account of SFI effects. Investigating the effect of presence of passengers on driving performance, Vollrath, Meilinger & Krüger (2002) found a correlation between passengers and a reduced accident risk. This protective effect was however lessened if a large amount of attention had to be directed towards the passengers. It has also been found that drivers with passengers had their speeding reduced (Baxter, Manstead, Stradling, Campbell, Reason & Parker, 1990), drivers accompanied by passengers drove at slower speeds compared to drivers who were alone in the vehicle (Lawshe, 1940), and that drivers in the presence of passengers also were more likely to come to a complete stop at stop signals (Feest, 1968). In a more recent study Lee & Abdel-Aty (2008) found a strong correlation between the presence of passengers and safer driving behavior and that when accidents did happen they were less serious if the driver was accompanied by passengers. They also note that the protective effects of the presence of passengers increased with the number of passengers present. However, the presence of passengers is not always protective since the results from this study also indicate that younger drivers had an increased risk of accident in the presence of young passengers. In addition to these findings, it has been shown that success on a driving test was significantly higher when the individual being examined only was accompanied by the examiner as opposed to when two or more students were in the same car during the examination (Rosenbloom, Shahar, Perlman, Estreich & Kirzner, 2007). However, few have studied the effects of a present examiner in the context of a driving simulator.

Worth noting is that none of the studies mentioned above seem to have isolated different mediating mechanisms of SFI (e.g. mere presence or evaluation apprehension). Following the criteria proposed by Guerin (1993) one finds that the studies do not control for interaction between driver and passenger and it is therefore unclear if they can claim to have fully isolated SFI from other psychological phenomena.

2.2 Simulated Driving Assessments

2.2.1 Benefits

Earlier research regarding the viability of using simulators for driving training has shown that practicing driving in a simulator significantly improved the driving performance of elderly people (Casutt, Theill, Keller & Jäncke, 2014). In addition to this, beneficial effects from practicing in a driving simulator can last for up to 18 months (Roenker, Cissell, Ball, Wadley & Edwards, 2003). The findings that simulator training translates to real life driving and that the beneficial effects of the training can last for several months indicate that simulator-based assessments are theoretically possible. A general rule of thumb for simulator-based driver's assessments proposed by Kappé, de Penning, Marsman & Roelofs (2012) is that: "tasks that can be trained well can also be assessed well".

Driving simulators have the potential to provide a safe way of assessing driving competence in a controlled and standardized manner. Simulators allow for collecting objective driver-response data, and allow scenarios to be customized thereby providing a greater flexibility compared to on-road driving. In simulator-based assessments time spent on unproductive (unassessed) driving could be minimized. Filtness et al., (2016) further note that simulators could prove beneficial if they can be used to expose drivers to an increased range of environments, including environments not likely to appear in on-road assessments. Performance during extreme weather conditions or reactions to different hazardous situations could for instance be assessed as they could safely be included in driving scenarios. Simulated driving assessments using objective data could also reduce the need for a potentially biased examiner, who runs a risk of having their judgments affected by contextual aspects or personal characteristics. It could also eliminate inter-examiner differences in interpretations of competency requirements. If the evaluator could be excluded from the situation ecological validity may also be improved as subjects are left to drive on their own, and there exists positive indications that recordings of simulated driving sessions could be sufficiently assessed by an evaluator at a later time (assessments based on off-road crashes, collisions, road edge excursions, failures to stop at stop signs, center-line crossings, illegal turns and driving above the speed limit) (Bédard, Parkkari, Weaver, Riendeau & Dahlquist, 2010). Of further relevance for using simulators as an evaluative tool it is worth to note that simulated driving performance has been shown to correlate with on-road performance in younger drivers preparing for examinations (de Winter et al, 2009), as well as older drivers (Lee, Cameron & Lee, 2003).

2.2.2 Considerations

Simulated environments are to date unable to fully emulate the on-road driving experience and simulated driving scenarios are unable to account for all events that can possibly occur in natural settings (Filtness et al., 2016). For driving simulators to be considered viable evaluation tools, it is crucial that they sufficiently reproduce the on-road driving experience (are of sufficient fidelity) and are valid in predicting on-road behavior (Mullen et al., 2011; Campos et al., 2017). The ability to predict on-road driving behavior from a person's behavior in the simulated environment is referred to as behavioral validity. There are a range of concerns regarding how fidelity and behavioral validity varies between different driving simulator setups. For instance, it has been found that participants driving in a simulated environment could be less emotionally engaged than participants driving in real life (Ekanayake, Backlund, Ziemke, Ramberg, Hewagamage & Lebram, 2013). If simulators are lacking in fidelity participants may approach situations differently and the perceived absence of an actual crash risk may impact driving behavior, in turn affecting the behavioral validity of the simulator (Filtness et al., 2016). Further, the hardware configuration of the simulator setup needs to be sufficiently similar to a real car if procedural aspects associated with real-world driving are to be assessed (Kappé et al., 2012). Regarding how behaviors translate between conditions, a common distinction is made between absolute and relative behavioral validity (Törnros, 1998; Mullen et al., 2011). Absolute validity refers to an exact correspondence of numerical behavioral data, while relative validity refers to correspondence of directionality and a similar magnitude of effect between on-road and simulated driving conditions. Simulated driving is generally believed to be able to approximate but not replicate driving behavior.

Relative validity has been established across several studies for a range of different behavioral measures and is considered sufficient for most purposes (Törnros, 1998; Mullen et al., 2011). Behavioral measures that have been found to correlate between simulated and on-road driving conditions include: speed, lateral position, risky traffic behaviors, breaking responses, reaction times, risk of future traffic violations for older drivers and age related changes in driving performance (Mullen et al., 2011). Further, with drive- and attentional-based accounts of SFI effects in mind, physiological measures as well as effects of divided attention on driving performance have been shown to correlate between simulated and on-road conditions (Mullen et al., 2011). It is however stressed that simulator validity always should be established for conditions and groups measured; validity should never be implicitly assumed since behaviors could be highly task-dependent and affected by individual differences (Mullen et al., 2011; Campos et al., 2017). A particular skill may be successfully evaluated using a particular setup, but unsuccessfully evaluated on another, and all meaningful differences between setups may not yet be fully understood (Filtness et al., 2016).

2.2.3 Current use

Following the previously stated considerations, it is perhaps unsurprising that simulators are not yet widely used for licensing purposes and that no such uses for novice drivers could be found. However, in Ecuador it has been reported that experienced drivers can choose to complete their assessment for a truck driver's license in a simulator instead of being assessed on a circuit closed for traffic (Filtness et al., 2016).

It has been demonstrated that simulators can be used in order to distinguish between experienced and novice drivers (Damm, Nachtergaële, Meskali & Berthelon, 2011), but further evidence is thought to be needed in order to confidently determine if a novice driver has acquired sufficient knowledge (Filtness et al., 2016). Even if state-of-the-art simulators currently are insufficient in fully establishing driver competence, it's highlighted that assessment in driving simulators could serve as an intermediate step in order to pre-screen drivers for risky behaviors, reducing but not eliminating the need for resource-heavy on-road assessments (Campos et al., 2017). Simulators could be used to assess reactions to simulated accident scenarios (Damm et al., 2011), and to conduct evaluations of attention maintenance and hazard perceptions (Filtness et al., 2016), thereby aiming to lower the crash risk for newly licensed drivers.

3 Study 1: The Effect of Mere Presence and Evaluation Apprehension on Driving Performance

The aim of study one was to investigate if driving performance in a simulated environment would be affected by evaluative status and the mere presence of another individual. Based on the previously reviewed literature on: SFI, the effects of mere presence, and evaluation apprehension, it was hypothesized that the interaction between evaluation and the presence of another individual would affect the selected measures of driving performance.

3.1 Method

3.1.1 Participants

41 (23 male and 18 female) participants were recruited ($N = 41$). The participants were between the ages of 21 and 44, the only prerequisite was that the participants currently possessed a valid driver's license. On average the participants had held a valid license for approximately 8 years ($M = 7.61$). To estimate driving experience participants were asked how often they drive: 43.9% reported to drive at least once per year, 26.8% at least once per month, 24.4% at least once per week and 2 individuals reported very infrequent driving. The majority of the participants were students from the University of Gothenburg. All participants were rewarded with a cinema voucher for their participation. A consent form was gathered for all participants informing them of the purpose of the study, what the data collected would be used for and that the individual identity of the participants would remain anonymous (see Appendix 1). The participants were also informed that they could choose to abort the study and withdraw their data without consequence at any point during the study. The consent form was presented to all participants before both study one and study two. The participants were debriefed after the study was completed, no participants aborted the study or chose to withdraw their results.

3.1.2 Equipment

A Samsung SyncMaster 400MX-2 LCD television screen (40", 1080p Full HD) was used to present the simulated scenario at a distance of approximately 1.5 meters from the participant. The peripherals consisted of a Logitech G27 Racing Wheel with dual-motor force feedback and accompanying pedals of the same brand, together with a Corbeau gaming chair. The vehicle audio was played through a Logitech THX speaker (one subwoofer and two stereo speakers) placed behind the chair. The setup is shown in Figure 1.

The simulator software was provided by VTI. It is based on Open Source and in-house developed code adhering to OpenDRIVE® for the description of the logical road network; and run using the in-house developed graphical image generator VISIR based on OpenSceneGraph (Swedish National Road and Transport Research Institute).

Figure 1. Simulator setup.

3.1.3 Design

A between groups design was used in order to separate the two independent variables: evaluative status (evaluation or no evaluation), and mere presence (presence, no presence). These variables combine to create the four groups: (1) No presence, No evaluation; (2) No presence, Evaluation; (3) Presence, No evaluation; (4) Presence, Evaluation. The conditions were ordered by block randomization and participants were randomly assigned to each block.

In order to compare the four groups each of the participants drove an identical scenario set in a simulated city environment. A total of 24 events occurred during the 13 minute drive ranging from pedestrian crossings from both sides of the street; both expected and unexpected, parked cars pulling out in front of the driver, passing of buses at bus stops with a range of variations (bus letting driver pass, bus pulling out in front, child running into the street in front, soccer ball bouncing out across the street in front of the bus). From these events four different measurements were recorded: driver reaction times (ms), time headway (s), max retardation (m/s^2) and mean speed (m/s).

Driver reaction times were measured from the visual presentation of certain sudden events to the initialisation of the breaks by the participants. The mean reaction time to unexpected hazards in real-world driving situations (time to detect hazard and initiate action) have been found to be about 1.1 seconds, with most drivers being able to react in 2 seconds or less (Fambro, Koppa, Picha & Fitzpatrick, 1998). Reaction times exceeding 2 seconds are therefore considered outliers and excluded from the analysis. Furthermore, a lower boundary was set at 100 ms as faster reaction times were deemed impossible due to the setup and sensitivity of the equipment.

Time Headway (THW) was measured as the amount of time required for the front axle of the participants' car to travel the distance needed to reach the same position as the object in front. It has been shown that a short THW can be interpreted as an indicator for a potentially dangerous situation (Vogel, 2003). THW together with the maximum retardation measures give an indication as to how the participants approached the different events.

Maximum retardation during these events was measured as the maximum deceleration achieved approaching the object in front, and mean speed across two shorter sections of the scenario were also recorded. The above described four measurements were chosen in part due to the setup of the provided software, however as mentioned by Mullen et al., (2011), some have also been correlated as transferable behavioral measurements between simulated and on-road driving.

All participants were presented with the same written instructions in regards to how the simulated car worked (automatic gearbox, location of the rearview mirror and where the speedometer was located). They were also instructed to follow the main road straight ahead and to drive as they would if it were a real-world situation. When the scenario was completed all participants were instructed to retrieve the experiment leader from an office located on a different floor of the building approximately 30 meters away from the room containing the simulator setup. In order to evoke a sense of evaluation within the participants of the evaluation apprehension groups, different instructions were presented to groups one and three, and groups two and four (see Appendix 2). Groups one and three were presented with instructions simply stating that their performance would not be evaluated at any point during the study, whereas groups two and four were presented with additional instructions more focused on the participants and their driving performance. Furthermore, these specific instructions included information stating that their driving performance would be objectively measured and immediately compared against the current criteria for passing a swedish driving examination, this in order to determine if the participant was proficient enough to pass a driving test today.

It was decided that the participants in the groups without evaluative instructions were not to be given a secondary task or reason as to why they were participating in the study until after the entirety of the drive and subsequent survey was completed. This was done since it has been argued that giving participants a reason as to why they are performing a given task introduces a secondary confounding task (Klauer, Herfordt, & Voss, 2008). This secondary task might for example take the form of impression formation if the reason given for participating would be to evaluate the task performed. The mere presence condition was set up in accordance to the 12 criteria established by Guerin (1993) relating to the isolation of SFI and mere presence conditions. In this case the person included in the mere presence condition was facing away from both the simulator screen and participants and did not interact with the subject during the entirety of the drive. The positioning of the mere presence individual in relation to the simulator is shown in Figure 2.

Figure 2. Positioning of mere presence individual.

3.1.4 Procedure

The participants received approximately five minutes of training on rural and urban roads with little traffic and without any simulated events taking place. They then received the instructions (Appendix 2) as to whether or not they would be judged on their performance. When the instructions had been presented the participants were allowed to ask questions if they had any. Following this, the simulated city scenario lasting for approximately 13 minutes was initiated. For the mere presence conditions another person remained present in the room working on a different unrelated task. For all of the conditions the experiment leader was present in the room during the training scenario in event of any questions but always absent during the main scenario. After the driving scenario was completed the participant would come and collect the experiment leader in order to finish the contribution to study 1 and to continue with the survey for study 2.

3.2 Results

Out of the 24 total events THW was measured in 18 events. The measurements were compiled and the mean THW (s) was calculated for each participant (N = 41). The same was done for reaction times (ms) measured in 12 events. Seven outliers were excluded from the response time analysis (6 > 2000ms and 1 < 100ms). Max retardation (m/s^2) from 16 events and mean speed (m/s) across two longer sections of the simulated road were compiled in the same way (see Appendix 3 for a summary of the descriptive statistics). Levene's test for homogeneity of variance was not significant for any variable. Shapiro-Wilks test for normality of distribution was met for reaction times and max retardation. However, in group 2 (no presence, evaluation) mean speed measures were not normally distributed ($p = .024$). Nor was THW measures in group 3 (presence, no evaluation) ($p = .002$), these measures were however normally distributed in the remaining three groups.

In order to compare the four levels of the independent variables (no presence/no evaluation, no presence/evaluation, presence/no evaluation, presence/evaluation) on each of the four dependant variables; 4 two-way ANOVA were performed. The comparisons showed no statistically significant interaction between mere presence and evaluation on any of the four dependant variables. THW = ($F(1,37) = 0.173$, $p = .680$, $\eta_p^2 = .005$), reaction time = ($F(1,37) = 0.446$, $p = .508$, $\eta_p^2 = .012$), max retardation = ($F(1,37) = 0.405$, $p = .529$, = .011), mean speed = ($F(1,37) = 0.00004$, $p = .995$, $\eta_p^2 < .000001$). Furthermore, no significant main effects of mere presence or evaluative status were found on any of the four dependant variables. These results are summarized in Appendix 4.

3.3 Discussion

The results found do not support the hypothesis that mere presence and evaluation apprehension together would affect driving performance as measured by THW, reaction times, max retardation and mean speed. The measures recorded did not show any significant interaction effects or main effects of mere presence or evaluation apprehension. This study therefore indicates that the effects of social presence could be negligible when driving in a simulator.

As stated above the normality of distribution was not met in groups 2 and 3 for mean speed and THW respectively. However, factorial ANOVA can be considered somewhat robust to this issue of normal distribution in larger sample sizes (>30 or 40) as stated by Ghasemi & Zahediasl (2012). It is therefore assumed that the lack of a normal distribution found specifically in groups 2 and 3 for these measurements did not have any adverse effects on the final analysis. Another comment in regards to the data is that due to the way in which the simulator and software used for registering the measurements was set up, 208 of 1908 (11%) potential measures were not recorded. This loss of data was however spread relatively evenly across conditions and variables indicating no systematic issues, once again leading to the conclusion that this did not have any adverse effects on the final analysis.

The absence of significant interactions and main effects between evaluation and mere presence may very well be a result of the sample size as the effect sizes found in the analysis strongly indicate that any potentially significant effects would require vastly greater sample sizes. This is however a possible indication that the potential effect of mere presence and evaluation apprehension on the practical application of assessments in simulated environments may be of relatively low importance as discussed further in the following sections.

Another potentially limiting factor is mentioned in Guerin (1986), where it is discussed that experimental situations are inherently evaluative. It could be possible that the participants inferred that the mere presence individual filled some form of evaluative roll even though it was specified that no evaluation of their performance would take place. This is however merely speculative since it is difficult to know what expectations each participant brings to the experimental environment.

As noted by Uziel (2007) an orientation towards social presence is potentially a better predictor of the SFI effects; individuals with a positive orientation tend to have their performance facilitated by the presence of conspecifics and individuals with a negative orientation tend to have their performance impaired. The present study did not account for individual differences in social orientation and this could therefore constitute a limiting factor. Furthermore, explanatory models taking norms into account might be better suited to explain the SFI effects. Different individuals may react differently depending on the social situation, relationship with the other person present and the norms associated with this (Mullen, Bryant & Driskell, 1997; Baxter et al., 1990). In relation to this, Baxter et al. (1990) speculate that their findings of variation in driving performance with passenger age could be due to these norms. Even though Lee & Abdel-Aty (2008) do not extensively discuss the possible causes of their findings; it seems that explanations taking social norms into account also would explain their finding that young drivers accompanied by young passengers have a higher risk of accident than other groups. It would therefore be of interest for future studies into this topic to account for individual differences causing the mere presence effect or evaluation apprehension as well as to consider the perspective of social orientation.

The assumption that driving in a simulator is a well-learned task for participants with a driver's license and that their performance should be facilitated might constitute another limitation with the study. It is possible that participants who drive less frequently still perceive the driving task as complex and therefore experience impairment of their performance, thereby causing bidirectional and non-noticeable effects of social presence. Future studies could consider replacing the driver's license prerequisite with a more stringent frequency of driving criteria to see if this produces different results. It could also be of further interest to investigate if social presence would affect novice drivers without driving experience and if task complexity could still constitute a mediating factor.

Finally, it is unknown if the validity of the simulator setup used had been established prior to this study. It is a possibility that the simulator was unable to evoke normal driving responses and that other habituated responses (e.g internalized responses from a history of using simulators or a habit of playing computer games) were subjected to the social presence effects. The issue of validity is however partly approached in study two.

3.4 Conclusion

The goal of study one was to investigate if driving performance in a simulated environment would be affected by evaluative status and the mere presence of another individual. No significant interaction or main effects of mere presence or evaluation apprehension on the performance of participants driving in a simulated environment were found on any of the selected measures (reaction times, THW, max retardation, mean speed). The results of this study indicate that mere presence and evaluation apprehension are not factors affecting driving performance in a simulated environment to any significant level, lending no support to the hypothesis of the present study. The practical implications of this are discussed further in the general discussion of this report.

4 Study 2: Survey

The aim of the survey study was to gauge how participants perceive their on-road driving to certain elements present in everyday driving situations and traffic, and how these ratings compare to how they perceived their driving in the simulator on the same elements. This was done to study the participants' self-reported behavioral validity as well as their judgments about the overall fidelity of the simulation in order to gain a deeper understanding regarding the viability of driving simulators as tools for assessments. In particular looking at a subset of elements present in the Swedish driving test for eligibility B.

4.1 Method

4.1.1 Participants

The same 41 (23 male and 18 female) participants between the ages of 21 and 44 used in the simulator study were also asked to answer a survey before (survey one) and after (survey two) the simulated drive ($N = 41$).

4.1.2 Design

The two surveys were designed in accordance with the good questionnaire practices suggested by Shaughnessy, Zechmeister & Zechmeister (2012). Repeated statements were used in order to increase the reliability of respondents self-reported attitudes, and the repeated statements were worded in an opposite direction to combat response bias effects.

Survey one consisted of several items used to gauge the participants' prior driving experience, perceived driving skill and more specific statements in regards to how participants handle different traffic situations. The items were formulated as statements where the participant could answer on a likert scale (1-5) ranging from strongly disagree to strongly agree respectively. The items were used to create a baseline of values for which the results of survey two could be compared against. The items were focused around four categories in survey one and two: attention, planning, adaptive driving, and traffic rules with an additional fifth category in survey two; simulator fidelity. Survey two consisted of the simulator equivalent items from survey one as well as another eight statements about the fidelity of the simulated environment and instruments. Questions about age and gender were included at the end of survey two in order to avoid any potential effects of stereotype threat; as it has been shown that stereotype threat has had significant effects on simulated driving performance (Joanisse, Gagnon & Voloaca, 2013; Yeung & Von Hippel, 2008). All participants were presented with the same items in the same order. The survey was created and managed using Google Forms (https://www.google.com/forms/).

The statements included in both surveys related to how participants assessed their own ability to adapt their driving to elements in the surrounding environment such as: roadworks, pedestrians, other drivers, to what degree they estimated their compliance with traffic regulations, and their perceived level of focus/distraction whilst driving.

The statements included to assess the simulator fidelity in survey two were focused around how the participants reacted to and perceived certain elements of the environment such as: sense of speed, behavior of the simulated elements (pedestrians, other drivers), and the correspondence of the controls to the movement of the car (pedals, steering). As the physical setup used in this study can be considered relatively basic (one screen visual presentation, game-based peripherals, very limited physical feedback), these statements were included in order to estimate the level of perceived fidelity of the simulation. As mentioned earlier it has been shown that establishing simulator fidelity is of considerable importance if behaviors are to be translated from on-road driving to the simulated environment (Filtness et al., 2016; Campos et al., 2017). Furthermore, it has been shown that physical feedback is of importance in order to develop a realistic illusion of speed (Backlund, Engström, Johannesson & Lebram, 2010). The setup used in this study did not include advanced physical feedback therefore statements in regard to the participants' perception of speed were included in order to further assess fidelity of the setup. All statements and questions used in survey one and two are listed under Appendix 5.

4.1.3 Procedure

The two-part survey was issued to participants on a laptop computer before and after participating in study one. Before answering survey one, participants were required to read and approve a consent form. An experiment leader was then present for the duration of the survey and answered any questions the participants might have had. After they had answered survey one, participants took part in the experiment conducted for study one and following its completion finished the second part of the survey.

4.2 Results

The responses of all participants for survey one and two were compiled separately for further comparison ($N = 41$). Each aspect of the participants' judgments about the simulator was divided into separate categories and responses were compared between survey one and survey two. This was done for categories: attention, planning, adaptive driving, and traffic rules. Category five (simulator fidelity) was only present in survey two and therefore analysed separately. All statements and their scores have been converted to have the same directionality, e.g. a reverse-worded statement: "I do not agree with…" with a score of 1 has been converted to the directional opposite "I agree with…" with a score of 5.

For categories one through four a comparison was performed using a paired samples *t* test. Response ratings from the participants' self judgments in these categories were compared from survey one to survey two. In category one the participants' judgments of estimated levels of attention during on-road driving ($M = 3.86$, $SD = 0.53$) was compared with the same items after the simulated scenario ($M = 4.17$, $SD = 0.74$). This comparison revealed a significant change in rating, participants estimated that their attention to the task of driving was higher after the scenario as compared to before ($t(40) = -3.438$ $p = .001$). In category two the participants' judgments of estimated ability to plan their driving ahead of time and ability to anticipate certain situations before driving in the simulated scenario ($M = 3.59$, $SD = 0.83$) was compared with after the simulated scenario ($M = 3.51$, $SD = 0.92$). This comparison revealed an insignificant change in mean rating ($t(40) = -0.927$ $p = .360$). Category three revealed no significant changes between ratings on adaptive driving before ($M = 4.13$, $SD = 0.64$) compared with after ($M = 4.26$, $SD = 0.55$) ($t(40) = 0.400$ $p = .691$). Category four compared the participants' estimations of how well they believe themselves to follow traffic laws and regulations in on-road situations ($M = 4.46$, $SD = 0.74$) and in the simulated environment ($M = 4.59$, $SD = 0.59$) revealing no significant change ($t(40) = -1.188$ $p = .242$).

Responses for category five (simulator fidelity) were only gathered from survey two; as such the following results are reported using descriptive statistics. The rating with the highest number of participant responses is reported as well as the mean value for all responses in the category (1 = Strongly Disagree, 5 = Strongly Agree). The results are summarized in Figure 3.

Four statements regarding the participants' impressions of the pedals and steering wheel were summed and the responses were as follows: 21 respondents (51.2%) rated the realism of the pedals as a 2; one point short of strongly disagreeing ($M = 2.44$), 15 respondents (36.6%) rated the correspondence between the pedals and the car as a 2; one point short of strongly disagreeing with that the sensitivity of the pedals corresponded with the response of the simulated car ($M = 2.85$). Relating to this, 17 respondents (41.5%) also placed their rating one point short of strongly disagreeing with that braking in the simulator felt realistic compared to braking in a real car ($M = 2.45$). On the statements pertaining to the participants' judgments about the steering of the car, 19 participants (46.3%) rated the correspondence of movement between the steering wheel and car as 3; the midpoint of the scale ($M = 2.76$). On judgments regarding the realism of the steering wheel 20 individuals (48.8%) rated the realism a 2; one point short of strongly disagreeing ($M = 2.41$). Indicating that overall individuals felt that the correspondence of the steering was adequate whereas the correspondence of the pedals and the realism of both pedals and steering was lacking.

Furthermore, three statements in regards to participants' impressions of the other simulated elements (pedestrians and other cars) were surveyed as well as the overall realism of the simulated environment. On the statement pertaining to if participants perceived the pedestrians in the simulation to behave as expected 13 respondents (31.7%) rated this as a 2; one point short of strongly disagreeing, whilst the same number of respondents rated this as a 4; one point short of strongly agreeing ($M = 3.10$). On the statement pertaining to if participants perceived the other cars to behave as expected 20 respondents (48.8%) placed their rating as a 4; point short of strongly agreeing ($M = 3.41$). When asked about the general realism of the simulator the majority of respondents (41.7%) rated this a 3; the middle point of the scale ($M = 2.95$). These results together indicate that there could be room for improvement regarding the realism of the simulator, especially with the simulated pedestrians.

When asked about the perception of speed a majority (70.7%) reported that it was difficult to judge their current speed ($M = 1.39$). Furthermore, when asked if it was easy to stay within the speed limit 13 participants (31.7%) reported a 2; one point short of strongly disagree (*M* = 2.39). Taken together these two items indicate that the simulator failed to create a valid illusion of speed.

Finally on the statement pertaining to if participants behaved in the simulator as they would have in real life, a majority of respondents (53.7%) placed their rating one point short of strongly agreeing (*M* = 3.66). Indicating that even though the simulation seemingly has shortcomings the self reported behavior of the participants within the simulated environment has some correspondence with their own estimated behavior in on-road situations.

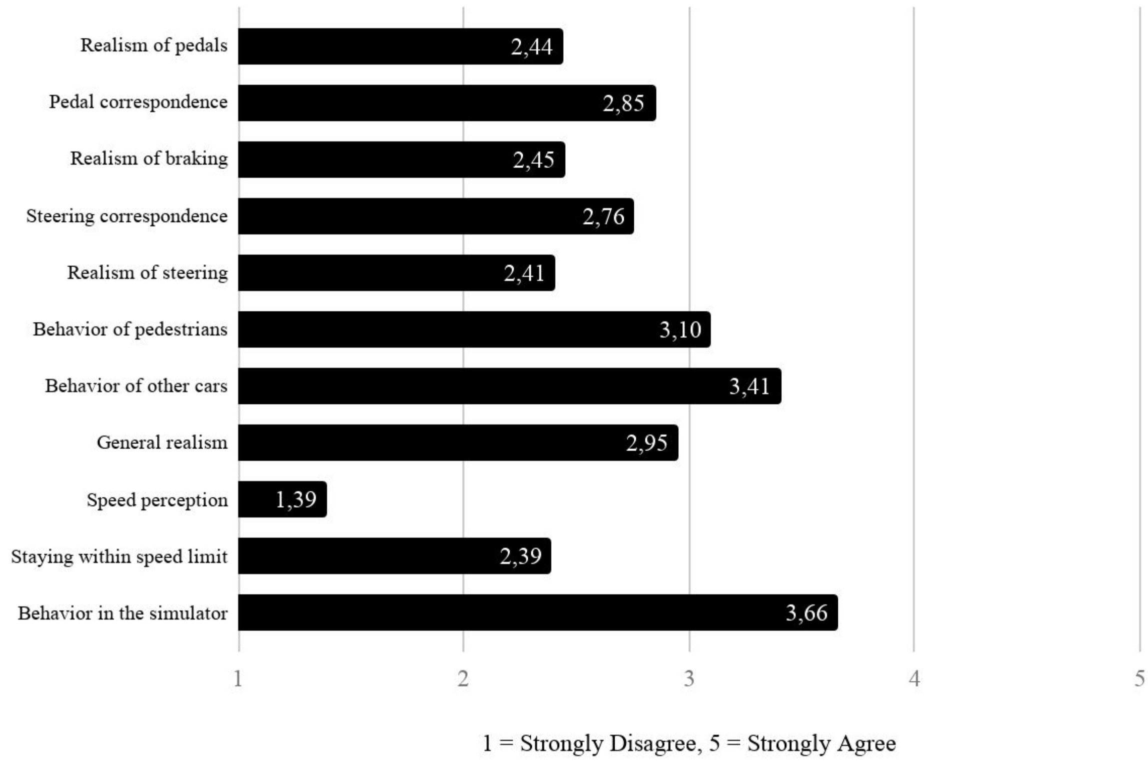

Figure 3. Mean ratings for simulator fidelity.

4.3 Discussion

The results from the survey are discussed primarily as an indicator of how certain behavioral elements, as would be found in the current Swedish driving test, translate from on-road driving to simulated driving and how these elements are approached by participants. As Reimer et al. (2006) showed, self-reported measures can correlate with objective measures from a simulated environment. Furthermore, self-reports can be used to get insights into on-road driving behaviors (Ben-Ari, Hager & Prato, 2016), supporting the validity of the self-reported measures.

From the survey it was found that the participants reported a significant difference in their levels of focus and attention between on-road driving and driving in the simulator. This was reported as a lower level of focus in relation to on-road driving and a higher level of focus during the simulated drive. These measures could be an indication that the simulation provided an environment that was less distracting than how the participants perceive on-road driving. If it is the case that the simulated environment provides a lower level of distraction as compared to on-road driving this could be beneficial when conducting assessments. A lower baseline level of distraction together with the fact that simulated environments constitute a more easily controlled setting could provide a reasonable platform to be used for manipulating the level of distraction to suit a wide range of different on-road situations. However, it is important to note that this result is merely an indication as the difference between conditions were not fully controlled. To confidently state a behavioral relationship between any of the survey categories would require further studies. There is an obvious issue that the two constructs (on-road versus simulated driving) the surveys aimed to measure were separated in time. Impressions of the simulator were assessed only a moment after the subjects completed the simulated driving scenario, and were thus assessed closer in time than any impressions towards real-world driving events. Due to recency effects, recently encountered events are typically easier to recall than more distant ones and impressions related to more distant events could in accordance with the availability heuristic be based on a subset of more easily retrieved situations (Hardman, 2009). This casts doubt on the validity of the comparison of the two constructs. To ensure that the results are representative it would be beneficial to control for this temporal factor.

In regards to the insignificant differences in ratings of planning ability, ability to adapt, and adherence to traffic rules; the results indicate that the participants seemingly treat all these categories the same in on-road situations as they did in the simulated environment. These findings do not rule out that some elements included in the Swedish class B driving test could be assessed using a simulator protocol (traffic behaviors towards unprotected pedestrians, zebra crossings, signal controlled crossings or effective braking). Furthermore, the results give positive indications of at least relative behavioral validity between the simulator setup and on-road driving (also supporting the findings of study one). The fact that a majority of participants reported that their behavior in the simulator corresponded to how they would have acted in real life provides further support for this interpretation of the results.

The low ratings of simulator realism and fidelity questions the use of the current simulator setup to confidently be able to assess all procedural skills as it does not fully emulate all aspects of on road driving. Therefore, with driving assessments in mind it would be of interest to continue to develop the understanding related to transference of on-road driving skill to simulated environments and what specific criterion of fidelity need to be met in order to achieve sufficient behavioral validity.

As with all survey research it is also possible that participants may have responded in a socially desirable way (Shaughnessy et al., 2012). Ratings of traffic law obedience is a measure that could be particularly vulnerable for social desirability bias. Previous findings indicate however that any social desirability effects could be expected to be relatively small (Lajunen & Summala, 2003). Reverse-worded statements were also used in both surveys as to get more accurate self-report measures and to counter other response bias effects. It is worth noting however that this is a questioned practice since it is also thought to increase the risk of confusing respondents and that a lack of attention could make participants inattentive of the reversals (van Sonderen, Sanderman & Coyne, 2013).

4.4 Conclusion

The aim of study two was to gather responses on self-reported driver behavior and the ability of the simulator to emulate the driving experience. The results indicate that the participants' impressions of the simulator fidelity were relatively low. However, the responses do indicate that the participants' impressions of driving in the simulator are similar to their self-reported recollections of on-road driving. This lends support to the viability of assessing at least certain elements of the Swedish driving test for a driver's license with eligibility B in a simulated environment.

5 General Discussion

The possibility of performing driving assessments in a simulator-based environment can entail many advantages over the current resource-heavy on-road assessment system. Competency assessments could be performed in a more secure, standardized and controlled setting, providing an objectively reliable means to evaluate novice drivers. Furthermore, simulated environments allow for both more efficient and detailed assessments as variable driving conditions can be included in a way not currently possible. One of the main purposes of this report was to investigate the use of driving simulators and simulated environments as a tool for assessing driving competence. Are simulator-based driving conditions and on-road driving conditions sufficiently similar for this to be practically viable and equally important is driver behavior sufficiently translated between the two conditions?

Previous findings have shown that physiological measures could correlate between simulated and on-road conditions (Mullen et al., 2011). This indicates that the experimental findings are at least theoretically transferable between simulated and on-road assessments. This claim can be more confidently stated as self-reports assessing behavioral validity of the simulator setup showed that the participants responded relatively strongly that their behavior for the simulated drive was consistent with their previous on-road behavior. The simulator fidelity was however assessed as being quite low, questioning the use of the current setup to assess all procedural skills associated with driving.

The effects of mere presence and evaluation apprehension were hypothesized to be factors that potentially could be of importance when transfering assessments from an on-road situation to a simulated one. However, as the results of study one did not find any meaningful differences in driving performance due to these factors; it is concluded that mere presence and evaluation apprehension are factors that seemingly have negligible if any effects on driving performance as measured by the dependant variables chosen for this study. From the perspective of mere presence and evaluation apprehension, it is concluded that the presence of another person or evaluative instructions during simulated assessments are seemingly superfluous and therefore not factors that need to be accounted for. However, the inclusion of an individual actively and noticeably evaluating the participant could still have a compounding effect. The role of the person present and the type of situation could as previously noted have arousal increasing effects that affects performance (Mullen et al., 1997). Further, the norms that the person represents could also potentially affect performance (Baxter et al., 1990). A closer resemblance between the testing environment and an actual assessment situation could also elicit more typical responses from participants (Guerin, 1993). A recommendation for future studies is therefore to approach the question from this broader perspective, as it cannot be ruled out that the presence of an active evaluator has additive consequences for individuals taking a driver's test.

A concluding remark is that the findings of this report lends tentative support to the viability of assessing certain elements from the Swedish class B driving test in a simulated environment.

6 References

Allport, F. H. (1920). The influence of the group upon association and thought. *Journal of Experimental Psychology, 3*(3), 159-182. http://dx.doi.org/10.1037/h0067891

Backlund, P., Engström, H., Johannesson, M., & Lebram, M. (2010). Games for traffic education: An experimental study of a game-based driving simulator. *Simulation & Gaming, 41*(2), 145-169. https://doi.org/10.1177/1046878107311455

Bates, J, L., Filtness, A., & Watson, B. (2018). Driver Education and Licensing Programs. In Lord, D., & Washington, S. (Eds.)*, Safe Mobility: Challenges, Methodology and Solutions* (p. 13-36). Bingley: Emerald Publishing Limited.

Baxter, J., Manstead, A., Stradling, S., Campbell, K., Reason, J., & Parker, D. (1990). Social facilitation and driver behaviour. *British Journal of Psychology, 81*(3), 351-360. https://doi.org/10.1111/j.2044-8295.1990.tb02366.x

Belletier, C., Normand, A., & Huguet, P. (2019). Social-Facilitation-and-Impairment Effects: From Motivation to Cognition and the Social Brain. *Current Directions in Psychological Science*. https://doi.org/10.1177/0963721419829699

Bédard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of Driving Performance Using a Simulator Protocol: Validity and Reproducibility. *American Journal of Occupational Therapy 64*(2), 336-340. https://dx.doi.org/10.5014/ajot.64.2.336

Ben-Ari, O.T., Hager, A.E., & Prato, C.G. (2016). The value of self-report measures as indicators of driving behaviors among young drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 39*, 33-42. https://doi.org/10.1016/j.trf.2016.03.005

Bond, C. F. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology, 42*(6), 1042-1050. http://dx.doi.org/10.1037/0022-3514.42.6.1042

Campos, J. L., Bédard, M., Classen, S., Delparte, J. J., Hebert, D. A., Hyde, N., Law, G., Naglie, G., Yung, S. (2017). Guiding Framework for Driver Assessment Using Driving Simulators. *Frontiers in psychology, 8*, 1428. https://dx.doi.org/10.3389/fpsyg.2017.01428

Casutt, G., Theill, N., Martin, M., Keller, M., & Jäncke, L. (2014). The drive-wise project: Driving simulator training increases real driving performance in healthy older drivers. *Frontiers in Aging Neuroscience, 6*(MAY), 85. https://doi.org/10.3389/fnagi.2014.00085

Cottrell, N., Wack, D., Sekerak, G., & Rittle, R. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology, 9*(3), 245-250. http://dx.doi.org/10.1037/h0025902

Damm, L., Nachtergaële, C., Meskali, M., & Berthelon, C. (2011). The Evaluation of Traditional and Early Driver Training With Simulated Accident Scenarios. *Human Factors, 53*(4), 323–337. https://doi.org/10.1177/0018720811413765

De Winter, J. C.F., De Groot, 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. https://doi.org/10.1080/00140130802277521

Fambro, D. B., Koppa, R. J., Picha, D. L., & Fitzpatrick, K. (1998). Driver Perception–Brake Response in Stopping Sight Distance Situations. *Transportation Research Record, 1628*(1), 1–7. https://doi.org/10.3141/1628-01

Feest, J. (1968). Compliance with Legal Regulations: Observation of Stop Sign Behavior. *Law & Society Review, 2*(3), 447-461. DOI: 10.2307/3052898

Filtness, A., Tones, M., Bates, L., Watson, B., & Williamson, A. (2016). *How would changing driver training in the Queensland licensing system affect road safety?* (Deliverable 2: Simulators for skill acquisition training and assessment, and their impact on road safety). Queensland: The Centre for Accident Research & Road Safety.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism, 10*(2), 486-9. https://dx.doi.org/10.5812%2Fijem.3505

Guerin, B. & Innes, J.M. (1984). Explanations of social facilitation: A review. *Current Psychological Research & Reviews, 3*(2). https://doi.org/10.1007/BF02686548

Guerin, B. (1986). Mere presence effects in humans: A review. *Journal of Experimental Social Psychology, 22*(1), 38-77. http://dx.doi.org/10.1016/0022-1031(86)90040-5

Guerin, B. (1993). *European monographs in social psychology. Social facilitation.* New York, NY, US: Cambridge University Press; Paris, France: Editions de la Maison des Sciences de l'Homme. http://dx.doi.org/10.1017/CBO9780511628214

Hardman, D. (2009). *Judgment and decision making: Psychological perspectives.* Malden: Blackwell Publishing; Leicester, England: British Psychological Society.

Harkins, S. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology, 23*(1), 1-18. http://dx.doi.org/10.1016/0022-1031(87)90022-9

Harkins, G.S. (2006). Mere Effort as the Mediator of the Evaluation–Performance Relationship. *Journal of Personality and Social Psychology, 91*(3), 436-455. http://dx.doi.org/10.1037/0022-3514.91.3.436

Herman, C. (2015). The social facilitation of eating. A review. *Appetite, 86*, 61-73. https://doi.org/10.1016/j.appet.2014.09.016

Ekanayake, H.B, Backlund, P., Ziemke, T., Ramberg, R., Hewagamage, K.P & Lebram, M., (2013). Comparing Expert Driving Behavior in Real World and Simulator Contexts. *International Journal of Computer Games Technology, 2013*(2013), 87-100. http://dx.doi.org/10.1155/2013/891431

Huguet, P., Barbet, I., Belletier, C., Monteil, J., Fagot, J., & Gauthier, Isabel. (2014). Cognitive Control Under Social Influence in Baboons. *Journal of Experimental Psychology: General, 143*(6), 2067-2073. https://doi.org/10.1037/xge0000026

Joanisse, M., Gagnon, S., & Voloaca, M. (2013). The impact of Stereotype Threat on the simulated driving performance of older drivers. *Accident Analysis and Prevention, 50*, 530-538. https://doi.org/10.1016/j.aap.2012.05.032

Kappé, B., de Penning, L., Marsman, M., & Roelofs, E. (2012). Assessment in driving simulators: Where we are and where we go. In *PROCEEDINGS of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 183-189). Big Sky MT, United States: University of Iowa.

Klauer, K. C., Herfordt, J., & Voss, A. (2008). Social presence effects on the stroop task: Boundary conditions and an alternative account. *Journal of Experimental Social Psychology, 44*(2), 469-476. https://doi.org/10.1016/j.jesp.2007.02.009

Klopfer, P. (1958). Influence of social interactions on learning rates in birds. *Science (New York, N.Y.), 128*(3329), 903. DOI: 10.1126/science.128.3329.903

Lawshe, C. (1940). Studies in Automobile Speed on the Highway. *Journal of Applied Psychology, 24*, 297. http://dx.doi.org/10.1037/h0055561

Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(2), 97-107. https://doi.org/10.1016/S1369-8478(03)00008-1

Lee, C., & Abdel-Aty, M. (2008). Presence of passengers: Does it increase or reduce driver's crash potential? *Accident Analysis and Prevention, 40*(5), 1703-1712. https://doi.org/10.1016/j.aap.2008.06.006

Lee, H.C., Cameron, D., & Lee, A.H. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis and Prevention, 35*(5), 797-803. https://doi.org/10.1016/S0001-4575(02)00083-0

Mullen, B., Bryant, B., & Driskell, J. E. (1997). Presence of others and arousal: An integration. *Group Dynamics: Theory, Research, and Practice, 1*(1), 52-64. http://dx.doi.org/10.1037/1089-2699.1.1.52

Mullen, N., Charlton, J., Devlin, A., & Bédard, M. (2011). Simulator Validity: Behaviors observed on the simulator and on the road. In Fisher, D., Rizzo, M., Caird, J., Lee, J.D., (Eds.), *Handbook of Driving Simulation for Engineering, Medicine, and Psychology* (p. 13-1-13-15). Boca Raton, FL, USA: CRC Press

Muller, D., & Butera, F. (2007). The focusing effect of self-evaluation threat in coaction and social comparison. *Journal of Personality and Social Psychology, 93*(2), 194-211. https://doi.org/10.1037/0022-3514.93.2.194

Reimer, B., D'Ambrosio, L.,A., Coughlin, J. F., Kafrissen, M. E., & Biederman, J. (2006). Using self-reported data to assess the validity of driving simulation data. *Behavior*

Research Methods, 38(2), 314-24. Retrieved from https://search-proquest-com.ezproxy.ub.gu.se/docview/204304707?accountid=11162

Roenker, D., Cissell, G., Ball, K., Wadley, V., & Edwards, J. (2003). Speed-of-Processing and Driving Simulator Training Result in Improved Driving Performance. *Human Factors: The Journal of Human Factors and Ergonomics Society, 45*(2), 218-233. https://doi.org/10.1518/hfes.45.2.218.27241

Rosenbloom, T., Shahar, A., Perlman, A., Estreich, D., & Kirzner, E. (2007). Success on a practical driver's license test with and without the presence of another testee. *Accident Analysis and Prevention, 39*(6), 1296-1301. https://doi.org/10.1016/j.aap.2007.03.015

Seitchik, A., Brown, A., & Harkins, S. (2017). Social Facilitation: Using the Molecular to Inform the Molar. In Harkins, G, S., Williams, D, K., & Burger, J., (Eds.), *The Oxford Handbook of Social Influence*. Oxford: Oxford University Press. Retrieved 9 Apr. 2019, Retrieved from http://www.oxfordhandbooks.com.ezproxy.ub.gu.se/view/10.1093/oxfordhb/9780199859870.001.0001/oxfordhb-9780199859870-e-8

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin & Review, 17*(1), 52-58. https://doi.org/10.3758/PBR.17.1.52

Shaughnessy, J. J., Zechmeister, B. E., & Zechmeister, S. J. (2012). *Research Methods in Psychology* (9th ed.). New York: McGraw-Hill

Spence, K. W. (1958). A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *American Psychologist, 13*(4), 131-141. http://dx.doi.org/10.1037/h0045054

Swedish National Road and Transport Research Institute. (n.d.). *Simulator technology at VTI.* Retrieved 2019-05-06, from https://www.vti.se/en/research-areas/driving-simulation1/simulator-technology/

Trafikverket. (2018) *Körkort för personbil behörighet B*. Retrieved 2019-05-02, from https://trafikverket.ineko.se/Files/sv-SE/17594/Ineko.Product.RelatedFiles/100820_korkort_for_personbil_behorighet_b_.pdf

Transportstyrelsen. (2012) *Transportstyrelsens föreskrifter om förarprov, behörighet B; (konsoliderad elektronisk utgåva).* Retrieved 2019-05-02, from https://www.transportstyrelsen.se/TSFS/TSFS%202012_43k.pdf

Triplett, N. (1898). The Dynamogenic Factors in Pacemaking and Competition. *The American Journal of Psychology, 9(4),* 507-533. http://dx.doi.org/10.2307/1412188

Törnros, J. (1998). Driving behaviour in real and simulated road tunnel validation study. *Accident Analysis & Prevention 30*(4), 497-503. http://dx.doi.org/10.1016/S0001-4575(97)00099-7

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality, 41*(3), 579-601. https://doi.org/10.1016/j.jrp.2006.06.008

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain. *PloS one*, *8*(7), e68967. https://doi.org/10.1371/journal.pone.0068967

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. *Accident Analysis and Prevention, 35*(3), 427-433. https://doi.org/10.1016/S0001-4575(02)00022-2

Yeung N.C., & Von Hippel, C. (2008). Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. *Accident Analysis and Prevention, 40*(2), 667-674. https://doi.org/10.1016/j.aap.2007.09.003

Zajonc, R. B. (1965). Social facilitation. *Science, 149*(3681), 269-274. http://dx.doi.org/10.1126/science.149.3681.269

Zajonc, R., Heingartner, A., Herman, E., & Mcguire, William J. (1969). Social enhancement and impairment of performance in the cockroach. *Journal of Personality and Social Psychology, 13*(2), 83-92. http://dx.doi.org/10.1037/h0028063

7 Appendices

7.1 Appendix 1 - Study 1 & 2: Consent Form

Consent to Participate in Research Study:

The purpose of the study is to investigate the viability of using a simulator to assess driving ability.

The study will take about 20 minutes, and you will be asked to:
1) Fill out one questionnaire before driving (this), and fill out one questionnaire after driving.
2) Drive through a short simulated city scenario.

A requirement for participating in this study is that you have a driver's license.

The decision to participate in this study is entirely up to you, and you may cancel your participation at any time without penalty. This study is anonymous. We will not be collecting or retaining any information about your identity. We will not include any information in any report we may publish that would make it possible to identify you. You will receive a cinema voucher for participating.

The study is to be included in a bachelor thesis by students at Gothenburg University. For further questions about this study ask an experiment leader directly.

7.2 Appendix 2 - Study 1: Translated Participant Instructions

(Originally presented in Swedish, Instructions are shown for groups 1 and 4 as the only difference between instructions consisted of the statements pertaining to evaluation.)

Group 1 (no presence, no evaluation)

You are now going to drive in a simulated environment for roughly 12 minutes. The car is automatic so you only need to steer, throttle up and brake. Upon completion of the drive you will be asked a few questions about your experience.

All you need to do is follow the road straight ahead, you will be notified on the screen when the scenario is over and when to stop. Drive as you normally would.

We will be waiting in an office downstairs so we won't be able to see when the scenario is over, please come and collect us when the scenario is over.

Your driving will not be evaluated.

Group 4 (Presence, Evaluated)

You are now going to drive in a simulated environment for roughly 12 minutes. The car is automatic so you only need to steer, throttle up and brake. Upon completion of the drive you will be asked a few questions about your experience.

All you need to do is follow the road straight ahead, you will be notified on the screen when the scenario is over and when to stop. Drive as you normally would.

The simulator will be collecting information during your drive which will be compared against the criterion for the Swedish driving test B, this comparison will be done to see how you would have performed on an examination today.

We will be waiting in an office downstairs so we won't be able to see when the scenario is over, please come and collect us when the scenario is over.

Keep in mind that this will be assessed.

7.3 Appendix 3 - Study 1: Descriptive Statistics

Table 1

Mean and standard deviations for all dependant variables across all four conditions

Mere Presence	Evaluative Status	THW (s)		Reaction Times (ms)		Max Retardation (m/s²)		Mean Speed (m/s)		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	n
No presence	No eval.	3.63	1.13	586.73	117.45	-5.21	1.10	11.83	1.85	10
No presence	Evaluation	3.77	0.86	548.89	83.96	-4.67	1.41	12.05	2.26	10
Presence	No eval.	3.20	0.84	604.97	157.77	-4.96	1.35	11.74	1.71	11
Presence	Evaluation	3.58	0.72	616.28	90.94	-4.90	1.04	11.97	2.00	10

7.4 Appendix 4 - Study 1: Analysis of Variance

Table 2

Analysis of variance for main and interaction effects on the four variables

Variable		*df*	*F*	η_p^2	*p*
THW	Mere Presence	1	1.202	.031	.280
	Evaluative Status	1	0.851	.022	.362
	Mere Presence x Evaluative Status	1	0.173	.005	.680
Reaction Time	Mere Presence	1	1.355	.036	.252
	Evaluative Status	1	0.131	.004	.720
	Mere Presence x Evaluative Status	1	0.446	.012	.508
Max Retardation	Mere Presence	1	0.001	.000	.975
	Evaluative Status	1	0.604	.016	.442
	Mere Presence x Evaluative Status	1	0.405	.011	.529
Mean Speed	Mere Presence	1	0.020	.001	.887
	Evaluative Status	1	0.136	.004	.715
	Mere Presence x Evaluative Status	1	0.00004	.000001	.995

Note. Nothing significant at the $p < 0.05$ level.

7.5 Appendix 5 - Study 2: Translated Survey Questions

(Originally presented in Swedish)

Survey 1

Background information
What year did you take your driver's license?
On average, how frequently do you drive?
How confident are you in your driving ability?

Category One - Attention
When driving a car I am fully focused on the task of driving.
I have a hard time concentrating on the task when driving a car.
I get easily distracted when driving in a city environment.

Category Two - Planning
I find it easy to plan my actions ahead of time when driving.
I frequently find myself braking hard in reaction to unforeseen events.

Category Three - Adaptive Driving
When driving a car I adapt my driving to nearby pedestrians.
I am aware of other drivers and adapt accordingly when driving a car
When driving a car I am aware of construction sites and adapt accordingly.

Category Four - Traffic Laws
When driving When driving a car I obey traffic laws to the best of my ability.

Survey 2

Category One - Attention
When driving in the simulation I was fully focused on the task of driving.
I had a hard time concentrating on the task when driving in the simulator.
I was easily distracted by the simulated city environment.

Category Two - Planning
I found it easy to plan my actions ahead in the simulated environment.
I frequently had to brake hard to unforeseen events when driving in the simulation.

Category Three - Adaptive Driving
When driving in the simulation I adapted my driving to nearby pedestrians.
I was aware of other drivers and adapted accordingly when driving in the simulation.
When driving in the simulation I noticed the construction sites and adapted accordingly.

Category 4 - Traffic Laws
When driving in the simulation i obeyed traffic laws to the best of my ability.

Category 5 - Simulator Fidelity

Peripherals
The movement of the steering wheel was consistent with the movement of the car.
Compared to a real car the steering did not feel realistic.
The sensitivity of the pedals was consistent with the movement of the car.
The pedals did not feel realistic compared to those in a real car.
Braking in the simulator felt realistic.

Speed Perception
It was difficult to perceive how fast i was driving.
I found it easy to drive according to the speed limit.

Behavioral components
During the simulation I acted as I would have in real life.
The pedestrians in the simulation acted as I expected.
The other cars in the simulation acted as I expected.
Overall, driving in the simulator felt realistic.

Additional Questions
What is your age?
What is your gender?
Did you experience any discomfort while driving the simulator (such as nausea)?